



Approches multivaluées et supervisées en morphologie mathématique et applications en analyse d'image

Sébastien Lefèvre

► To cite this version:

Sébastien Lefèvre. Approches multivaluées et supervisées en morphologie mathématique et applications en analyse d'image. Interface homme-machine [cs.HC]. Université de Strasbourg, 2009. tel-00515901

HAL Id: tel-00515901

<https://theses.hal.science/tel-00515901>

Submitted on 8 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approches multivaluées et supervisées
en morphologie mathématique
et applications en analyse d'image

Mémoire présenté par

Sébastien Lefèvre

le 9 décembre 2009

en vue de l'obtention de

l'Habilitation à diriger des recherches

devant le jury composé de :

Président :

Patrick Gros Directeur de recherche, INRIA Rennes Bretagne Atlantique

Rapporteurs :

Rémy Mullot Professeur, université de La Rochelle
Philippe Salembier Professeur, université polytechnique de Catalogne (Espagne)
Pierre Soille Chercheur, centre commun de recherche – JRC (Italie)

Examineurs :

Hans Burkhardt Professeur, université Albert-Ludwig de Freiburg (Allemagne)
Christian Ronse Professeur, université de Strasbourg
Nicole Vincent Professeur, université Paris Descartes

Remerciements

Dans le monde de la Recherche, il n'est pas suffisant d'avoir des idées et de s'assurer de leur bien-fondé : recevoir la reconnaissance de ses pairs est également important, puisque cela permet de s'assurer de l'intérêt des travaux scientifiques réalisés, et de motiver les travaux futurs. Aussi je tiens à remercier les membres du jury pour avoir bien voulu accepter de consacrer une partie de leur temps à l'étude de mes travaux et de parcourir tant de kilomètres pour venir à ma soutenance : Pierre Soille, Philippe Salembier, et Rémy Mullot, qui ont accepté la tâche de rapporteur, mais aussi Hans Burkhardt, Patrick Gros, et Nicole Vincent qui ont accepté de participer au jury.

Pour mener à bien des travaux de recherche, il faut disposer d'un environnement de qualité. Au sein du LSIIT, Fabrice Heitz (directeur du laboratoire), Ernest Hirsch (responsable de l'équipe MIV), et Christian Ronse (responsable du thème "morphologie mathématique") m'ont permis de disposer d'un tel environnement. Une aide précieuse m'a été apportée au quotidien par les personnels administratifs et techniques, notamment Sylvie, Jean-Marc et Stéphane. La fonction d'enseignant-chercheur n'est bien sûr pas limitée aux travaux de recherche, et j'ai eu la chance de bénéficier d'un accueil particulièrement chaleureux au Département d'Informatique et à l'UFR de Mathématique et d'Informatique. Je tiens ici à remercier Catherine Mongenet et Pascal Schreck (respectivement ex-directrice de l'UFR et directeur du département) pour leur confiance et leur soutien, ainsi que Marie-Claire "hello de lu", Guigui et Christophe, Sandrine, Saïd, et Claudine pour leurs efficacités administratives et techniques ainsi que pour leur gentillesse. En complément de cet environnement strasbourgeois, je suis profondément reconnaissant envers Nicole Vincent et Florence Sèdes, qui n'ont cessé de m'apporter de précieux conseils et m'ont garanti un soutien inaltérable, malgré la distance.

HDR signifie "Habilitation à Diriger des Recherches". Je remercie donc ceux qui m'ont fait confiance pour guider (ce terme me semble plus fidèle à la réalité) leurs travaux de recherche, et notamment : Erchan, Sébastien, Benjamin et Jonathan. Ils m'ont fait découvrir avec plus ou moins de joie toutes les facettes du rôle d'encadrant, pas toujours liées à la Science mais toujours fortement chronophages ! Florent et Régis ont également travaillé sous ma responsabilité sur des sujets plus applicatifs et je les remercie pour leur investissement.

J'ai eu la chance de "m'échapper" de mon laboratoire (et de France) pour quelques jours ou quelques mois et d'être systématiquement accueilli dans d'excellentes conditions. Je suis donc très reconnaissant à Martino Pesaresi du JRC à Ispra, et à Hans Burkhardt de l'Université Albert-Ludwig à Freiburg-im-Breisgau, pour la chance qu'ils m'ont offerte de connaître des environnements complètement différents des habitudes strasbourgeoises et françaises. Les échanges avec eux et leurs équipes (Elodie à Ispra, et tous les membres ou ex-membres de l'équipe LMB que j'ai côtoyés à Freiburg : Margret, Alex, Olaf, Thorsten, Marco, Janis, et bien d'autres) m'ont beaucoup apporté sur les plans professionnel et personnel, et les collaborations initiées lors de mes séjours sont sources de perspectives particulièrement intéressantes. Je pense qu'il en sera de même avec ma nouvelle terre d'accueil, et je remercie Patrick Gros et l'équipe TEXMEX de l'IRISA pour la confiance qu'ils m'ont accordée et leur chaleureux accueil à Rennes.

Les collaborations scientifiques dans lesquelles j'ai été impliqué, qu'elles soient ou non pluridisciplinaires, ont été très enrichissantes, en termes d'applications de mes travaux fondamentaux, d'ouverture à d'autres domaines, ou encore de contacts humains. Pour n'en citer que quelques uns, je tiens à remercier Anne Puissant, Annett Wania, David Sheeren, et Olivier Lézoray. Plus localement, la qualité scientifique et humaine des échanges avec différents membres du laboratoire justifie d'ajouter quelques noms à cette liste, mais il m'est bien impossible d'être exhaustif ! Citons néanmoins et par ordre alphabétique : Alain, Alex(s), Caro, Cédric, Christophe, Dominique, Germain, Guillaume, Pascal(s), Pierre(s), Philippe, Vincent(s), et bien d'autres. Sans oublier Stéphane, bien sûr, bien plus qu'un simple collègue et compagnon sur les routes (dont celle de l'habilitation)...

Enfin, ces années passées en Alsace à s'investir dans l'enseignement et la recherche m'ont surtout apporté l'essentiel... Merci à toi, Séverine, pour ton soutien sans faille malgré l'horizon pas toujours dégagé, et pour tes plus belles réussites : Elsa, Rémi, Adam et Alice. Vous m'avez tous les cinq offert à la fois la motivation et les à-côté indispensables à l'aboutissement de ces travaux, et je vous dédie ce mémoire.

Préambule

Ce document présente une synthèse des activités de recherche menées au sein du Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (UMR 7005 CNRS–UDS¹) depuis mon recrutement comme Maître de Conférences en Informatique en 2003. Ces travaux sont le fruit d'une reconversion thématique dans le domaine de la morphologie mathématique, complètement effective depuis 2004–2005. Ce document est organisé en quatre chapitres et comporte en annexe certaines des publications jugées représentatives de ces activités de recherche (la table des matières est fournie pour sa part à la fin du document).

Contexte

Le premier chapitre a pour but de rappeler le contexte dans lequel se sont déroulés ces travaux de recherche. J'y illustre notamment ma reconversion thématique en décrivant mes travaux de thèse sur l'analyse et l'indexation de séquences vidéo, qui ont été menés dans le cadre d'une collaboration industrielle et ont donc été fortement guidés par des considérations applicatives (section 1.1).

Au contraire, ce sont des questions fondamentales qui ont initialement motivé les travaux relatifs à la morphologie mathématique que j'ai réalisés à Strasbourg (seul ou en collaboration avec des jeunes, ou moins jeunes, chercheurs). Comment prendre en compte des données multivaluées en morphologie ? Comment intégrer des connaissances dans les approches morphologiques ? Ces questions fondamentales ont néanmoins débouché sur la résolution de problèmes applicatifs dans différents domaines : imagerie couleur (pour la recherche par le contenu), télédétection, et imagerie astronomique (section 1.2).

Ce chapitre est complété par un curriculum vitae, (section 1.3), une description de mes activités d'enseignement et responsabilités pédagogiques (section 1.4), et de mes diverses activités concernant l'animation de la recherche (section 1.5) : j'y décris notamment mes activités en termes d'encadrement de jeunes chercheurs, d'insertion dans la communauté scientifique, et de gestion de projets et contrats de recherche. Je termine ce chapitre par la liste de mes publications (section 1.6).

Travaux fondamentaux

Le second chapitre présente nos² contributions en morphologie mathématique au travers de leurs aspects fondamentaux.

Une courte introduction à la morphologie mathématique est tout d'abord fournie (section 2.1), dans l'objectif de rappeler les bases nécessaires à une bonne compréhension de la suite du manuscrit, des problématiques abordées comme des solutions proposées, et d'introduire les notations utilisées par la suite.

La section 2.2 traite du problème de la prise en compte d'images multivaluées en morphologie mathématique. De telles images peuvent être rencontrées sous différentes formes (images en couleurs, multispectrales, multitemporelles, etc.), et dans de nombreux domaines. Cependant leur traitement par les approches morphologiques est loin d'être trivial, du fait notamment de l'absence d'ordre vectoriel universel. Nous discuterons donc les principaux ordres vectoriels proposés dans la littérature, et nous nous intéresserons plus particulièrement

¹UDS signifie ici Université de Strasbourg. Avant le 01/01/2009, il s'agissait de l'ULP – Université Louis Pasteur, mais pour des raisons de simplicité la distinction ULP/UDS sera omise dans ce document.

²En général, le pronom personnel de la première personne sera utilisé au singulier lorsque sont concernés des travaux où seul l'auteur est impliqué, et au pluriel lorsqu'ils résultent de collaborations, ou d'encadrement de jeunes chercheurs, et qu'ils sont donc le fruit d'un travail collectif.

aux ordres totaux tels que l'ordre lexicographique, pour lequel nous avons proposé différentes variantes équilibrées. Nous étudierons également d'autres solutions ne faisant pas explicitement appel à une relation d'ordre mais procédant par décomposition en images binaires ou à niveaux de gris.

Indépendamment de la nature des images à traiter (qu'elles soient mono ou multivaluées), il est souvent capital d'exploiter au mieux les connaissances disponibles (le plus souvent fournies par un expert ou utilisateur) sur le problème à traiter et les images considérées pour aboutir à des résultats pertinents. La section 2.3 porte donc sur l'intégration de connaissances en morphologie mathématique, et débute par une description des différents types de connaissances pouvant être introduites. Les connaissances étant généralement formulées au travers des éléments structurants (le principal paramètre des opérateurs morphologiques), nous discuterons l'apport des différents types d'élément structurant pour l'intégration de connaissance. Nous considérerons ensuite le cas d'une connaissance limitée au nombre de classes d'intérêt dans l'image, et montrerons comment celle-ci peut être prise en compte en morphologie à l'aide d'un processus de classification non-supervisée. Enfin, nous nous focaliserons sur l'utilisation de connaissances fournies sous la forme de classes de pixels ou d'objets, que nous exploiterons au travers d'une procédure de classification supervisée.

Applications

Dans le troisième chapitre, nous adoptons un autre point de vue pour décrire nos travaux en les présentant au travers des domaines d'application auxquels ils se rattachent.

La section 3.1 porte sur l'analyse d'images couleur, notamment dans un but de description et reconnaissance. Dans ce contexte, nous nous sommes intéressés au traitement morphologique d'images couleur et à la prise en compte de l'information de teinte. Nous avons par la suite abordé les problèmes classiques de détection d'objets, de segmentation en régions et de description d'image couleur à l'aide de solutions morphologiques adaptées. Nous présentons également dans cette section le système d'annotation et de recherche par le contenu que nous avons élaboré dans le cadre d'un projet soutenu par Oséo – ANVAR.

En télédétection (section 3.2), nous nous sommes focalisés sur deux environnements spécifiques que sont les zones urbaines et le littoral, en partenariat avec des géographes notamment, dans le cadre de projets ANR. Pour chacun de ces environnements, nous avons cherché à exploiter au mieux les connaissances de l'expert pour fournir des résultats pertinents en détection et segmentation d'objets géographiques.

L'imagerie astronomique fait l'objet de la section 3.3. Même si les problèmes traités ici sont relativement similaires à ceux abordés dans les sections précédentes, la nature des données introduit de nouvelles difficultés. En particulier, l'hétérogénéité des structures à identifier, et le niveau de bruit présent dans les images doivent être correctement pris en compte. Dans cette section, nous décrirons donc les solutions morphologiques mises en œuvre pour la segmentation, la détection et la classification de galaxies dans des images astronomiques.

La section 3.4 complète ce chapitre par la présentation de la plate-forme PELICAN (Polyvalent Extensible Library for Image Computing and ANalysis), fruit d'un travail non directement lié à la morphologie mathématique, mais qui me semble mériter d'être décrit ici. Ce projet de plate-forme pour l'analyse d'image a été initié localement en 2004 et a depuis nécessité des efforts importants en développement et ingénierie logicielle. De plus, il s'inscrit dans les thèmes de recherche du groupe de travail SCATI des GDR ISIS et I3. Les principes fondateurs de PELICAN (la polyvalence, l'extensibilité) s'appuient sur une vision particulière des logiciels de traitement d'image puisque nous souhaitons ici décorrélérer la conception et l'utilisation des algorithmes de traitement d'image.

Programme de recherche

Le quatrième et dernier chapitre présente les principaux axes de recherche que je souhaite développer dans les prochaines années. Naturellement tous les thèmes fondamentaux ou applicatifs auxquels je me suis intéressé sont porteurs de perspectives (celles-ci sont précisées dans les sections correspondantes des chapitres 2 et 3), et il est ambitieux (voire irréaliste ?) de croire que chacune d'entre elles pourra être étudiée avec attention.

J'ai identifié trois orientations principales de mes travaux qui ont pour origine différentes collaborations, notamment avec l'équipe LMB de l'Université de Freiburg et l'équipe TEXMEX de l'IRISA à Rennes. La première perspective concerne les notions d'invariance et d'imprécision en morphologie mathématique, et les

applications qui en découlent en imagerie du vivant. Une seconde orientation porte sur l'analyse de séquences vidéo et l'indexation multimédia au moyen de descripteurs morphologiques locaux. Enfin, le dernier axe que je souhaite étudier est une perspective à plus long terme et propose d'élargir les domaines d'application de la morphologie mathématique, à d'autres types, voire à tout type de données (et non nécessairement des images).

Ce chapitre comporte également une synthèse des contributions et les conclusions qui s'imposent au travail présenté dans ce mémoire.

Chapitre 1

Contexte

Ce chapitre a pour but de préciser le contexte dans lequel se sont déroulés mes travaux de recherche, présentés au titre de l'habilitation à diriger des recherches. Après avoir donné un historique de mes thèmes de recherche permettant de comprendre le changement de thématique intervenu suite à mon recrutement à l'Université de Strasbourg (section 1.1), je présente les principaux thèmes de recherche auxquels je me suis intéressé au sein du LSIIT (section 1.2). La suite de ce chapitre est plus factuelle et comporte, outre un curriculum vitae (section 1.3), une description de mes activités d'enseignement et charges administratives en section 1.4, et d'animation de la recherche (encadrement de jeunes chercheurs, implication dans la communauté scientifique, gestion de contrats et projets de recherche) en section 1.5. Ce chapitre se termine par la liste de mes publications (section 1.6).

1.1 Mon parcours : de l'applicatif au fondamental avec des applications

Cette section dresse un historique de mes activités de recherche, depuis ma thèse de doctorat en analyse et indexation de séquences vidéo (guidée par des considérations applicatives, sans doute induites par mon parcours universitaire : DUT puis formation d'ingénieur) jusqu'aux travaux en morphologie mathématique présentés au titre de l'HDR (où l'étude de quelques problèmes fondamentaux a débouché sur des solutions pertinentes dans différents contextes applicatifs).

Les références utilisées dans ce chapitre sont relatives à la liste de mes publications (section 1.6, pages 28 à 33). Dans les chapitres suivants, les références utilisées pourront être relatives à cette liste ou à la bibliographie générale (pages 125 à 134), selon qu'il s'agisse respectivement de travaux impliquant l'auteur ou non.

1.1.1 Analyse et indexation de séquences vidéo

Dans le cadre de ma thèse de doctorat réalisée de 1999 à 2002 au Laboratoire d'Informatique de l'Université de Tours¹ en collaboration avec la société Atos Origin (CIFRE), j'ai mené des recherches en indexation de séquences vidéo sous la direction de Nicole Vincent. Notre but était plus précisément d'y détecter en temps réel des événements prédéfinis. J'ai poursuivi ces travaux en 2002-2003 en tant qu'ATER dans ce même laboratoire, et jusqu'en 2004-2005 au LSIIT.

Afin de permettre la détection d'événements prédéfinis dans des séquences vidéo, nous avons proposé une architecture générique pouvant être adaptée à différents contextes et ainsi donner lieu à des systèmes spécifiques d'indexation vidéo répondant à des problématiques variées : détection de buts dans des retransmissions télévisées de matchs de football, vidéosurveillance de salles informatiques sécurisées, obtention de statistiques liées à la fréquentation d'un site public ou privé, analyse de scènes ferroviaires pour assurer la sécurité des personnes, etc. Cette architecture se compose de traitements élémentaires qui doivent être appliqués sur les données vidéo pour extraire des informations permettant leur indexation [C1]. De nombreuses applications d'une telle architecture nécessitant un fonctionnement en temps réel, ces traitements doivent être caractérisés par une faible complexité algorithmique [C3].

¹EA 2101, <http://www.li.univ-tours.fr>

Parmi ces traitements, la segmentation dans le domaine temporel des séquences vidéo est nécessaire pour séparer les différents plans. Après une analyse approfondie des différentes méthodes publiées dans la littérature [B1], nous avons proposé une méthode rapide, robuste aux conditions variables d'illumination, et permettant de traiter des séquences acquises avec des caméras en mouvement [B2, C2, D1]. Une fois la segmentation en plans effectuée, chaque plan peut être analysé indépendamment des autres afin d'interpréter la scène observée. La scène étant le plus souvent composée d'un arrière-plan et d'objets en mouvement, nos travaux ont ensuite porté sur l'analyse de ces deux parties de l'image.

Dans le cadre de l'étude de l'arrière-plan de la scène, deux problèmes principaux ont été abordés. Le premier consistait à séparer les objets présents dans la scène de l'arrière-plan, et une approche multirésolution pour la segmentation temps réel de scènes d'extérieur a été proposée [C2, D4]. La structuration du fond a également été étudiée sous l'hypothèse d'une modélisation de l'arrière-plan par des segments, le problème consistant alors en la détection de ces segments dans les images [D10]. Une approche locale et rapide a été proposée [D5] pour résoudre ce problème.

Les objets en mouvement présents dans la scène peuvent également être analysés au moyen d'algorithmes de suivi d'objet. Afin de permettre un suivi de qualité tout en limitant la complexité de l'approche utilisée, nous avons choisi de suivre les différents objets avec des approches adaptées. Plus précisément, nous avons différencié quatre types d'objets : les objets rigides pouvant être appris, les objets rigides pour lesquels un apprentissage est impossible, les objets de petite taille, et enfin les objets non rigides. Dans le cas d'objets rigides avec apprentissage, une méthode de suivi basée sur les chaînes de Markov cachées peut être utilisée [D11]. L'apprentissage est effectué hors ligne tandis que le suivi, consistant en la recherche d'une occurrence des images apprises, est effectué en ligne. Les objets de petite taille, facilement confondables avec du bruit présent dans l'image, doivent faire l'objet d'une attention particulière. Ainsi, nous avons proposé de les détecter et de les suivre par des approches adaptées, soit locales, soit globales [D9]. Dans le cas d'objets non-rigides, les contours actifs semblent être des outils plus appropriés. Cependant les modèles proposés dans la littérature ne permettent pas d'effectuer un suivi en temps réel. Nous avons donc proposé un algorithme original de suivi d'objet par contours actifs qui permet de suivre en temps réel des objets non-rigides (comme des personnes) dans des séquences d'images couleur acquises avec une caméra en mouvement [D3, D8, D12]. Cet algorithme combine notamment une analyse multirésolution des images et une capacité du modèle à changer de topologie.

La séquence vidéo peut aussi contenir d'autres types d'informations qu'il est pertinent d'analyser. Afin d'analyser les informations sonores, une approche en deux étapes a été envisagée. La première étape consiste en la classification de chaque extrait sonore (d'une durée de l'ordre de la seconde). Il est ensuite possible d'effectuer une analyse approfondie de l'extrait sonore en fonction de la classe à laquelle il appartient. La méthode de classification utilisée combine un algorithme de K-means et des chaînes de Markov cachées multidimensionnelles [D7, G1]. Comparée à des approches classiques, elle donne de meilleurs résultats [D6]. Les données textuelles (texte incrusté dans les images) peuvent également être extraites par détection des zones de texte incrusté. Nous avons fait ici le choix de combiner différentes approches pour assurer de meilleurs taux de reconnaissance [D13].

En conclusion, j'ai étudié pendant plusieurs années différents problèmes inhérents à la détection d'événements dans les séquences vidéo et me suis intéressé à un vaste panel de méthodes de traitement d'image. Mes thématiques se sont ensuite recentrées sur la morphologie mathématique.

1.1.2 Morphologie mathématique

J'ai rejoint en septembre 2003 le LSIIT (UMR 7005 CNRS–ULP) et intégré l'équipe Modèles, Images et Vision (MIV)² afin de renforcer les activités de recherche en morphologie mathématique animées par Christian Ronse. J'ai donc été naturellement amené à effectuer progressivement une reconversion thématique dans ce domaine (reconversion effective depuis 2004-2005), afin de pouvoir initier de nouvelles pistes de recherche et d'élargir les domaines d'application considérés (jusqu'alors principalement centrés sur l'imagerie médicale). En effet, à mon arrivée en 2003, le groupe de morphologie mathématique était constitué d'un enseignant-chercheur permanent (C. Ronse) et de trois doctorants (Julien Lamy, Benoit Naegel, Nicolas Passat) menant

²<http://lsiit-miv.u-strasbg.fr>

leurs travaux de thèse en imagerie médicale au travers de collaborations étroites avec d'autres laboratoires (IRCAD³ pour J. Lamy et B. Naegel, IPB⁴ pour N. Passat). Christian Ronse mène pour sa part, et depuis de nombreuses années, des travaux théoriques de tout premier plan dans le domaine de la morphologie mathématique.

Dans ce contexte, j'ai souhaité développer un nouveau thème de recherche fondamentale portant sur le traitement de données multivaluées en morphologie mathématique. Ma principale motivation à traiter ce problème encore ouvert était sans doute le nombre d'applications pouvant potentiellement bénéficier d'avancées sur ce sujet. Malheureusement, ce thème n'a pu faire réellement l'objet de travaux appliqués en imagerie médicale au travers de collaborations avec les autres partenaires du programme IRMC⁵, les images fournies par ces partenaires étant le plus souvent monovaluées. J'ai donc été amené à m'orienter vers de nouveaux domaines d'applications plus propices à la mise en oeuvre des avancées obtenues en amont en morphologie mathématique multivaluée. Parmi ces domaines d'applications, l'analyse d'image couleur pour la description et la recherche par le contenu a fait l'objet d'une convention CIFRE suite à une sollicitation émanant d'une jeune société alsacienne. Celle-ci n'ayant pas pu assurer le financement même partiel du doctorant (Erchan Aptoula), ce projet a finalement réussi à aboutir grâce au soutien d'Oséo-ANVAR. Une autre sollicitation est venue de collègues de l'équipe FDBT⁶ du LSIIT à la recherche de compétences en traitement d'image (segmentation) pour des projets ANR en télédétection en partenariat avec des géographes. Je me suis donc intéressé à ce nouveau domaine d'application particulièrement en adéquation avec le thème de la morphologie mathématique multivaluée. Cette collaboration avec l'équipe FDBT s'est également traduite par le co-encadrement d'un doctorant (Sébastien Derivaux) et l'ouverture vers un nouveau sujet de recherche, l'intégration de connaissances en morphologie mathématique, en particulier à l'aide de procédures de classification automatique (ce thème de recherche ne s'est par la suite pas limité à la télédétection, avec notamment des applications en recherche d'images couleur par le contenu). Ce sujet a aussi été abordé au travers de collaborations avec des géographes (par exemple, encadrement avec Anne Puissant du LIVE⁷ du stage de master recherche de Jonathan Weber). Enfin, suite à un séminaire donné par Christophe Collet présentant les travaux du groupe PASEO⁸ de l'équipe MIV, je me suis intéressé à l'utilisation des approches morphologiques en imagerie astronomique. Cette collaboration avec Christophe Collet (mais aussi plusieurs astronomes des observatoires de Strasbourg et Nice) s'est traduite notamment par deux co-encadrements de stage de master recherche (Erchan Aptoula et Benjamin Perret), et ma participation à l'encadrement de la thèse de Benjamin Perret sur les aspects morphologiques.

Ces derniers travaux, en cours lors de la rédaction de ce manuscrit, portent sur les représentations morphologiques arborescentes construites à partir d'images multivaluées en s'appuyant sur les notions de connexité ou d'hyperconnexité. Les arbres de coupe semblent en effet particulièrement adaptés pour l'identification des structures caractéristiques des objets astronomiques. Je suis également impliqué, de façon plus importante, dans la thèse de Jonathan Weber sur la fouille de séquences vidéo par une approche objet, réalisée en partenariat avec une SSII alsacienne et Pierre Gançarski de l'équipe FDBT. Nous nous intéressons notamment aux problématiques relatives à la segmentation morphologique de séquences vidéo par des approches efficaces et partiellement guidées, à l'apprentissage de segmentations en objets, et à l'intégration de ces tâches dans un processus de fouille de données. Outre ces deux encadrements de thèse, mes travaux actuels s'inscrivent dans le cadre de différentes collaborations, permises par plusieurs mobilités récentes (à Freiburg-im-Breisgau, Allemagne) ou en cours (délégation INRIA à Rennes). Celles-ci sont sources de perspectives aux travaux menés jusqu'à présent sur Strasbourg : citons par exemple l'étude de la notion d'invariance en morphologie mathématique, l'élaboration d'outils morphologiques pour l'indexation d'images et de séquences vidéo, ou encore l'application de la morphologie mathématique à des données non-image.

³Institut de Recherche contre les Cancers de l'Appareil Digestif, <http://www.ircad.fr>

⁴Institut de Physique Biologique, aujourd'hui Laboratoire d'Imagerie et de Neurosciences Cognitives LINC – UMR 7191, <http://alsace.u-strasbg.fr/ipb>

⁵Imagerie et Robotique Médicale et Chirurgicale, <http://irmc.u-strasbg.fr>

⁶Fouille de Données et Bioinformatique Théorique, auparavant Apprentissage et Fouille de Données, <http://lsiit.u-strasbg.fr/fdbt-fr>

⁷Laboratoire Image, Ville, Environnement — UMR 7011 CNRS-ULP, <http://imaville.u-strasbg.fr>

⁸Probabilistic Analysis for Space & Earth Observations, <http://lsiit-miv.u-strasbg.fr/paseo>

1.2 Résumé des travaux en morphologie mathématique

Nos activités de recherche en morphologie mathématique ont essentiellement porté sur deux thèmes fondamentaux (le traitement de données multivaluées et l'intégration de connaissances en morphologie mathématique) et trois applications (recherche d'image par le contenu, télédétection, imagerie astronomique). Les différentes contributions que nous avons apportées sont résumées dans cette section, qui donne également un plan des prochains chapitres de ce mémoire.

1.2.1 Travaux fondamentaux

Extension de la morphologie mathématique aux images multivaluées

La morphologie mathématique, qui connaît depuis 40 ans un vif succès dans la communauté de l'analyse et du traitement des images, offre un cadre rigoureux et des outils efficaces pour l'analyse spatiale et non-linéaire des images. Son application aux images binaires et aux images en niveaux de gris s'effectue très simplement en s'appuyant sur la théorie des ensembles ou mieux, celle des treillis. Cependant, son extension au cas des images multivaluées (où chaque pixel est représenté par un vecteur et non plus par un scalaire) n'est pas triviale et reste un problème ouvert. En effet, il n'existe pas de solution unique et unanimement acceptée pour ordonner des vecteurs. Les opérateurs morphologiques ne peuvent donc pas être utilisés directement sur des données multivaluées contenant par exemple une dimension spectrale ou temporelle.

Dans le cadre de la thèse de doctorat d'Erchan Aptoula, nous avons dressé un panorama des solutions proposées dans la littérature et identifié leurs limites [B3]. Nous nous sommes ensuite attachés à définir de nouveaux ordres vectoriels, en nous focalisant en particulier sur les ordres lexicographiques [A3] pour lesquels il est nécessaire de diminuer le caractère asymétrique. Plusieurs approches originales ont été proposées, en se basant par exemple sur des fonctions spatiales ou spectrales de quantification [B4], ou en appliquant au problème de l'ordonnancement de vecteurs le principe d'alpha-troncature fréquemment utilisé en filtrage [B5, D21].

Outre l'utilisation d'ordres vectoriels pour formuler les opérateurs morphologiques dans un cadre multivalué, nous avons étudié d'autres approches consistant à décomposer une image sous forme d'un ensemble d'images binaires ou à niveaux de gris. Dans le premier cas, il est possible d'appliquer sur chaque image binaire une séquence d'opérateurs morphologiques (binaires) [D17]. Dans le second cas, en définissant des paramètres (éléments structurants) spécifiques à chaque composante ou bande spectrale, on peut traiter chaque composante comme une image à niveaux de gris [D25]. Ces deux approches ont été appliquées dans le contexte des projets ANR FoDoMuST et Ecosgil, et ont notamment fait l'objet des stages de master de Jonathan Weber. Dans les deux cas, il est nécessaire ensuite d'unifier les résultats obtenus indépendamment sur chaque composante ou image monovaluée pour produire le résultat final. Dans le contexte de la segmentation, la prise en compte simultanée de plusieurs composantes est délicate mais permet d'obtenir un résultat intéressant, tout en s'affranchissant du problème du choix de l'ordre vectoriel [D15, D19].

Intégration de connaissances dans les traitements morphologiques

Indépendamment de la nature des images considérées, la création des systèmes d'analyse d'image par morphologie mathématique nécessite le plus souvent une expertise du domaine et une connaissance très fine du problème pour pouvoir choisir, combiner, et paramétrer les opérateurs morphologiques à bon escient. De ce fait, les méthodes morphologiques ne peuvent généralement pas être réutilisées dans un contexte différent de celui pour lequel elles ont été élaborées, et ne respectent que très peu la contrainte de généricité souhaitée en analyse d'image. Ce problème n'est bien sûr pas spécifique à la morphologie mathématique et est récurrent en traitement d'image [A1].

Afin de limiter l'implication de l'expert et la spécialisation des algorithmes, nous avons étudié comment il était possible de prendre en compte des connaissances ou des informations a priori dans les traitements morphologiques. Celles-ci peuvent être issues directement de l'image à traiter, comme par exemple sous la forme de régions auxquelles est associée une sémantique particulière fournie par l'utilisateur. Dans le contexte de la segmentation des images (ou partition des images en régions d'intérêt) par la méthode de ligne de partage

des eaux, l'utilisation de ces régions exemples peut intervenir à différentes étapes du processus. Lors d'un prétraitement, ces régions peuvent composer des ensembles d'apprentissage pour un processus de classification supervisée floue, ce qui permet d'obtenir un nouvel espace de représentation des données où chaque pixel est associé aux degrés d'appartenances aux différentes classes. La segmentation peut alors s'effectuer dans cet espace de représentation, en utilisant par exemple des informations spectrales [D14] ou texturales [E13] lors de la phase de classification. Il est également possible d'exploiter les régions exemples directement au sein de l'algorithme de segmentation en modifiant celui-ci et en l'adaptant à l'existence de plusieurs cartes d'appartenances des pixels aux classes [D19]. Enfin, ces régions peuvent aussi être utilisées comme zones de validation lors d'un post-traitement dans le cadre d'une optimisation des paramètres de segmentation par un algorithme génétique [D20]. L'utilisation de zones exemples lors d'un pré ou post-traitement s'inscrit dans les travaux de thèse de doctorat de Sébastien Derivaux, en partenariat avec des membres de l'équipe FDBT du LSIIT. Le paradigme de l'apprentissage (qu'il soit supervisé, semi-supervisé, actif) est également abordé dans la thèse de doctorat de Jonathan Weber (toujours en partenariat avec l'équipe FDBT) dans le but de produire des segmentations de séquences vidéo en objets sémantiques (du point de vue de l'utilisateur).

Dans certains cas, la connaissance disponible est plus limitée. Ainsi, l'expert peut connaître, à défaut d'une description des classes d'intérêt présentes dans l'image à analyser, uniquement leur nombre. Il est alors possible d'effectuer une classification non-supervisée de l'image à traiter et d'appliquer les traitements morphologiques sur l'image classée. Si la classification est dure, chaque pixel est affecté à une et une seule classe, et l'image classée peut être représentée sous la forme d'un ensemble d'images binaires [D17]. Si la classification est floue, chaque pixel est représenté par les degrés d'appartenances aux différentes classes [E9]. Lorsque le nombre d'objets est connu, il est possible de construire une méthode de segmentation automatique et itérative, fonctionnant sur le principe de l'algorithme K-Means [E15, A4].

En outre, la connaissance peut également être introduite au moyen des éléments structurants. Dans le contexte de la détection d'objets, ceux-ci représentent les motifs recherchés au travers d'une information de taille ou de forme [D17], de réponse spectrale [D25], de profil d'intensité [B9], etc.

1.2.2 Applications

Analyse d'images couleur pour la recherche par le contenu

Face à l'accroissement continu de la place des images et vidéos dans les systèmes d'information (professionnels ou destinés au grand public), les utilisateurs ont besoin d'outils pour organiser leurs données et y accéder efficacement. L'automatisation de ces tâches peut être assurée en exploitant le contenu des données de manière appropriée, par exemple à l'aide d'approches morphologiques. Dans le cas d'images couleur, ceci implique la définition d'opérateurs morphologiques adaptés. L'analyse d'images couleur par morphologie mathématique a fait l'objet de la thèse de doctorat d'Erchan Aptoula, et outre la proposition d'ordres vectoriels compatibles avec une représentation des images en couleurs [A3], nous avons plus spécifiquement étudié comment prendre en compte correctement l'information de teinte [B8].

Nous nous sommes ensuite intéressés à la description du contenu des images couleur par des approches morphologiques, en proposant de nouveaux descripteurs relatifs aux informations de couleur [B10] et de texture [D16, D22]. Ces descriptions illustrent l'intérêt de la morphologie mathématique pour les espaces-échelle [A2] et sont notamment pertinentes lorsque leur nature est multidimensionnelle [D18, B6]. Nous avons également élaboré plusieurs approches pour la segmentation (guidée [D19] ou non [D23]) et la détection d'objets dans les images couleur [B7].

L'application principale de ces différents travaux en analyse d'images couleur a concerné l'annotation semi-automatique et la recherche d'images par le contenu, projet soutenu par Oséo-ANVAR dans le cadre d'un contrat d'aide à l'innovation 2006 – 2009. Ce projet se caractérise par une annotation semi-automatique réalisée hors ligne, où l'utilisateur fournit des exemples et contre-exemples pour définir chaque mot-clé présent dans la base. À partir de ces données, le système est en mesure de définir les meilleurs descripteurs à utiliser pour associer ce mot-clé aux images de la base, et l'utilisateur peut intervenir sur les résultats obtenus afin d'améliorer la qualité de l'annotation. La recherche est ensuite effectuée en ligne et fournit immédiatement la liste des images associées à un mot-clé prédéfini.

Télédétection

Comme dans de nombreux domaines, la morphologie mathématique n'est pas utilisée en télédétection à la hauteur de ses possibilités. Les outils morphologiques sont principalement employés dans un but de filtrage et de classification d'images panchromatiques (ou images à niveaux de gris). Cependant, les avancées théoriques récentes offrent de nouveaux débouchés dans ce domaine, et nous nous intéressons en particulier aux applications en imagerie à très haute résolution spatiale, dans les espaces urbains et côtiers. Dans ces domaines, nous cherchons à utiliser la morphologie mathématique pour résoudre différents problèmes récurrents que sont la segmentation en régions, la détection d'objets d'intérêt, ou encore la description et la classification des données (à l'échelle des pixels ou des régions).

L'interprétation d'images à très haute résolution spatiale s'effectue généralement en deux étapes : tout d'abord la segmentation extrait des images les régions qui seront ensuite étiquetées (interprétées) par une méthode de classification (supervisée par exemple). Il s'agit donc d'un domaine d'application particulièrement pertinent pour nos travaux en segmentation morphologique guidée par des exemples, quelle que soit la nature des images considérées. Ainsi, nous nous sommes intéressés aux approches sans connaissance spatiale [D14, D20, G3] et à leur application aux espaces urbains dans le cadre de la thèse de doctorat de Sébastien Derivaux, tandis que l'utilisation de marqueurs pour la segmentation a fait l'objet d'une collaboration plus récente avec Anne Puissant du LIVE [C5] dans le contexte de l'analyse du littoral par télédétection, et d'une participation à une compétition internationale [D27] pour la détection automatique de bâtiments.

Dans le contexte de la télédétection appliquée aux espaces urbains, nos travaux s'effectuent en partenariat avec l'équipe FDBT et des géographes du LIVE, notamment au travers du projet ACI Masse de Données FoDoMuST 2004 – 2007. Outre les travaux en segmentation guidée par des exemples, nous nous sommes intéressés à la détection des zones de bâti dans les images panchromatiques à très haute résolution spatiale. Pour résoudre ce problème, nous avons décidé de représenter les données par un ensemble d'images binaires via l'utilisation d'une classification non-supervisée [D17]. La transformée en tout-ou-rien est ensuite appliquée pour rechercher dans ces images des formes prédéfinies [C4].

Dans le contexte de la télédétection appliquée aux zones côtières, nos travaux s'effectuent au sein du projet ANR Jeunes Chercheurs Ecosgil 2005 – 2008, en partenariat avec des géographes du LIVE et de plusieurs laboratoires de l'Université de Caen – Basse Normandie. Outre nos travaux en segmentation guidée [C5], nous nous intéressons à la détection d'objets géographiques du littoral. Ainsi, nous avons élaboré une approche morphologique pour la détection d'objets dans des images multispectrales [D25], approche qui a été appliquée avec succès à la détection du trait de côte dans des images à très haute résolution spatiale [D24].

Imagerie astronomique

En imagerie astronomique, la communauté scientifique connaît peu les outils morphologiques. Ce constat s'explique certainement par le niveau de bruit présent dans les images, parfois très important. Nos travaux dans ce domaine s'effectuent en étroite collaboration avec Christophe Collet et le thème PASEO de l'équipe MIV, ainsi qu'avec les Observatoires de Strasbourg et de Nice – Côte d'Azur.

A l'aide de la morphologie mathématique, nous cherchons à résoudre deux problèmes principaux. Tout d'abord nous souhaitons obtenir une classification générale des galaxies. Une solution a été élaborée dans le cadre du stage de DEA d'Erchan Aptoula [D15], mais celle-ci n'est basée que sur des critères morphologiques (segmentation morphologique de la galaxie puis description des zones de formation stellaire). Afin d'améliorer la caractérisation des objets astronomiques, il semble nécessaire de considérer la nature spectrale des données conjointement avec les informations spatiales, et de combiner les approches morphologiques avec des modèles statistiques (inférence bayésienne). Ce problème est abordé dans le cadre de la thèse de doctorat de Benjamin Perret où les contributions en morphologie mathématique concernent les arbres de coupe produits à partir d'images multivaluées, et à partir de différents types de connexité.

L'autre problème que nous abordons dans le domaine de l'imagerie astronomique est celui de la détection de certaines galaxies particulières. Ainsi, nous nous sommes focalisés sur les galaxies à faible brillance de surface dans le cadre du stage de master recherche de Benjamin Perret. En intégrant la modélisation du signal (objets, bruit) dans l'approche morphologique de détection d'objet (la transformée en tout ou rien), nous avons proposé

une nouvelle méthode apte à détecter des objets même en cas de très faible rapport signal-bruit [B9, D26].

Développement d'une plate-forme pour l'analyse d'image

En complément des travaux scientifiques décrits précédemment, et liés exclusivement à des développements théoriques et applicatifs de la morphologie mathématique, j'ai initié localement en 2004 un projet de plate-forme pour l'analyse d'image, intitulé PELICAN (Polyvalent Extensible Library for Image Computing and ANalysis). L'élaboration de cette plate-forme nécessite des efforts importants en développement et ingénierie logicielle, tout en s'intégrant dans les thèmes de recherche du groupe de travail SCATI (Systèmes Complexes pour l'Analyse et le Traitement d'Images) des GDR ISIS et I3. Ce groupe de travail s'intéresse à la notion de systèmes pour le traitement et l'analyse d'image, en faisant appel aux outils et méthodes relevant du traitement d'image, de la fusion de données, de l'intelligence artificielle, de l'ingénierie de la connaissance, de la cognition et de la perception visuelle.

L'intérêt de notre proposition, outre bien sûr de servir de dépôt commun d'algorithmes de traitement d'image pour les chercheurs et étudiants, est de permettre l'accès à ces traitements sous différentes formes, contrairement aux solutions existantes dédiées le plus souvent à un seul type de public : développeurs, experts en traitement d'image, grand public. Favoriser la reproductibilité et la réutilisation des solutions de traitement d'image est un des objectifs de SCATI.

Notre plate-forme est définie par une architecture en trois couches : noyau, algorithmes, et interfaces. Le noyau ou cœur de la plate-forme définit les structures de données et sert de médiateur entre les interfaces et les algorithmes. Les algorithmes correspondent aux méthodes de traitement d'image développées par les chercheurs et étudiants, définies selon les spécifications introduites dans le noyau. Enfin, les interfaces permettent aux utilisateurs finaux d'accéder à tout ou partie des traitements (ou algorithmes) disponibles dans la plate-forme, d'une façon complètement générique puisque les interfaces ne font pas directement référence aux algorithmes mais communiquent uniquement avec le noyau. Cette architecture permet de garantir à la plate-forme les propriétés de polyvalence (tout type d'utilisateur ou d'usage) et d'extensibilité (ajout dynamique de fonctionnalités : algorithmes ou interfaces).

1.2.3 Synthèse

Les travaux de recherche décrits ici sont souvent le fruit d'un travail en collaboration, notamment au travers de l'encadrement de jeunes chercheurs. Ainsi, j'ai pu encadrer ou participer à l'encadrement de deux thèses soutenues (Erchan Aptoula, 2005–2008 ; Sébastien Derivaux, 2005–2009) et deux thèses en cours (Benjamin Perret et Jonathan Weber, toutes deux débutées en 2007), de quatre DEA ou master recherche (Fateh Beder, 2004 ; Erchan Aptoula, 2005 ; Benjamin Perret et Jonathan Weber, 2007), et d'un master en apprentissage (Florent Sollier, 2007–2008).

En outre, la plupart de ces travaux ont été réalisés dans le cadre de projets auxquels j'ai participé : ACI Masses de Données Fodomust 2004–2007, ANR Jeunes Chercheurs ECOSGIL 2005–2008, soutien Oséo-ANVAR 2005–2009, collaboration CIFRE avec RBS 2007–2010, projet BQR de l'UDS Écoulements gravitaires 2008–2009. J'ai notamment assuré la direction scientifique de deux de ces projets (Oséo et CIFRE).

Enfin, ces travaux ont fait l'objet de 8 publications en revues internationales, de 2 publications en revues nationales, de 3 chapitres d'ouvrage, de 13 communications en conférences internationales, et de 8 communications en conférences nationales. 2 articles ont également été soumis en revue internationale et sont en cours d'analyse par les comités éditoriaux (dont 1 en révision). La liste de ces publications est donnée en pages 28 à 33.

1.3 Curriculum vitae

État civil

Nom et prénom : **LEFEVRE Sébastien**
 Date et lieu de naissance : 17 octobre 1977 à Lille (Nord)
 Nationalité : Française
 Situation familiale : Marié, 4 enfants
 Adresse professionnelle : IRISA, équipe TEXMEX
 Campus de Beaulieu 35042 Rennes
 Tél. : +33 (0)2 99 84 71 91 – Sec. +33 (0)2 99 84 74 37
 Adresses électronique : sebastien.lefevre@irisa.fr et lefevre@unistra.fr
 Page personnelle : <http://dpt-info.u-strasbg.fr/~lefevre>

Position Administrative Actuelle (depuis septembre 2003)

Maître de Conférences en Informatique à l'Université de Strasbourg (UDS) :

Enseignant à l'UFR de Mathématique et d'Informatique, Département Informatique
 Chercheur au Laboratoire des Sciences de l'Image, de l'Informatique, et de la Télédétection
 (LSIIT, UMR 7005 CNRS – UDS), dans l'équipe MIV (Modèles, Image et Vision)

Titulaire de la **Prime d'Encadrement Doctoral et de Recherche** depuis octobre 2006

Bénéficiaire d'un **Congé pour Recherches (CRCT)** pour les seconds semestres 2007–08 et 2008–09

Accueil en **Délégation à l'INRIA Rennes Bretagne Atlantique** (équipe TEXMEX) en 2009–2010.

Formation Universitaire

1994 Baccalauréat série C

Lycée Montesquieu, Herblay (Académie de Versailles)

1996 Diplôme Universitaire de Technologie en Informatique

Institut Universitaire de Technologie d'Orsay, Université Paris XI

1999 Diplôme d'Ingénieur en Génie Informatique, filière Communication Homme-Machine

Université de Technologie de Compiègne

1999 Diplôme d'Études Approfondies en Contrôle des Systèmes, Mention Très Bien

Université de Technologie de Compiègne

Titre du mémoire : *Élaboration et validation d'un outil automatique de segmentation et de reconstruction volumique sur imagerie IRM de tumeurs cérébrales*

Laboratoire d'Accueil : UMR 6615 – IRCOM-SIC, Poitiers

Responsable de DEA : Christine Fernandez-Maloigne

2002 Doctorat en Informatique, Mention Très Honorable avec les Félicitations du Jury

Université de Tours

Titre du mémoire : *Détection d'événements dans une séquence vidéo*

Laboratoire d'Accueil : EA 2101 – Laboratoire d'Informatique

Directeurs de thèse : Nicole Vincent et Christian Proust

Financement : CIFRE avec Atos Origin (Blois)

Jury : Christophe Brunet (Atos Origin), Serge Miguët, Henri Nicolas, Jack-Gérard Postaire (Rapporteur), Christian Proust, Florence Sèdes (Rapporteur), Georges Stamon (Président), Nicole Vincent.

Expérience professionnelle

Seules les expériences liées à une activité de recherche sont décrites ici.

Mars – Août 1999

Stage de recherche (DEA)

Laboratoire IRCOM-SIC et CHU de Poitiers. Travaux de recherche relatifs à la segmentation d'images IRM (et notamment la phase initiale de pré-segmentation du cerveau, obtenue par filtrage, seuillage, et méthode des contours actifs).

Septembre 1999 – Août 2002

Doctorant CIFRE

Université François Rabelais, Tours. Activités de recherche dans l'équipe Reconnaissance de Formes et Analyse d'Image du Laboratoire d'Informatique, travaux de doctorat en analyse et traitement des images et séquences vidéo. Activités d'enseignement dans l'École d'Ingénieurs en Informatique pour l'Industrie (E3i) principalement dans les domaines : bases de données, réseaux, programmation. Participation à la vie universitaire et aux tâches collectives.

Septembre 2002 – Août 2003

Attaché Temporaire d'Enseignement et de Recherche

Université François Rabelais, Tours. Activités de recherche dans l'équipe Reconnaissance de Formes et Analyse d'Image du Laboratoire d'Informatique, coresponsabilité du projet Faiveley en vidéosurveillance ferroviaire, et travaux de recherche en analyse et traitement des images et séquences vidéo. Activités d'enseignement dans l'École Polytechnique de l'Université de Tours (EPU-Tours, ex-E3i), principalement dans les domaines : vision et multimédia, parallélisme, logique, LaTeX. Participation à la vie universitaire et aux tâches collectives.

depuis Septembre 2003

Maître de Conférences en Informatique

Université de Strasbourg. Activités de recherche en traitement d'image (morphologie mathématique) au LSIIT (équipe MIV). Activités d'enseignement à l'UFR de Mathématique et d'Informatique, principalement dans les domaines de la conception et programmation orientée objet et du traitement d'image. Participation à la vie universitaire et aux tâches collectives. Mobilités à Freiburg-im-Breisgau (2007–2008) et à Rennes (2009–2010), respectivement sur les thèmes des invariants et de l'imagerie biologique, et de l'indexation multimédia.

1.4 Activités d'enseignement et responsabilités pédagogiques

En tant que Maître de Conférences en Informatique à l'Université de Strasbourg, j'ai naturellement été impliqué dans des activités d'enseignement et d'administration à vocation pédagogique que je détaille ici.

1.4.1 Enseignements

Je décris ici les principaux thèmes pour lesquels j'assure des enseignements, donne une liste des projets à vocation pédagogique que j'ai encadrés, et enfin, dresse un historique de mes activités d'enseignement.

Disciplines enseignées (2007–2009)

En lien avec mes thématiques de recherche en **Image – Vision – Multimédia**, j'ai assuré le cours de Traitement et Codage d'Image en 2ème année du Master d'Informatique, spécialité professionnelle Informatique de l'Image et du Calcul Intensif. L'objectif de ce cours était de fournir aux étudiants les connaissances et compétences en traitement d'image leur permettant de résoudre des problèmes rencontrés dans un contexte industriel. Dans ce contexte, je dressais dans ce cours un large panorama des méthodes de traitement d'image et l'illustrais par des séances de travaux pratiques où les étudiants avaient l'occasion d'élaborer des solutions à des problèmes réels de grande ampleur. Ce cours faisait également l'objet d'un stage court de formation continue intégré à l'offre de l'UDS. Auparavant, j'intervenais dans l'enseignement de Traitement d'Image en Master Informatique 1ère année de l'UDS et dans les enseignements d'Applications Multimédia et de Vision en formation d'ingénieur à l'Université de Tours.

En lien avec mes activités en **Génie Logiciel et Conception / Programmation Orientée Objet** menées dans le cadre des projets auxquels je participe (cf. section 1.5.3) et notamment la plate-forme Pelican, j'ai assuré le cours de Programmation Orientée Objet en 2ème année du parcours Compétence Complémentaire en Informatique offert à l'ensemble des Masters de l'UDS. Dans ce cours destiné à des scientifiques néo-informaticiens, j'abordais le génie logiciel à l'aide du paradigme orienté objet, depuis l'analyse et la modélisation UML jusqu'à la réalisation de logiciels à l'aide du langage Java et de l'environnement Eclipse. Auparavant, j'intervenais également dans les enseignements de Conception et Programmation Orientée Objet en Licence d'Informatique et d'Outils de Génie Logiciel en Master d'Informatique de l'UDS, et dans les enseignements de Java en formation d'ingénieur à l'Université de Tours.

Dans les deux domaines décrits précédemment, j'ai réalisé la majorité des supports de cours et me suis beaucoup impliqué dans la définition des programmes pédagogiques. Je devais également assurer la responsabilité de plusieurs modules relatifs à ces thématiques dans différentes spécialités de Master proposés pour la première fois à l'Université de Strasbourg en 2009 – 2010 : Informatique et Sciences de l'Image, Observation de la Terre et Géomatique, Ingénierie du Logiciel et des Connaissances. Mon accueil en délégation INRIA à partir de septembre 2009 retardera mon intervention dans ces nouvelles formations.

Encadrement de projets d'étudiants

Outre les activités d'encadrement de jeunes chercheurs décrites en section 1.5.1, j'ai encadré plusieurs projets de second cycle à vocation pédagogique (50 à 100 heures de travail étudiant) dans différentes formations de l'Université de Tours de 1999 à 2003 puis de l'UDS depuis 2003 :

- en Master Informatique (UDS) avec 8 projets de 2ème année et 16 projets de 1ère année (TER) ;
- en IUP (UDS) avec 4 projets de 2ème année ;
- en formation d'ingénieur (EPU Tours) avec 4 projets de 3ème année, 39 projets de 2ème année, et 46 projets de 1ère année.

Historique de mes activités d'enseignement

- 1999 – 2002** Vacataire à l'Université de Tours 100 à 150 h / an
Enseignements sous forme de vacations dans les 3 années de formation d'ingénieur E3i et en formation continue (DUAC : Diplôme Universitaire d'Actualisation des Connaissances). Disciplines enseignées en CM/TD/TP : bases de données, administration système et réseau, Matlab, langage Java et interfaces graphiques, technologies Web, vision par ordinateur et applications multimédia.
- 2002 – 2003** ATER à l'Université de Tours 192 h / an
Service d'enseignement statutaire dans les 3 années de formation d'ingénieur EPU (ex-E3i), ainsi que dans le DEA d'Informatique et la formation (modules transversaux) de l'École Doctorale Science, Santé et Technologie de l'Université de Tours. Disciplines enseignées en CM/TD/TP : vision par ordinateur, applications et indexation multimédia, Matlab, LaTeX, Parallélisme, Logique.
- 2003 – 2006** MCF à l'Université de Strasbourg environ 192 h / an
Service d'enseignement statutaire au sein du département informatique de l'UFR de Mathématique et d'Informatique. Enseignements dans les filières suivantes : DEUG Sciences, Licence de Mathématiques, IUP Génie Mathématique et Informatique, Maîtrise et DESS d'Informatique. Depuis 2005 (réforme LMD) : enseignements et responsabilités de modules en Master d'Informatique et Licence MPC (Maths-Physique-Chimie). Stage de formation continue (5 jours) en techniques de bases en analyse et traitement des images, intégré dans l'offre de l'UDS depuis 2005 et assuré trois fois depuis. Disciplines enseignées en CM/TD/TP : bases de données, génie logiciel, conception orientée objet, algorithmique et programmation, traitement d'image.
- 2007 – 2009** MCF avec CRCT à l'Université de Strasbourg CRCT, 96 h / an
Service d'enseignement statutaire réduit (96 h annuelles) suite à l'octroi de deux semestres de CRCT (1 en 2007–2008 et 1 en 2008–2009). Enseignements et responsabilités de modules de Programmation Orientée Objet en Master 2 CCI et Traitement et Codage d'Image en Master 2 IICI.

1.4.2 Charges administratives et fonctions d'intérêt collectif

Master CCI

Depuis 2005, j'ai été responsable pédagogique du Master de Sciences, Parcours Compétence Complémentaire en Informatique (Master CCI) à l'Université de Strasbourg. Cette formation était destinée à des étudiants scientifiques non informaticiens ayant au minimum validé une première année de Master, afin de leur permettre d'acquérir des compétences en informatique complémentaires à leurs compétences initiales dans un autre domaine scientifique. Plus fréquemment, les étudiants étaient titulaires d'un diplôme de Master ou de Doctorat et souhaitaient effectuer une reconversion thématique (les débouchés en informatique étant nombreux). Mes activités en tant que responsable pédagogique ont porté sur :

- la coordination de l'équipe pédagogique (définition des emplois du temps, organisation des jurys),
- les actions de communication (réalisation d'affiches, organisation de campagnes de communication),
- la gestion des recrutements en partenariat avec les commissions pédagogiques (chaque dossier étant évalué par la commission pédagogique de l'UFR de Mathématique et d'Informatique mais également par la commission pédagogique de la discipline initiale du candidat),
- la participation au groupe de travail national CCI regroupant les responsables de ces filières dans les différentes universités françaises, ainsi que la rencontre des tutelles au Ministère de l'Education Nationale pour discuter des problèmes relatifs au Master CCI dans le contexte de la réforme LMD,
- et, bien sûr, l'accueil des étudiants, leur suivi tout au long de l'année universitaire et leur orientation (stages et débouchés) en fonction de leur parcours initial.

Bénéficiant d'un Congé pour Recherches (CRCT) en 2007-2008 et 2008-2009, je n'ai pas assumé la charge officielle de responsable pédagogique ces deux années, tout en restant disponible pour aider mon remplaçant.

Projet de Master 2id

La formation d'une année de Master 2 CCI étant par essence incompatible avec le schéma LMD, nous avons décidé en 2007 à l'UFR de Mathématique et d'Informatique de ne pas demander la réhabilitation de cette formation en 2009 mais plutôt de travailler sur une nouvelle offre de formation en 2 ans intitulée Master Informatique Inter Disciplinaire (2id). J'ai été à l'initiative de ce projet et par la suite coresponsable du projet avec Pierre Tellier, Maître de Conférences en Informatique à l'UDS. Mes activités en tant qu'initiateur et coresponsable du projet ont porté sur :

- l'animation du groupe de réflexion au sein du département informatique,
- la définition du contenu de la formation,
- la rédaction de la maquette d'habilitation.

Malgré un fort soutien de l'Université de Strasbourg et de différents acteurs socio-économiques locaux, ce projet n'a pas reçu l'accueil espéré de la part des instances nationales. Néanmoins, je reste persuadé de l'intérêt de notre proposition. Nul doute que nous retravaillerons celle-ci dans le futur pour proposer une formation adaptée aux scientifiques intéressés par l'informatique.

— Synthèse des responsabilités et activités collectives —

1999 – 2002	Représentant élu des Doctorants au Conseil du Laboratoire d'Informatique, Tours
2003 – 2009	Correspondant Specif pour l'Université de Strasbourg
2004 – 2008	Rapporteur de stages de DEA / Master Recherche de l'Université de Strasbourg
2005 – 2007	Responsable pédagogique du Master Compétence Complémentaire en Informatique (CCI)
2006 – 2008	Membre suppléant élu de la Commission de Spécialiste 27ème section de l'UDS
2007 – 2009	Coresponsable du projet de création de Master Informatique Inter Disciplinaire (2id)
depuis 2008	Membre élu du Conseil d'Administration de Specif
2009	Membre extérieur d'un Comité de Sélection MCF 27 de l'Université Paris Descartes

1.5 Animation de la recherche

1.5.1 Encadrement d'activités de recherche

L'encadrement de jeunes collègues chercheurs, étudiants en second ou troisième cycle, fait partie intégrante de l'activité de recherche. Les travaux que j'ai encadrés s'inscrivent donc tout naturellement dans les différents thèmes de recherche que je développe au LSIIT, et qui feront l'objet des prochains chapitres.

Les publications indiquées ici utilisent les références introduites dans la section 1.6.

— Synthèse des activités d'encadrement —

2	thèses soutenues co-encadrées (75 % et 33 %)
2	thèses en cours co-encadrées (50 % et 30 %)
4	stages de DEA / master recherche encadrés ou co-encadrés (entre 50 et 100 %)
1	master en apprentissage encadré (100 %)
3	stages de master 2 CCI encadrés (100 %)
1	stage de master 1ère année co-encadré (50 %)
8	projets de fin d'études co-encadrés (50 %)

Co-encadrement de Thèses de Doctorat

Remarque : un astérisque signale les rapporteurs des thèses co-encadrées.

1. Erchan Aptoula, *Analyse d'images couleur par morphologie mathématique. Application à la description, l'annotation, et à la recherche d'images.*
Thèse débutée en octobre 2005 et soutenue le 10 juillet 2008. Financement : contrat Oséo-ANVAR.
Jury de thèse : C. Ronse, J. Serra*, C. Fernandez-Maloigne*, M. Tajine*, B. Sankur, S. Lefèvre.
Prix de thèse 2009 de la Société des Amis des Universités de l'Académie de Strasbourg.
Publications : 1 chapitre de livre [A3], 6 articles en revues internationales [B3, B4, B5, B8, B7, B10], 4 communications en conférences internationales [D16, D21, D22, D23] et 1 en conférence nationale [E12].
Situation actuelle : service militaire (depuis novembre 2008).
Encadrement à 75% avec Pr. Christian Ronse (Directeur).
Nature de l'encadrement : j'ai défini le sujet de recherche, effectué la recherche du financement et assuré l'encadrement de l'étudiant dans sa quasi-intégralité tout au long des 3 années de thèse.
2. Sébastien Derivaux, *Construction et classification d'objets à partir d'images de télédétection par une approche itérative guidée par des connaissances du domaine.*
Thèse débutée en novembre 2005, soutenue le 5 mai 2009. Financement : allocation MESR.
Jury de thèse : J. Korczak, F. Sèdes*, M. Thonnat*, P. Collet*, B. Crémilleux ; invités : A. Puissant, S. Lefèvre, C. Wemmert.
Publications : 2 en conférences internationales [D14, D20], 4 en conférences nationales [E10, E11, E13, E14], et 1 article soumis [G3] (et des publications en classification que je n'ai pas cosignées).
Situation actuelle : Ingénieur Consultant (CDI) chez Arias Associates, Strasbourg depuis 2009.
Encadrement à 33% avec l'équipe FDBT : Pr. Jerzy Korczak (Directeur) et Dr. Cédric Wemmert.
Nature de l'encadrement : j'ai participé à la définition du sujet de recherche et assuré l'encadrement de l'étudiant pour les travaux en segmentation d'image (par morphologie mathématique). Mon intervention lors des travaux complémentaires en classification est négligeable.
3. Benjamin Perret, *Classification des galaxies : caractérisation spatiale et spectrale par inférence de haut niveau sémantique.*
Thèse débutée en octobre 2007. Financement : allocation MESR.
Encadrement à 30% avec Pr. Christophe Collet (Codirecteur), Dr. Eric Slezak (HDR, Codirecteur, Observatoire de la Côte d'Azur), Dr. Vincent Mazet.
Nature de l'encadrement : j'ai participé à la définition du sujet de recherche et assuré l'encadrement de

l'étudiant pour les travaux en morphologie mathématique. Mon intervention lors des travaux complémentaires en inférence bayésienne et modélisation statistique a été négligeable.

4. Jonathan Weber, *Structuration automatique d'ensembles d'images et de séquences vidéo.*

Thèse débutée en décembre 2007. Financement : CIFRE avec RBS.

Publications : 1 article soumis [G2].

Encadrement à 50 % avec l'équipe FDBT : Dr. Pierre Gançarski (HDR, Directeur 2007–2009).

Nature de l'encadrement : *j'ai défini le sujet de recherche, effectué la recherche du financement et assuré la majeure partie de l'encadrement de l'étudiant depuis le début de sa thèse. Depuis octobre 2009, j'assure la direction officielle de la thèse à un taux de 67 %.*

Co-encadrement de stages de DEA / Master Recherche

Ces stages ont été effectués par des étudiants du DEA puis master recherche en informatique de l'Université de Strasbourg.

Année	Etudiant	Titre des travaux	Taux	Publications
2004	F. Beder	<i>Détection de changements de plans par morphologie mathématique</i>	100 %	
2005	E. Aptoula	<i>Segmentation d'images astronomiques multibandes par morphologie mathématique</i>	50 % ¹	[D15]
2007	B. Perret	<i>Détection et caractérisation de galaxies fantômes en imagerie astronomique bibande</i>	50 % ¹	[B9, D26]
2007	J. Weber	<i>Détection d'objets géographiques du littoral dans des images satellites multispectrales</i>	50 % ²	[D24, D25]

¹ Les deux stages de Erchan Aptoula et Benjamin Perret en morphologie mathématique appliquée à l'imagerie astronomique ont été co-encadrés avec Christophe Collet, Professeur 61e au LSIIT et responsable du thème MIV-PASEO.

² Le stage de Jonathan Weber en morphologie mathématique appliquée à la télédétection a été co-encadré avec Anne Puissant, Maître de Conférences 23e au LIVE, Strasbourg.

Autres activités d'encadrement à l'UDS

Printemps 2004 Loïc Fortchantre, *Participation à la conception d'un système générique d'analyse d'image*, Stage de DESS Compétence Complémentaire en Informatique (4 mois), Encadrement à 100 %.

Printemps 2004 Abdelhak Mokthari, *Réalisation d'une bibliothèque d'opérateurs de morphologie mathématique*, Stage de DESS Compétence Complémentaire en Informatique (4 mois), Encadrement à 100 %.

Été 2006 Jonathan Weber, *Détection de bâtiments dans des images satellites*, Stage de Master Informatique 1ère année, Encadrement à 50 % avec Dr. David Sheeren, MCF à l'ENSAT Toulouse (Post-doc en 2006 au LSIIT). Publications relatives à ce projet [C4, D17, E9].

2007-2008 Florent Sollier, *Développement d'une plate-forme d'analyse d'image*, Master Ingénierie Logicielle en Apprentissage (3 semaines par mois pendant 12 mois), Encadrement à 100 %. La charge de maître d'apprentissage nécessite une implication plus forte que lors des projets, stages ou thèses.

2008-2009 Régis Witz, *Annotation et recherche d'image par le contenu*, Ingénieur CDD (10 mois) recruté dans le cadre du contrat Oséo, Encadrement à 100 %.

Printemps 2009 Aurélie Mignot, *Classification morphologique en télédétection par approche évolutionnaire*, Stage de Master Compétence Complémentaire en Informatique (3 mois), Encadrement à 50 % avec Dr. Alexandre Blansché, ATER à l'UDS (équipe FDBT).

Projets de fin d'études d'ingénieur

Ces projets de recherche ont été effectués par des étudiants de l'EPU Tours lors de la dernière année de leur formation à raison de 2 jours par semaine pendant 8 mois. Ils ont tous été encadrés à 50 % en collaboration avec le Professeur Nicole Vincent, et porté sur des thèmes relatifs à mes travaux doctoraux.

Année	Etudiant	Titre des travaux	Publications
2000	E. Bouton	<i>Suivi d'objet dans une séquence d'images</i>	[D11]
2000	J. Holler	<i>Segmentation par plans d'une séquence vidéo</i>	[B1, D1, E1]
2001	B. Maillard	<i>Reconnaissance d'événements dans une bande sonore</i>	[D6, D7, E4]
2001	J. Dury	<i>Extraction de texte dans des séquences vidéo</i>	
2002	C. Jousse*	<i>Détection de lignes dans des images couleur en temps réel</i>	[D5, D10]
2002	T. Foissotte	<i>Détection et suivi d'objets de petite taille</i>	[D9]
2003	C. L'Orphelin	<i>Localisation du texte dans les séquences vidéo</i>	[E5]
2003	F. Lagoueyte	<i>Reconstruction de scène à partir de droites extraites</i>	

* Cédric Jousse a poursuivi son projet de fin d'études par un stage de DEA que j'ai également co-encadré avec Nicole Vincent.

1.5.2 Rayonnement scientifique

Je détaille ici mon insertion dans la communauté scientifique : implication dans l'association Specif, activités éditoriales et participations à l'organisation de manifestations, relectures et expertises, participations aux GDR et séminaires, ainsi que mes principales collaborations.

Implication dans l'association Specif

L'association Specif (Société des Personnels Enseignants et Chercheurs en Informatique de France) est la société savante française pour l'enseignement supérieur et la recherche en informatique. L'objectif de Specif est de faire connaître et reconnaître la spécificité de la discipline informatique, dans les domaines de la carrière, de l'enseignement et de la recherche, quel que soit le mode ou le lieu d'exercice (Universités, CNRS, INRIA) et de contribuer à diffuser la connaissance de la discipline informatique et d'en faire mesurer toute la portée scientifique (cf. le site web : <http://www.specif.org>).

Conscient de l'intérêt d'une telle association, je suis adhérent à Specif depuis mes premières années de thèse (2000). Suite à ma nomination sur un poste de Maître de Conférences en 2003, j'ai accepté la charge de correspondant Specif pour l'Université de Strasbourg (jusqu'en 2009, date à laquelle j'ai laissé cette charge à un collègue suite à mon départ en délégation pour Rennes). Le rôle du correspondant est d'assurer une transmission des informations depuis l'association vers ses membres et les collègues de la discipline, et inversement de faire remonter les informations locales au CA de Specif. Il s'agit également d'effectuer le suivi local des adhésions, de solliciter de nouvelles adhésions et de répondre aux questions des collègues sur l'association.

Plus récemment, j'ai renforcé mon implication dans l'association en rejoignant en 2008 son conseil d'administration composé d'une vingtaine de membres. En tant que membre du CA, je suis impliqué dans le pilotage de l'association : participation aux réunions (au moins 6 par an au CNAM à Paris) et aux groupes de travail, définition et suivi des actions telles que congrès annuel, journées thématiques, bulletins et lettres, etc. Une de mes premières actions au sein du CA a été la rédaction d'un aide-mémoire d'une dizaine de pages pour l'organisation du congrès annuel, dont j'ai coordonné l'organisation en 2008 à Strasbourg.

Cette édition 2008 du congrès Specif s'est tenue à l'Université de Strasbourg du 9 au 11 janvier 2008 dans les locaux de l'UFR de Mathématique et d'Informatique. Il a réuni environ 80 participants (dont plusieurs collègues européens) sur le thème de la dimension européenne dans la recherche et l'enseignement en informatique. Plus de détails sont donnés sur le site web de la manifestation : <http://specif2008.u-strasbg.fr>.

La responsabilité de l'organisation de ce congrès s'est traduite par les tâches suivantes : préparation du dossier de candidature, discussion des thèmes et invités avec le CA de Specif, définition du budget et recherche de soutiens financiers, élaboration du programme scientifique et social, animation du comité d'organisation,

gestion des invités et de la logistique, suivi administratif et financier, accueil des participants et gestion de la manifestation, rédaction d'un compte-rendu publié dans le bulletin Specif.

Activités éditoriales et organisations de manifestations

Edition du numéro spécial : Machine Learning in Image Processing

A l'initiative d'Olivier Lezoray (Professeur, Université de Caen – Basse Normandie), et en collaboration avec Christophe Charrier (MCF, Université de Caen – Basse Normandie) et Hubert Cardot (Professeur, Université de Tours), j'ai édité un numéro spécial [A1] de la revue EURASIP Journal on Advances in Signal Processing (JASP) sur le thème "Machine Learning in Image Processing" (MLIP), en relation avec mes travaux de recherche. Une description de ce numéro spécial et de ses thématiques est donnée à l'adresse <http://www.hindawi.com/journals/asp/volume-2008/si.mlip.html>.

Les actions menées dans le cadre de cette responsabilité éditoriale ont concerné la préparation et la diffusion de l'appel à contributions, la réception et l'étude des articles soumis, le choix des relecteurs, le suivi du processus éditorial et enfin la rédaction de l'éditorial.

Participations à des comités de programme et d'organisation

J'ai participé à plusieurs comités de programme de congrès internationaux ou nationaux en lien avec mes thématiques de recherche.

acronyme	libellé	lieu et année
EGC	Journées francophones Extraction et Gestion de Connaissances	Strasbourg, 2009
IMAGAPP	International Conference on Imaging Theory and Applications	Lisbonne, 2009
ISMM	International Symposium on Mathematical Morphology	Groningen, 2009
EGC	Journées francophones Extraction et Gestion de Connaissances	Hammamet, 2010
ICISP	International Conference on Image and Signal Processing	Trois-Rivières, 2010
IMAGAPP	International Conference on Imaging Theory and Applications	Angers, 2010

J'ai été également membre de plusieurs comités d'organisation de manifestations scientifiques se déroulant en Alsace sous l'égide du LSIIT et de l'UDS, en lien avec mes thématiques de recherche.

acronyme	libellé	lieu et année
SAGEO	International Conference on Spatial Analysis and Geomatics	Strasbourg, 2006
ORASIS	Congrès Jeunes Chercheurs en Vision par Ordinateur	Obernai, 2007
EGC	Journées francophones Extraction et Gestion de Connaissances	Strasbourg, 2009

Activités de relecture et d'expertise

revues internationales et nationales

ACM Computing Reviews (ACM)	5 évaluations
IEEE Transactions on Image Processing (IEEE)	4 évaluations
IEEE Transactions on Signal Processing (IEEE)	1 évaluation
Pattern Recognition Letters (Elsevier)	6 évaluations
Image and Vision Computing (Elsevier)	5 évaluations
Digital Signal Processing (Elsevier)	1 évaluation
Multimedia Tools and Applications (Kluwer)	3 évaluations
Journal of Mathematical Imaging and Vision (Springer)	1 évaluation
Eurasip Journal on Advances in Signal Processing (Hindawi)	1 évaluation
International Journal of Imaging Systems and Technology (Wiley)	1 évaluation
Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées	1 évaluation
Remote Sensing Journal	1 évaluation
Traitement du Signal	1 évaluation

conférences internationales*

Advanced Concepts for Intelligent Vision Systems (ACIVS) 2003 à 2006	5 évaluations
Int. Symposium on Mathematical Morphology (ISMM) 2005 et 2007	7 évaluations
Int. Conference on Image Processing (ICIP) 2006 à 2009	15 évaluations
Int. Geoscience and Remote Sensing Symposium (IGARSS) 2008 à 2010	9 évaluations
Int. Conference on Acoustics, Speech and Signal Processing (ICASSP) 2009 à 2010	10 évaluations

conférences nationales*

Colloque International de Géomatique et d'Analyse Spatiale (SAGEO) 2006	4 évaluations
Atelier Fouille de Données Complexes, Congrès EGC, 2007	1 évaluation

* Par ailleurs, j'ai systématiquement évalué des articles soumis aux conférences dont j'étais membre du comité de programme.

expertises nationales et régionales

Région Bretagne (2005) pour les bourses de thèse	1 expertise
ANR Concepts Systèmes et Outils pour la Sécurité Globale (CSOSG) (2007,2008)	3 expertises

Participations à des GDR et séminaires invités**Participation à des GDR**

Je suis membre de plusieurs GDR (Groupement De Recherche) en lien avec mes thématiques de recherche :

GDR ISIS (Information – Signal – Image – viSion) pour l'ensemble de mes travaux, que ce soit sur les aspects fondamentaux en morphologie mathématique ou les applications dans différents domaines,

GDR I3 (Information – Interaction – Intelligence) pour mes travaux relatifs à l'intégration de connaissances dans les traitements morphologiques et à notre plate-forme d'analyse d'image,

GDR SIGMA (Systèmes d'Information Géographique – Méthodologies et Applications) pour les travaux et applications en télédétection.

A ce titre, j'ai été orateur lors de plusieurs réunions thématiques :

Date	GDR	Thème	Titre de l'exposé
avril 2006	ISIS	Imagerie multivariée	<i>Morphologie mathématique multivariée</i>
novembre 2007	I3 & SIGMA	JIGOT ¹	<i>Profils morphologiques multispectraux pour la classification d'images de télédétection</i>
mars 2008	ISIS & I3	SCATI ²	<i>Segmentation morphologique guidée par des exemples</i>

¹ Journées sur l'Information Géographique et l'Observation de la Terre

² Groupe de Travail : Systèmes Complexes pour l'Analyse et le Traitement d'Images

Remarque : J'ai également présenté mes travaux doctoraux lors de deux réunions thématiques du GDR ISIS en indexation multimédia (2001 et 2003).

En outre, je suis membre de l'EPML IRMC (Imagerie et Robotique Médicale et Chirurgicale) et du Pôle de Compétitivité à vocation mondiale "Innovations Thérapeutiques", mais mon implication dans ce programme est pour l'instant très limitée. Toutefois, comme indiqué dans la section perspectives de mes travaux de recherche, l'imagerie médicale est un des domaines d'applications sur lesquels je pourrais concentrer mes efforts à moyen terme, dans le cadre d'IRMC.

Séjours et séminaires invités

J'ai effectué un séjour de 3 mois d'avril à juin 2008 (à raison de 3 jours par semaine) en tant que chercheur invité à l'Université de Freiburg-im-Breisgau (Allemagne), dans l'équipe de Traitement d'Image et de Reconnaissance des Formes (LMB) du Professeur Hans Burkhardt. Cette collaboration s'est poursuivie en 2008–2009 sous forme de visites régulières du laboratoire. Elle concerne les thèmes des invariants morphologiques et de l'utilisation de la morphologie mathématique en imagerie biologique 3-D.

Depuis septembre 2009, je bénéficie d'un accueil en délégation INRIA à l'IRISA (Rennes), au sein de l'équipe TEXMEX spécialisée dans l'exploitation des documents multimédia. Cette collaboration devrait me permettre de m'intéresser à la segmentation morphologique de séquences vidéo, aux descripteurs morphologiques locaux pour l'indexation multimédia, et enfin à l'application de la morphologie mathématique à des données non-image.

Outre ces séjours de recherche, j'ai également été invité à donner plusieurs séminaires relatifs à mes travaux de recherche :

Date	Lieu	Titre de l'exposé
janvier 2005	Laboratoire ERIC, Lyon	<i>Introduction à la morphologie mathématique : principaux outils et applications</i>
mars 2005	Laboratoire L3i, La Rochelle	<i>Génération de descripteurs spatiaux et temporels par analyse multi-images</i>
février 2008	Institut ISPC du JRC, Ispra (Italie)	<i>Multivariate and supervised approaches for mathematical morphology in remote sensing</i>
avril 2008	Université de Freiburg (Allemagne)	<i>Knowledge-driven morphological approaches for image segmentation and object detection</i>
mai 2008	Université de Freiburg (Allemagne)	<i>Scale-space features with mathematical morphology</i>
avril 2009	Equipe LSIT-MIV, Strasbourg	<i>Classification, segmentation et morphologie mathématique</i>

Remarque : J'ai également donné des séminaires dans plusieurs laboratoires (Tours en 1999, 2002, 2003 ; Rouen, Poitiers, Caen en 2003 ; Strasbourg en 2004) dans le but de présenter mes travaux doctoraux.

Collaborations scientifiques

J'attache une grande importance à l'exercice d'une activité de recherche en collaboration avec d'autres chercheurs, que ce soit dans le cadre de projets mono ou pluridisciplinaires. Je dresse ici une liste de mes principales collaborations internationales, nationales, et locales, en précisant la nature et le thème de chaque collaboration. J'indique notamment les collaborations qui se sont traduites par des séjours de recherche invités (de quelques jours à quelques mois), des publications cosignées (cf. section 1.6), ou des co-encadrements d'étudiants en projet de master (non cités dans la section 1.5.1). Pour être tout à fait exhaustive, cette liste doit être complétée par les co-encadrements (section 1.5.1) et les projets (section 1.5.3) passés ou en cours (mais aussi des collaborations en cours qui n'ont encore fait l'objet ni de publication cosignée, ni de co-encadrement, ni de séjour invité, et qui ne sont pas indiquées ici).

collaborations internationales

H. Burkhardt	Université de Freiburg	séjour	<i>invariants morphologiques ; approches morphologiques en imagerie biologique</i>
M. Pesaresi	JRC-IPSC, Ispra	séjour	<i>profils morphologiques multispectraux ; extraction automatique du bâti par télédétection</i>

collaborations nationales

P. Gros	IRISA, Rennes	séjour	<i>approches morphologiques en indexation multimédia</i>
D. Sheeren	INPT Toulouse	publications	<i>détection et classification morphologique de structures (bâti, haies) en télédétection</i>
F. Sèdes	IRIT, Toulouse	publications *	<i>descripteurs spatio-temporels pour l'indexation de séquences vidéo</i>
N. Vincent	LIPADE, Paris	publications *	<i>outils rapides pour l'indexation de séquences vidéo</i>

collaborations locales (hors équipe MIV)

P. Gañarski, A. Blansché	LSIIT-FDBT	encadrements	<i>classification morphologique non-supervisée par algorithmes génétiques</i>
S. Genaud	LSIIT-ICPS	encadrements	<i>mise en œuvre d'algorithmes morphologiques sur les grilles de calcul</i>
A. Puissant	LIVE	publications	<i>approches morphologiques en télédétection ; applications aux espaces urbains et côtiers</i>
V. Agnus	IRCAD (IRMC)	encadrements	<i>segmentation de structures anatomiques du foie par morphologie mathématique floue</i>

* Les collaborations avec Florence Sèdes et Nicole Vincent concernent des travaux menés sur le thème de l'indexation vidéo, dans le prolongement de mes travaux doctoraux.

1.5.3 Activités contractuelles et projets de recherche

Mes travaux de recherche sont souvent liés à des projets de recherche contractualisés avec des partenaires publics (ANR, Oséo-ANVAR, Université) ou privés (CIFRE, collaborations de recherche). Ces projets ont permis d'une part d'étudier l'impact des travaux de recherche et la pertinence des solutions proposées dans des contextes applicatifs réels, et d'autre part d'aider à la réalisation des travaux par l'obtention de financements, notamment pour couvrir les besoins en développement logiciel. Avant de donner la liste exhaustive des projets auxquels j'ai participé, je présente plus en détail les projets où mon implication a été la plus forte.

PELICAN

Conscient des manques de capitalisation et de valorisation des travaux de recherche en traitement d'image, j'ai initié en 2004 le projet PELICAN (décrit en section 3.4) dont le but est d'offrir une plate-forme logicielle pour l'analyse et le traitement des images, découplant complètement le développement des algorithmes de leur utilisation. Plus de 25 personnes (étudiants principalement) ont participé à ce projet.

Les tâches que j'assume en tant que responsable du projet sont variées : animation et suivi du projet (réunions, wiki), encadrement des développeurs et participation au développement, validation des spécifications fonctionnelles et techniques, définition des orientations stratégiques et positionnement dans la communauté scientifique, recherche de partenaires en vue d'une contractualisation.

Oséo-ANVAR

En lien avec mes travaux doctoraux en indexation vidéo, j'ai été sollicité en 2005 par une jeune société de la région pour initier une collaboration de recherche sur le thème de la recherche d'image par le contenu. Dans ce contexte, j'ai donc été amené à définir un programme de recherche sur l'annotation semi-automatique et la recherche d'images couleur par des méthodes morphologiques, projet que nous avons déposé auprès de l'ANRT

afin d'obtenir un financement de thèse CIFRE. Le financement ayant été accordé mais la société partenaire ne pouvant finalement pas assumer ses engagements, j'ai obtenu avec l'aide de la cellule de valorisation de l'Université de Strasbourg un financement d'Oséo-ANVAR en 2006 (et pour une durée de 3 années) sous la forme d'une aide au transfert pour permettre au doctorant de continuer ses travaux doctoraux et de soutenir sa thèse en juillet 2008.

En tant que responsable de ce projet, j'ai défini le programme de recherche ainsi que les spécifications fonctionnelles et techniques du système, encadré le doctorant, initié des actions de communication autour de notre projet et participé à la recherche de partenaires dans le but d'une valorisation ultérieure, et enfin assuré le suivi administratif et financier du projet.

CIFRE RBS

Suite à notre partenariat dans le cadre du projet Oséo-ANVAR décrit précédemment, la cellule de valorisation de l'UDS m'a contacté pour répondre à une demande de collaboration de la part de la société RBS dans le domaine des moteurs de recherche d'image. Afin de répondre à la fois aux besoins de la société et aux objectifs scientifiques du laboratoire, j'ai proposé un programme de recherche portant sur la structuration automatique d'ensembles d'images et de séquences vidéo. Pour mener à bien ce programme associant traitement d'image et fouille de données, j'ai demandé à Pierre Gañarski (MCF HDR, équipe FDBT) d'y participer et d'assurer pendant les deux premières années la direction de la thèse qui a débuté fin 2007 sur un financement CIFRE. Mon rôle en tant qu'initiateur puis coresponsable du projet avec Pierre Gañarski a porté sur la définition du sujet de recherche, la négociation du contrat de collaboration avec la société partenaire, l'encadrement du doctorant, et enfin le suivi administratif et financier.

— Synthèse des projets et contrats de recherche —

responsabilités complètes de projets ¹			
Plate-forme PELICAN	Projet collaboratif interne au LSIIT	depuis 2004	
Annotation d'image	Aide au transfert Oséo-ANVAR	2006 – 2009	144 000 €
Structuration de vidéos ^a	CIFRE RBS	2007 – 2010	36 000 €
responsabilités scientifiques partielles ²			
Vidéosurveillance ferroviaire ^b	Collaboration de recherche Faiveley	2002 – 2003	29 000 €
FoDoMuST	ACI Masses de Données	2004 – 2007	434 000 €
participations à des projets contractualisés ³			
Ecosgil	ANR Jeunes Chercheurs	2005 – 2008	52 000 €
Ecoulements gravitaires	BQR UDS avec IPGS et LIVE	2008 – 2009	15 000 €

^a La responsabilité du contrat relatif à la convention CIFRE avec RBS est partagée avec Pierre Gañarski (MCF HDR, équipe FDBT) avec qui je co-encadre la thèse de Jonathan Weber.

^b J'ai été impliqué dans ce projet lors de mon année d'ATER à l'Université de Tours. Il ne concerne donc pas les thématiques de recherche que j'ai développées au LSIIT et présentées dans ce document.

¹ La responsabilité complète d'un projet inclut les tâches suivantes : initiation et définition du projet, négociation avec le partenaire, gestion financière et suivi du budget, responsabilité scientifique et animation du projet, suivi administratif, encadrement des personnels affectés au projet, et les tâches inhérentes à la responsabilité scientifique partielle et à la participation à un projet qui sont décrites ci-dessous.

² La responsabilité scientifique partielle d'un projet inclut les tâches suivantes : responsabilité et animation d'un ou plusieurs axes scientifiques du projet, participation au suivi administratif (pas de gestion financière), et les tâches inhérentes à la participation à un projet qui sont décrites ci-dessous.

³ La participation à un projet inclut les tâches suivantes : forte implication voire responsabilité des actions directement relatives à mes thématiques de recherche, participation à la définition du projet, à l'encadrement des développeurs (le cas échéant), à la réalisation du projet et au développement.

1.6 Liste de publications

Les publications relevant d'activités d'encadrement de recherche sont indiquées en soulignant les noms des auteurs ayant fait l'objet d'un (co-)encadrement.

Les facteurs d'impact (pour les revues) et taux de sélection (pour les conférences) sont donnés entre parenthèses lorsqu'ils sont connus. Sont également précisés les cas où la sélection s'est faite sur résumé.

Pour chaque type de publication, une barre horizontale sépare les publications relatives aux travaux présentés au titre de l'HDR de celles liées à mon expérience doctorale.

— Synthèse des publications —

4	Chapitres d'ouvrage et numéros spéciaux édités
10	Articles dans des Revues Internationales
5	Articles dans des Revues Nationales
27	Communications dans des Conférences Internationales avec Comité de Lecture et Actes
16	Communications dans des Conférences Nationales avec Comité de Lecture et Actes
2	Mémoires
3	Soumissions (Chapitres, Articles, ou Communications) dont 2 en révision

Chapitres d'ouvrage et numéros spéciaux édités

- [A4] **S. Lefèvre** : A new approach for unsupervised classification in image segmentation. *Dans* : Henri Briand, Fabrice Guillet, Gilbert Ritschard et Djamel Zighed, éditeurs : *Advances in Knowledge Discovery and Management*, Studies in Computational Intelligence. Springer-Verlag, 2009. (*Accepté pour publication en Octobre 2009, 19 pages*).
- [A3] E. Aptoula et **S. Lefèvre** : Multivariate mathematical morphology applied to colour image analysis. *Dans* : Jocelyn Chanussot, Kacem Chehdi et Christophe Collet, éditeurs : *Multivariate image processing : methods and applications*, pages 303–337. Hermès – John Wiley, 2009.
- [A2] **S. Lefèvre** : Image features from morphological scale-spaces. *Dans* : Dacheng Tao, Dong Xu et Xuelong Li, éditeurs : *Semantic Mining from Multimedia Databases*, pages 32–79. Idea Group Publishing, 2009.
- [A1] O. Lezoray, C. Charrier, H. Cardot et **S. Lefèvre**, éditeurs. *special issue on Machine Learning in Image Processing*. EURASIP Journal on Advances in Signal Processing, 2008.
-

Articles dans des revues internationales

- [B10] E. Aptoula et **S. Lefèvre** : Morphological description of colour images for content-based image retrieval. *IEEE Transactions on Image Processing*, 18(11):2505–2517, Novembre 2009. (*Facteur d'impact 2008 : 3.32*).
- [B9] B. Perret, **S. Lefèvre** et C. Collet : A robust hit-or-miss transform for template matching in very noisy astronomical images. *Pattern Recognition*, 42(11):2470–2480, Novembre 2009. (*Facteur d'impact 2008 : 3.28*).
- [B8] E. Aptoula et **S. Lefèvre** : On the morphological processing of hue. *Image and Vision Computing*, 27(9):1394–1401, Août 2009. (*Facteur d'impact 2008 : 1.50*).
- [B7] E. Aptoula, **S. Lefèvre** et C. Ronse : A hit-or-miss transform for multivariate images. *Pattern Recognition Letters*, 30(8):760–764, Juin 2009. (*Facteur d'impact 2008 : 1.56*).
- [B6] **S. Lefèvre** : Beyond morphological size-distribution. *Journal of Electronic Imaging*, 18(1):013010 (10 pages), Mars 2009. (*Facteur d'impact 2008 : 0.56*).

- [B5] E. Aptoula et **S. Lefèvre** : α -trimmed lexicographical extrema for pseudo-morphological image analysis. *Journal of Visual Communication and Image Representation*, 19(3):165–174, Avril 2008. (Facteur d'impact 2008 : 1.34).
- [B4] E. Aptoula et **S. Lefèvre** : On lexicographical ordering in multivariate mathematical morphology. *Pattern Recognition Letters*, 29(2):109–118, Janvier 2008. (Facteur d'impact 2008 : 1.56).
- [B3] E. Aptoula et **S. Lefèvre** : A comparative study on multivariate mathematical morphology. *Pattern Recognition*, 40(11):2914–2929, Novembre 2007. (Facteur d'impact 2008 : 3.28).
-
- [B2] **S. Lefèvre** et N. Vincent : Efficient and robust shot change detection. *Journal of Real Time Image Processing*, 2(1):23–34, Octobre 2007. (Facteur d'impact non disponible, revue créée en 2006).
- [B1] **S. Lefèvre**, J. Holler et N. Vincent : A review of real-time segmentation of uncompressed video sequences for content-based search and retrieval. *Real-Time Imaging*, 9(1):73–98, 2003. (Facteur d'impact 2006 : 0.99, revue arrêtée en 2006).

Articles dans des revues nationales

- [C5] A. Puissant, **S. Lefèvre**, R. Desguée et F. Levoy : Cartographie et suivi de l'évolution des schorres de la baie du mont-saint-michel (1986–2006) à partir d'images satellites haute résolution. *Photo-Interprétation*, 2008(3/4):3–11, 2008.
- [C4] D. Sheeren, **S. Lefèvre** et J. Weber : La morphologie mathématique binaire pour l'extraction automatique des bâtiments dans les images THRS. *Revue Internationale de Géomatique, Numéro Thématique : Informations géographiques : structuration, extraction et utilisation*, 17(3/4):333–352, 2007.
-
- [C3] **S. Lefèvre** et N. Vincent : Outils pour la recherche d'extraits pertinents dans les séquences vidéo. *Technique et Science Informatiques*, 25(10):1311–1340, 2006.
- [C2] **S. Lefèvre** et N. Vincent : Apport de l'espace teinte-saturation-luminance pour la segmentation spatiale et temporelle. *Traitement du Signal*, 23(1):59–77, 2006.
- [C1] **S. Lefèvre** et F. Sèdes : Indexation de séquences vidéo. indices liés au temps. *Document Numérique, Numéro thématique : Temps et Documents*, 8(4):41–48, 2004.

Communications dans des conférences internationales

- [D27] S. Aksoy, B. Ozdemir, S. Eckert, F. Kayitakire, M. Pesaresi, O. Aytekin, C. Borel, J. Cech, E. Christophe, S. Duzgun, A. Erener, K. Ertugay, E. Hussain, J. Inglada, **S. Lefèvre**, O. Ok, D. Koc San, R. Sara, J. Shan, J. Soman, I. Ulusoy et R. Witz : Performance evaluation of building detection and digital surface model extraction algorithms : Outcomes of the PRRS 2008 algorithm performance contest. *Actes de : IAPR International Workshop on Pattern Recognition in Remote Sensing (PRRS)*, Tampa, Etats-Unis, Décembre 2008.
- [D26] B. Perret, **S. Lefèvre**, C. Collet et B. Vollmer : Astronomical object detection with a robust hit-or-miss transform. *Actes de : EURASIP European Conference on Signal Processing (EUSIPCO)*, Lausanne, Suisse, Août 2008. (Taux de sélection : 50 %).
- [D25] J. Weber et **S. Lefèvre** : A multivariate hit-or-miss transform for conjoint spatial and spectral template matching. *Actes de : IEEE International Conference on Image and Signal Processing*, volume 5099 de *Lecture Notes in Computer Science*, pages 226–235, Cherbourg, France, Juillet 2008. Springer-Verlag. (Taux de sélection pour les présentations orales : 25 %).
- [D24] A. Puissant, **S. Lefèvre** et J. Weber : Coastline extraction in VHR imagery using mathematical morphology with spatial and spectral knowledge. *Actes de : ISPRS Congress*, Pékin, Chine, Juillet 2008. (Sélection sur résumé).

- [D23] E. Aptoula et **S. Lefèvre** : A basin morphology approach to colour image segmentation by region merging. *Actes de : Asian Conference on Computer Vision (ACCV)*, volume 4843 de *Lecture Notes in Computer Science*, pages 935–944, Tokyo, Japon, Novembre 2007. Springer-Verlag. (Taux de sélection pour les présentations orales : 8 %).
- [D22] E. Aptoula et **S. Lefèvre** : On morphological color texture characterization. *Actes de : International Symposium on Mathematical Morphology (ISMM)*, pages 153–164, Rio de Janeiro, Brésil, Octobre 2007. (Taux de sélection : 72 %).
- [D21] E. Aptoula et **S. Lefèvre** : Pseudo multivariate morphological operators based on α -trimmed lexicographical extrema. *Actes de : IEEE International Symposium on Image and Signal Processing and Analysis (ISPA)*, pages 367–372, Istanbul, Turquie, Septembre 2007. (Taux de sélection : 62 %).
- [D20] S. Derivaux, **S. Lefèvre**, C. Wemmert et J.J. Korczak : On machine learning in watershed segmentation. *Actes de : IEEE Workshop on Machine Learning in Signal Processing (MLSP)*, pages 187–192, Thessalonique, Grèce, Août 2007. (Taux de sélection : 49 %).
- [D19] **S. Lefèvre** : Knowledge from markers in watershed segmentation. *Actes de : International Conference on Computer Analysis of Image and Patterns (CAIP)*, volume 4673 de *Lecture Notes in Computer Science*, pages 579–586, Vienne, Autriche, Août 2007. Springer-Verlag. (Taux de sélection pour les présentations orales : 16 %).
- [D18] **S. Lefèvre** : Extending morphological signatures for visual pattern recognition. *Actes de : IAPR International Workshop on Pattern Recognition in Information Systems (PRIS)*, pages 79–88, Madère, Portugal, Juin 2007. (Taux de sélection pour les présentations orales : 31 %).
- [D17] **S. Lefèvre**, J. Weber et D. Sheeren : Automatic building extraction in VHR images using advanced morphological operators. *Actes de : IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Paris, France, Avril 2007. (Sélection sur résumé, Taux de sélection pour les présentations orales : 42 %).
- [D16] E. Aptoula et **S. Lefèvre** : Spatial morphological covariance applied to texture classification. *Actes de : International Workshop on Multimedia Content Representation, Classification and Security (IWMRCS)*, volume 4105 de *Lecture Notes in Computer Science*, pages 522–529, Istanbul, Turquie, Septembre 2006. Springer-Verlag. (Taux de sélection : 52 %).
- [D15] E. Aptoula, **S. Lefèvre** et C. Collet : Mathematical morphology applied to the segmentation and classification of galaxies in multispectral images. *Actes de : EURASIP European Conference on Signal Processing (EUSIPCO)*, Florence, Italie, Septembre 2006. (Taux de sélection : 63 %).
- [D14] S. Derivaux, **S. Lefèvre**, C. Wemmert et J.J. Korczak : Watershed segmentation of remotely sensed images based on a supervised fuzzy pixel classification. *Actes de : IEEE International Geosciences and Remote Sensing Symposium (IGARSS)*, Denver, Etats-Unis, Juillet 2006. (Sélection sur résumé, Taux de sélection : 45 %).
-
- [D13] **S. Lefèvre** et N. Vincent : Caption localisation in video sequences by fusion of multiple detectors. *Actes de : International Conference on Document Analysis and Recognition (ICDAR)*, pages 106–110, Séoul, Corée du Sud, Septembre 2005. (Taux de sélection : 79 %).
- [D12] **S. Lefèvre** et N. Vincent : Real time multiple object tracking based on active contours. *Actes de : International Conference on Image Analysis and Recognition (ICIAR)*, volume 3212 de *Lecture Notes in Computer Science*, pages 606–613, Porto, Portugal, Septembre 2004. Springer-Verlag. (Taux de sélection : 66 %).
- [D11] **S. Lefèvre**, E. Bouton, T. Brouard et N. Vincent : A new way to use hidden markov models for object tracking in video sequences. *Actes de : IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 117–120, Barcelone, Espagne, Septembre 2003. (Taux de sélection : 44 %).
- [D10] **S. Lefèvre**, C. Jeusse et N. Vincent : A comparison of line detectors for image background modelling. *Actes de : IAPR International Workshop on Pattern Recognition in Information Systems (PRIS)*, pages 185–190, Angers, France, Avril 2003.

- [D9] **S. Lefèvre**, T. Foissotte et N. Vincent : Detection and tracking of small objects by details removal. *Actes de : International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, Londres, Royaume-Uni, Avril 2003. (Taux de sélection : 68 %).
- [D8] **S. Lefèvre**, J.P. Gérard, A. Piron et N. Vincent : An extended snake model for real-time multiple object tracking. *Actes de : International Workshop on Advanced Concepts for Intelligent Vision Systems (ACIVS)*, pages 268–275, Gand, Belgique, Septembre 2002. (Taux de sélection : 70 %).
- [D7] **S. Lefèvre**, B. Maillard et N. Vincent : A two level classifier process for audio segmentation. *Actes de : IAPR International Conference on Pattern Recognition (ICPR)*, volume 3, pages 891–894, Québec, Canada, Août 2002. (Taux de sélection : 64 %).
- [D6] **S. Lefèvre**, B. Maillard et N. Vincent : 3 class segmentation for analysis of football audio sequences. *Actes de : IEEE International Conference on Digital Signal Processing*, volume 2, pages 975–978, Santorin, Grèce, Août 2002. (Taux de sélection : 80 %).
- [D5] **S. Lefèvre**, C. Dixon, C. Jousse et N. Vincent : A local approach for fast line detection. *Actes de : IEEE International Conference on Digital Signal Processing*, volume 2, pages 1109–1112, Santorin, Grèce, Août 2002. (Taux de sélection : 80 %).
- [D4] **S. Lefèvre**, L. Mercier, V. Tiberghien et N. Vincent : Multiresolution color image segmentation applied to background extraction in outdoor images. *Actes de : IS&T European Conference on Color in Graphics, Image and Vision (CGIV)*, pages 363–367, Poitiers, France, Avril 2002. (Taux de sélection : 32 %).
- [D3] **S. Lefèvre**, C. Fluck, B. Maillard et N. Vincent : A fast snake-based method to track football players. *Actes de : IAPR International Workshop on Machine Vision Applications (MVA)*, pages 501–504, Tokyo, Japon, Novembre 2000. (Sélection sur résumé).
- [D2] A.S. Capelle, O. Alata, C. Fernandez, **S. Lefèvre** et J.C. Ferrié : Unsupervised segmentation for automatic detection of brain tumors in MRI. *Actes de : IEEE International Conference on Image Processing (ICIP)*, volume 1, pages 613–616, Vancouver, Canada, Septembre 2000.
- [D1] **S. Lefèvre**, J. Holler et N. Vincent : Real time temporal segmentation of compressed and uncompressed dynamic colour image sequences. *Actes de : International Workshop on Real Time Image Sequence Analysis*, pages 56–62, Oulu, Finlande, Août 2000. (Taux de sélection : 54 %).

Communications dans des conférences nationales

Certaines communications dans des conférences nationales correspondent à des traductions en français de communications présentées dans des conférences internationales, dans un but de diffusion des résultats auprès de la communauté scientifique nationale. Ces communications sont indiquées en précisant entre parenthèses les références aux communications en conférences internationales.

- [E16] **S. Lefèvre** : Segmentation par ligne de partage des eaux avec marqueurs spatiaux et spectraux. *Actes de : Colloque GRETSI sur le Traitement du Signal et des Images*, Dijon, France, Septembre 2009. (Taux de sélection : 75 %).
- [E15] **S. Lefèvre** : Une nouvelle approche pour la classification non supervisée en segmentation d'image. *Revue des Nouvelles Technologies de l'Information, Numéro Thématique Extraction et Gestion des Connaissances (EGC'2009)*, volume RNTI-E-15, pages 127–138. Cépaduès Editions, Strasbourg, France, Janvier 2009. (Taux de sélection pour les présentations orales : 21 %).
- [E14] S. Derivaux, G. Forestier, C. Wemmert et **S. Lefèvre** : Construction de détecteurs d'objets urbains à partir d'une ontologie. *Actes de : Atelier Extraction de Connaissance à partir d'Images (ECOI), Journées EGC*, Sophia-Antipolis, France, Janvier 2008.
- [E13] S. Derivaux, **S. Lefèvre**, C. Wemmert et J.J. Korczak : Segmentation par ligne de partage des eaux basées sur des connaissances texturales. *Actes de : Colloque GRETSI sur le Traitement du Signal et des Images*, Troyes, France, Septembre 2007.

- [E12] E. Aptoula et **S. Lefèvre** : Pseudo-opérateurs morphologiques multivariés basés sur les extrema lexicographiques α -tronqués. *Actes de : Congrès Francophone de Vision par Ordinateur ORASIS'2007*, Obernai, France, Juin 2007. (Version française de [D21]).
- [E11] S. Derivaux, C. Wemmert, **S. Lefèvre** et J.J. Korczak : Paramétrisation de méthodes de segmentation par utilisation de connaissances et approche génétique. *Actes de : Atelier Extraction de Connaissance à partir d'Images (ECOI), Journées EGC*, Namur, Belgique, Janvier 2007. (Version française de [D20]).
- [E10] S. Derivaux, **S. Lefèvre**, C. Wemmert et J.J. Korczak : Apport d'une classification non supervisée floue à la segmentation par ligne de partage des eaux. *Actes de : International Conference on Spatial Analysis and Geomatics (SAGEO)*, Strasbourg, France, Septembre 2006.
- [E9] J. Weber, **S. Lefèvre** et D. Sheeren : Détection des bâtiments dans les images THRS avec la morphologie mathématique. *Actes de : International Conference on Spatial Analysis and Geomatics (SAGEO)*, Strasbourg, France, Septembre 2006. (Article court, version française de [D17]).
-
- [E8] **S. Lefèvre** et N. Vincent : Comparaison des détecteurs de lignes pour la modélisation de l'environnement. *Actes de : Journées Compression et Représentation des Signaux Audiovisuels (CORESA)*, pages 67–72, Rennes, France, Novembre 2005. (Version française de [D10]).
- [E7] **S. Lefèvre** et N. Vincent : Contours actifs pour le suivi d'objet en temps-réel : multi-topologies et multi-résolutions. *Actes de : Colloque GRETSI sur le Traitement du Signal et des Images*, pages 1125–1128, Louvain-La-Neuve, Belgique, Septembre 2005. (Version française de [D12]).
- [E6] **S. Lefèvre** et N. Vincent : Extraction d'indices spatiaux et temporels dans des séquences vidéo couleur. Dans : F. Cloppet, J.M. Petit et N. Vincent, éditeurs : *Revue des Nouvelles Technologies de l'Information, Numéro Thématique Extraction des connaissances : état et perspectives, volume RNTI-E-5*, pages 247–258. Cépaduès Editions, Paris, France, Janvier 2005. (Présenté initialement à l'Atelier Fouille de données complexes dans un processus d'extraction des connaissances, Journées Extraction et Gestion des Connaissances).
- [E5] **S. Lefèvre**, C. L'Orphelin et N. Vincent : Extraction multicritère de texte incrusté dans les séquences vidéo. *Actes de : Colloque International Francophone sur l'Ecrit et le Document (CIFED)*, La Rochelle, France, Juin 2004. (Version française de [D13]).
- [E4] **S. Lefèvre**, B. Maillard et N. Vincent : Deux niveaux et deux outils d'analyse pour une meilleure segmentation de données audio. *Actes de : Colloque GRETSI sur le Traitement du Signal et des Images*, Paris, France, Septembre 2003. (Version française de [D7]).
- [E3] **S. Lefèvre**, N. Vincent et C. Proust : Architecture et outils pour la recherche d'évènements dans les séquences vidéo. *Actes de : Congrès INFORSID en Bases de Données et Recherche d'Information*, Nancy, France, Juin 2003.
- [E2] **S. Lefèvre**, C. Fluck, B. Maillard et N. Vincent : Un modèle de contour actif pour le suivi rapide d'objets en mouvement. Application au suivi de joueurs de football. *Actes de : Colloque GRETSI sur le Traitement du Signal et des Images*, Toulouse, France, Septembre 2001. (Version française de [D3]).
- [E1] **S. Lefèvre**, J. Holler et N. Vincent : Segmentation temporelle de séquences d'images en couleur compressées et non-compressées en temps réel. *Actes de : Congrès Francophone de Vision par Ordinateur ORASIS'2001*, pages 329–338, Cahors, France, Juin 2001. (Version française de [D1]).

Mémoires

-
- [F2] **S. Lefèvre** : *Détection d'évènements dans une séquence vidéo*. Thèse de doctorat, Université de Tours, Décembre 2002. (219 pages).
- [F1] **S. Lefèvre** : *Elaboration et validation d'un outil automatique de segmentation et de reconstruction volumique sur imagerie IRM de tumeurs cérébrales*. Mémoire de DEA, Université de Technologie de Compiègne, Septembre 1999. (74 pages).

Soumissions d'articles, chapitres, communications

- [G3] S. Derivaux, G. Forestier, C. Wemmert et **S. Lefèvre** : Supervised segmentation using machine learning and evolutionary computation. *Pattern Recognition Letters*, 2009. (Soumis en Juillet 2009, 20 pages).
- [G2] J. Weber, **S. Lefèvre** et P. Gañarski : Video object mining : Issues and perspectives. *International Journal of Multimedia Tools and Applications*, 2009. (Soumis en Mars 2009, 13 pages, Révisé en Septembre 2009).
-
- [G1] **S. Lefèvre** et N. Vincent : A two level strategy for audio segmentation. *Digital Signal Processing*, 2009. (Soumis en Juillet 2005, 15 pages, Révision mineure en Mai 2009).

Chapitre 2

Travaux fondamentaux

Dans ce chapitre, nous rappelons les principaux concepts et outils de la morphologie mathématique (section 2.1) et détaillons nos contributions sous l’angle de leurs aspects fondamentaux : l’adaptation des approches morphologiques aux images multivaluées (section 2.2), et l’intégration de connaissances au sein de ces approches (section 2.3). L’application de ces travaux dans différents domaines fera l’objet du prochain chapitre.

2.1 Introduction à la morphologie mathématique

Les travaux présentés dans ce mémoire étant essentiellement relatifs à la morphologie mathématique et à son utilisation dans le contexte de l’analyse et du traitement des images, nous rappelons dans cette section les principes essentiels de cette théorie, ainsi que les principaux opérateurs et méthodes de traitement d’image qui en découlent, en nous focalisant notamment sur ceux sur lesquels nous avons basé l’essentiel de nos travaux. La plupart des notations utilisées dans la suite de ce mémoire sont également introduites dans cette section introductive. Le lecteur averti pourra directement se rendre à la page 48 où débute la description de nos contributions.

2.1.1 Introduction

La morphologie mathématique est une théorie introduite en 1964 par deux chercheurs français, Georges Matheron et Jean Serra (le second ayant effectué sa thèse de doctorat sous la direction du premier). Dans le contexte de la science des matériaux (et plus précisément de l’étude des milieux poreux), ceux-ci ont élaboré un formalisme complet pour l’analyse des images binaires à l’aide de la théorie des ensembles. Ils ont ainsi introduit une grande partie des opérateurs morphologiques actuels et leurs principaux ouvrages sur le sujet [Mat75, Ser82, Ser88] sont considérés comme précurseurs du domaine.

Dans ce formalisme, les images sont analysées au travers des structures spatiales qu’elles contiennent. Pour cela, on définit un voisinage ou motif de recherche (appelé élément structurant) et on parcourt l’image en observant en chaque pixel la réponse fournie par un opérateur donné, en considérant la zone de recherche représentée par l’élément structurant. Le résultat obtenu dépendra donc de l’opérateur utilisé mais aussi de l’élément structurant considéré. À partir de ce principe somme toute assez simple, il est possible de définir les opérateurs morphologiques de base (cf. section 2.1.3). Par exemple, l’érosion consiste à ne conserver un pixel en blanc que si l’ensemble des voisins sont également blancs. De façon duale, la dilatation produit un pixel blanc dès qu’un de ses voisins est blanc. Plus généralement, l’érosion et la dilatation peuvent être vues respectivement comme la recherche du minimum et du maximum local (i.e. dans le voisinage du pixel considéré). Ces opérateurs peuvent à leur tour être utilisés pour formuler des opérateurs plus complexes, permettant ainsi de résoudre des problèmes variés en traitement d’image : filtrage, détection de contours, segmentation, description de texture, etc.

Une vingtaine d’années a été nécessaire pour que cette théorie s’affranchisse du cadre ensembliste et soit formulée avec la théorie des treillis [Ser88, HR90, RH91]. Cette extension a permis de faciliter le traitement des images à niveaux de gris, mais aussi d’appliquer la morphologie mathématique sur toute image sur laquelle

peut être imposée une structure de treillis. Aujourd'hui, la morphologie mathématique offre un large choix d'opérateurs et d'algorithmes performants, basés par exemple sur des représentations des images sous forme de files d'attente, de graphes, d'arbres, etc. L'élément structurant, élément fondamental des premiers opérateurs morphologiques, n'est aujourd'hui plus une contrainte, et les opérateurs connexes s'en affranchissent. Le lecteur intéressé trouvera un historique des premiers travaux en morphologie mathématique dans un article [MS02] que les deux fondateurs de cette théorie avaient écrit quelques années avant sa publication¹.

La morphologie mathématique n'est évidemment pas la seule boîte à outils disponible en traitement d'image [GW08]. Elle se distingue des autres formalismes par certains avantages :

- les caractéristiques des approches morphologiques (analyse spatiale, traitement non-linéaire) leur permettent d'être un complément pertinent à de nombreuses autres techniques (statistiques par exemple) ;
- les opérateurs morphologiques sont simples à appréhender, en particulier dans le cas des images binaires ;
- le formalisme est riche et complet, s'appuyant sur de solides bases mathématiques ;
- il existe de nombreux algorithmes ou implantations efficaces des outils morphologiques ;

mais aussi quelques inconvénients :

- les avancées en morphologie mathématique restent très peu connues en dehors de cette communauté, la connaissance des opérateurs morphologiques en traitement d'image étant le plus souvent limitée aux opérateurs les plus basiques et au cas des images binaires ;
- les applications de la morphologie mathématique concernent essentiellement quelques domaines privilégiés (imagerie médicale, analyse de matériaux) alors que de nombreux autres domaines pourraient bénéficier de ces outils ;
- le fonctionnement des outils peut être moins intuitif dans le cas d'images à niveaux de gris, et le traitement morphologique d'images complexes (couleur ou multispectrales, multitemporelles, etc) est loin d'être répandu ;
- l'élaboration de solutions morphologiques avancées nécessite le plus souvent d'être un expert ou un utilisateur confirmé des techniques de morphologie mathématique ;

Cette liste n'est bien sûr pas exhaustive mais elle permet de mettre en lumière les raisons qui nous ont poussé à nous intéresser aux sujets traités dans ce document. En effet, il nous semble que la morphologie mathématique dispose d'un potentiel d'outils encore insuffisamment exploités en traitement d'image, notamment du fait d'une mauvaise prise en compte des images multivaluées et d'une difficulté à choisir ou guider les approches morphologiques. En complément, élargir le champ d'application de la morphologie mathématique en illustrant l'intérêt des approches morphologiques dans de nouveaux domaines est important pour mieux diffuser les travaux menés dans cette communauté.

2.1.2 Cadre théorique et notations

Formulation par la théorie des ensembles

La morphologie mathématique fut initialement appliquée aux images binaires. Ces images étant composées d'objets (ou formes) et d'un fond, elles peuvent être représentées par des ensembles de pixels (correspondant aux objets). La théorie des ensembles est alors une solution assez intuitive pour définir les opérateurs morphologiques. Ainsi, l'érosion et la dilatation peuvent être formulées en termes d'inclusion, d'intersection et/ou d'union, comme illustré en figure 2.1.

Dans le cas le plus fréquent d'éléments structurants représentés sous forme ensembliste, l'extension dite plate des opérateurs morphologiques aux images à niveaux de gris s'effectue en représentant de telles images comme des piles d'images binaires obtenues par seuillage avec des valeurs de seuil croissantes (le terme d'ombre [Ste86] est également utilisé). L'érosion et la dilatation sont alors appliquées sur chaque image binaire, et l'union (ou somme) des images obtenues permet de produire l'image résultat à niveaux de gris (figure 2.2).

Cependant, cette approche présente différents inconvénients. Ainsi, elle nécessite en théorie la construction de 256 images binaires pour représenter une image à niveaux de gris codée sur 8 bits, et l'application de 256 opérations morphologiques binaires pour obtenir le résultat équivalent sur l'image initiale à niveaux de gris.

¹Il s'agit du dernier article de Matheron, disparu en 2000.

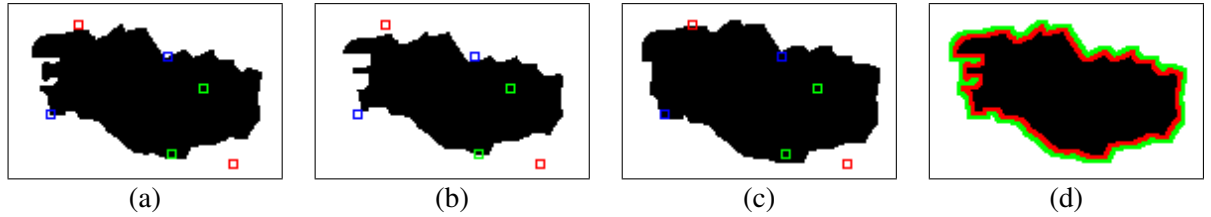


FIG. 2.1 – Morphologie mathématique ensembliste et images binaires : (a) une image binaire où l'objet est représenté en noir, avec différentes configurations locales (inclusion en vert, intersection en bleu, et disjonction en rouge), (b) érosion de l'image ne conservant que les configurations incluses dans l'objet (en vert), (c) dilatation de l'image conservant aussi les configurations intersectant l'objet (en bleu), et (d) comparaison entre érosion (suppression des pixels rouges) et dilatation (ajout des pixels verts).

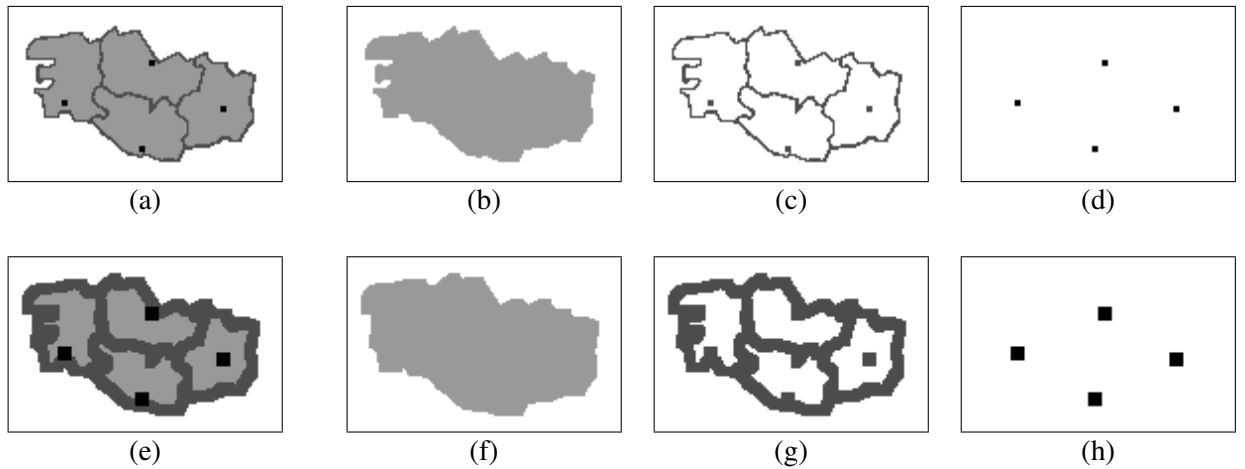


FIG. 2.2 – Morphologie mathématique ensembliste et images à niveaux de gris : image à niveaux de gris (a) et sa représentation par pile d'images binaires (b)-(d), et résultat d'une érosion (e) obtenue par empilement des images binaires érodées (f)-(h).

Fort heureusement, l'utilisation d'opérations de minimum et de maximum évite en pratique ces nombreux calculs. De plus, elle ne permet pas l'utilisation d'éléments structurants à niveaux de gris. Outre ces limites évidentes, elle est également caractérisée par différents problèmes théoriques (cf. en particulier [Ron90] pour une discussion sur le sujet). Il faut néanmoins noter que, malgré ses limites, la théorie des ensembles reste un moyen simple et intuitif pour formuler les opérations morphologiques de base (en particulier dans le cas d'images binaires) comme en témoignent les nombreux chapitres relatifs à la morphologie mathématique dans les ouvrages généraux de traitement des images (par exemple le très célèbre *Digital Image Processing* de Gonzalez et Woods [GW08]).

Formulation par la théorie des treillis

Pour s'affranchir des problèmes posés par une représentation ensembliste de la morphologie mathématique, la théorie des treillis complets, suggérée par Matheron et Serra [Ser88] est aujourd'hui le cadre théorique le plus communément accepté pour définir les opérateurs morphologiques. Nous ferons également ce choix dans la suite de ce document.

Un treillis complet est défini comme un ensemble partiellement ordonné $(\mathcal{T}, \leq)$ où tout sous-ensemble $\mathcal{S}$ de $\mathcal{T}$ possède un plus grand minorant ou infimum $\bigwedge \mathcal{S}$ et un plus petit majorant ou supremum $\bigvee \mathcal{S}$.

La relation $\leq$ est :

$$\text{(réflexive)} \quad \forall a \in \mathcal{T}, a \leq a \quad (2.1.1)$$

$$\text{(anti-symétrique)} \quad \forall a, b \in \mathcal{T}, a \leq b, b \leq a \implies a = b \quad (2.1.2)$$

$$\text{(transitive)} \quad \forall a, b, c \in \mathcal{T}, a \leq b, b \leq c \implies a \leq c \quad (2.1.3)$$

Dans le cas d'images binaires, on utilise généralement le treillis des ensembles (la relation d'ordre est $\subseteq$, l'infimum $\cap$ et le supremum $\cup$). Dans le cas des images à niveaux de gris, on considère le treillis des fonctions numériques où la relation d'ordre est $\leq$, tandis que l'infimum $\wedge$ et le supremum $\vee$ sont calculés simplement à l'aide des opérateurs de minimum et maximum si les niveaux de gris sont discrets et bornés (ce qui est le plus souvent le cas). Pour plus de simplicité, nous utiliserons le treillis des fonctions numériques pour toutes les images, qu'elles soient binaires ou à niveaux de gris. Cependant, dans le cas d'images multivaluées où à chaque pixel est associé un vecteur de niveaux de gris, la définition de la relation d'ordre entre vecteurs n'est pas triviale et fera l'objet de la section 2.2.

Notations

En s'appuyant sur le cadre théorique des treillis, nous pouvons définir l'ensemble des opérations morphologiques, dont nous rappellerons les plus connues dans cette section. Pour cela, et dans la suite de ce manuscrit, nous utiliserons les notations suivantes. Une image est définie comme une fonction $f : \mathcal{D} \rightarrow \mathcal{T} : \mathbf{p} \mapsto f(\mathbf{p})$, et on utilisera également la notation $\mathcal{F} = \mathcal{D}^{\mathcal{T}}$ pour représenter les fonctions de $\mathcal{D}$ dans $\mathcal{T}$. En cas d'ambiguïté, on pourra indiquer $\mathcal{T}$ et $\mathcal{D}$ par f , soit $f : \mathcal{D}_f \rightarrow \mathcal{T}_f$. Le domaine de définition $\mathcal{D}$ de la fonction f correspond à l'espace des coordonnées d'un pixel $\mathbf{p}$, avec $\mathcal{D} \subset \mathbb{N}^m$ pour une image à m dimensions. Chaque pixel $\mathbf{p}$ est donc défini comme un m -uplet de coordonnées, par exemple $\mathbf{p} = (x, y)$ pour une image 2-D, $\mathbf{p} = (x, y, z)$ pour un volume 3-D ou $\mathbf{p} = (x, y, t)$ pour une séquence d'images. Dans le cas des images 2-D, on pourra utiliser les notions de 4- et de 8-voisinage, respectivement définis comme l'ensemble des 4 pixels $\mathbf{p}' = (x', y')$ autour de $\mathbf{p}$ tel que $|x - x'| + |y - y'| = 1$ et l'ensemble des 8 pixels tel que $\max(|x - x'|, |y - y'|) = 1$. Nous noterons $\mathcal{N}_{\mathbf{p}}$ le voisinage de $\mathbf{p}$.

L'ensemble $\mathcal{T}$ correspond aux valeurs pouvant être affectées aux pixels, avec $\perp$ et $\top$ les plus petit et plus grand éléments de $\mathcal{T}$. Dans le cas de niveaux de gris discrets, on a $\mathcal{T} \subset \mathbb{Z}$ ou le plus souvent $\mathcal{T} \subset \mathbb{N}$ (par exemple $\mathcal{T} = \{0, \dots, 255\}$ dans le cas d'images à 256 niveaux de gris où la valeur de chaque pixel est codée sur un octet), tandis qu'on a $\mathcal{T} \subset \mathbb{R}$ dans le cas de niveaux de gris continus. Les images binaires peuvent être représentées sous la forme d'une image à deux niveaux de gris (0 pour le fond et 1 pour les objets), d'où $\mathcal{T} = \{0, 1\}$. Enfin, pour les images multivaluées telles que les images couleur ou multispectrales, on a $\mathcal{T} \subset \mathbb{N}^n$, (voire $\mathcal{T} \subset \mathbb{Z}^n$ ou $\mathcal{T} \subset \mathbb{R}^n$ si les pixels peuvent respectivement prendre des valeurs négatives ou réelles), avec n le nombre de bandes ou canaux. On notera $\mathbf{f}$ une telle image, où à chaque pixel est associé un vecteur de valeurs, i.e. $\mathbf{f}(\mathbf{p}) = \{f_1(\mathbf{p}), \dots, f_n(\mathbf{p})\}$.

L'opérateur de complémentation est formulé par $\mathcal{C} : \mathcal{F} \rightarrow \mathcal{F} : f \mapsto \mathcal{C}(f) = f^c$ avec $f^c(\mathbf{p}) = \top - f(\mathbf{p})$. La réflexion $\check{f}$ de f s'écrit $\check{f}(\mathbf{p}) = f(-\mathbf{p})$. Le dual de f noté f^* est défini par $f^*(\mathbf{p}) = \mathcal{C}(\check{f})(\mathbf{p}) = \top - f(-\mathbf{p})$. Notons que l'utilisation de $\top$ dans les définitions précédentes n'est justifiée que par l'emploi d'un intervalle de valeurs fini pour formuler $\mathcal{T}$: elle disparaît si on considère par exemple $\mathcal{T} = \overline{\mathbb{Z}} = \mathbb{Z} \cup \{+\infty, -\infty\}$ ou $\mathcal{T} = \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$. Enfin, il faut remarquer que la structure de treillis peut également être imposée sur les fonctions, et on aura par exemple $\forall \mathbf{p} \in \mathcal{D}, f \leq g \iff f(\mathbf{p}) \leq g(\mathbf{p})$.

2.1.3 Opérateurs et filtres morphologiques

Un opérateur morphologique est une transformation $\psi : \mathcal{F} \rightarrow \mathcal{F} : f \mapsto \psi(f)$ qui produit une image à partir d'une autre image. Les domaines de définition $\mathcal{D}$ de l'image initiale f et de l'image transformée $\psi(f)$ sont identiques, comme le sont généralement les plages de niveaux de gris $\mathcal{T}$ de f et $\psi(f)$. Ainsi, il devient possible d'appliquer récursivement un opérateur, et nous noterons cette application $\psi^{(n)} = \psi \psi^{(n-1)}$ pour n applications successives de ψ , avec $\psi^{(0)}$ l'identité. Il est aussi possible d'utiliser les opérateurs morphologiques pour en construire d'autres, par composition $\psi(f) = \psi_1(\psi_2(f))$ ou par différence $\psi(f) = \psi_1(f) - \psi_2(f)$.

Élément structurant

L'application d'un opérateur morphologique nécessite le plus souvent la spécification d'un voisinage pré-défini, ou l'utilisation d'un paramètre de forme, connu sous le nom d'élément structurant. Ce paramètre a une incidence majeure sur le résultat fourni par l'opérateur. Il est doté d'une origine $\mathbf{o}$ et peut être défini soit comme une fonction $b : \mathcal{D} \rightarrow \mathcal{T}$, soit comme un ensemble B . L'origine est le plus souvent le pixel de coordonnées nulles $\mathbf{o} = \mathbf{0}$, ce qui implique dans ce cas la définition d'un espace de coordonnées à valeurs positives ou négatives, i.e. $\mathcal{D} \subset \mathbb{Z}^m$. Par la suite, nous ferons l'hypothèse implicite que $\mathcal{D} \subset \mathbb{Z}^m$ pour les éléments structurants et $\mathcal{D} \subset \mathbb{N}^m$ pour les images. Le choix de la forme (carré, croix, segment, disque, etc.) et de la taille λ de l'élément structurant permettent d'affiner le comportement de l'opérateur morphologique en fonction du résultat recherché. Dans un souci de lisibilité, on pourra noter $\psi_\lambda = \psi_{B_\lambda}$ l'application d'un opérateur morphologique ψ avec un élément structurant de taille λ , voire seulement ψ lorsqu'il s'agit d'un élément structurant élémentaire isotropique (une croix ou un carré de 3×3 pixels, respectivement pour le 4- ou le 8-voisinage).

La représentation des éléments structurants sous forme de fonction est une généralisation du cas ensembliste où il faut définir la forme (ou domaine de définition $\mathcal{D}$ de la fonction) comme dans le cas ensembliste, mais également les valeurs des pixels $b(\mathbf{q})$ composant l'élément structurant. Dans ce cas, on utilise les termes d'élément structurant fonctionnel, d'élément structurant à niveaux de gris ou encore de fonction structurante. Cependant, le plus souvent, on définit les éléments structurants comme des ensembles, dans le cas des images binaires mais aussi des images à niveaux de gris (et on parle alors de morphologie mathématique plate, d'opérateurs et d'éléments structurants plats, ces derniers pouvant néanmoins être représentés par une fonction $b : \mathcal{D} \rightarrow \mathcal{T} : \mathbf{q} \mapsto b(\mathbf{q}) = 0$ où $\mathcal{D} = B$).

Nous discuterons dans la section 2.3.2 les différentes informations ou connaissances pouvant être apportées par l'utilisateur au travers des éléments structurants. Notons que les opérateurs morphologiques (ou structuraux) peuvent être généralisés et que l'usage des éléments structurants n'est plus obligatoire en morphologie mathématique.

Propriétés des opérateurs

Tous les opérateurs morphologiques sont caractérisés par un ensemble de propriétés, et c'est une des raisons qui ont fait le succès de la morphologie mathématique. Les principales propriétés sont :

$$\text{croissance} \quad f_1 \leq f_2 \implies \psi(f_1) \leq \psi(f_2) \quad (2.1.4)$$

$$\text{idempotence} \quad \psi(\psi(f)) = \psi(f) \quad (2.1.5)$$

$$\text{extensivité} \quad f \leq \psi(f) \quad (2.1.6)$$

$$\text{anti-extensivité} \quad \psi(f) \leq f \quad (2.1.7)$$

$$\text{complémentation (entre } \psi \text{ et } \psi^c) \quad \psi^c(f) = \mathbb{C}\psi(f) \quad (2.1.8)$$

$$\text{auto-complémentation} \quad \psi^c(f) = \psi(f) \quad (2.1.9)$$

$$\text{dualité (entre } \psi \text{ et } \psi^*) \quad \psi^*(f) = \mathbb{C}\psi\mathbb{C}(f) \quad (2.1.10)$$

$$\text{auto-dualité} \quad \psi^*(f) = \psi(f) \quad (2.1.11)$$

$$\text{invariance aux translations } g_{\mathbf{t}} \quad \psi(g_{\mathbf{t}}f) = g_{\mathbf{t}}\psi(f) \quad (2.1.12)$$

L'invariance est une propriété qui a une signification particulière dans le contexte de la morphologie mathématique. Ainsi, les opérateurs morphologiques sont invariants à certains groupes de transformations géométriques, tel que celui des translations noté $g_{\mathbf{t}}$ où $(g_{\mathbf{t}}f)(\mathbf{p}) = f(\mathbf{p} - \mathbf{t})$. Notons qu'il s'agit en fait d'une équivariance et non d'une invariance [BS01]. De plus, on dit qu'une image f est invariante à un opérateur morphologique ψ si $\psi(f) = f$, i.e. si l'opérateur ψ n'a pas d'effet sur f . Toutes les images f telles que $\psi(f) = f$ forment la classe invariante de ψ .

Opérateurs élémentaires

Les opérateurs élémentaires en morphologie mathématique sont l'érosion et la dilatation, dont nous rappelons ici la définition dans le cas d'éléments structurants plats ou à niveaux de gris. La formulation usuelle d'une

érosion plate ε d'une image f avec un élément structurant B est la suivante :

$$\varepsilon_B(f)(\mathbf{p}) = \bigwedge_{\mathbf{q} \in B} f(\mathbf{p} + \mathbf{q}), \quad \mathbf{p} \in \mathcal{D} \quad (2.1.13)$$

Il s'agit d'une opération croissante et distributive sur l'infimum, voire anti-extensive si l'élément structurant contient son origine (ce qui est très généralement le cas). L'érosion produit une image où les zones claires ont été réduites au profit des zones sombres, en propageant les minima locaux de f calculés dans le voisinage défini par B . Notons que si l'élément structurant est défini comme une fonction b , on utilise la formulation suivante :

$$\varepsilon_b(f)(\mathbf{p}) = \bigwedge_{\mathbf{q} \in \mathcal{D}_b} f(\mathbf{p} + \mathbf{q}) - b(\mathbf{q}), \quad \mathbf{p} \in \mathcal{D}_f \quad (2.1.14)$$

L'opération duale (au sens de la complémentation) de l'érosion est la dilatation, définie dans le cas d'un élément structurant plat B comme :

$$\delta_B(f)(\mathbf{p}) = \bigvee_{\mathbf{q} \in \check{B}} f(\mathbf{p} + \mathbf{q}), \quad \mathbf{p} \in \mathcal{D} = \bigvee_{\mathbf{q} \in B} f(\mathbf{p} - \mathbf{q}), \quad \mathbf{p} \in \mathcal{D} \quad (2.1.15)$$

avec $\check{B} = \{-q \mid q \in B\}$ la réflexion de B . Dans le cas d'une fonction structurante b , on a la définition suivante :

$$\delta_b(f)(\mathbf{p}) = \bigvee_{\mathbf{q} \in \mathcal{D}_b} f(\mathbf{p} - \mathbf{q}) + b(\mathbf{q}), \quad \mathbf{p} \in \mathcal{D}_f \quad (2.1.16)$$

La dilatation est croissante et distributive sur le supremum, voire extensive si l'élément structurant contient son origine. Son effet est l'opposé de l'érosion, puisqu'elle va réduire les zones sombres au profit des zones claires en propageant les maxima locaux. L'élément structurant étant le plus souvent isotropique ($B = \check{B}$), on assimile fréquemment l'application d'une érosion et d'une dilatation au calcul respectif d'un minimum et d'un maximum local dans une fenêtre définie par la forme de l'élément structurant B .

Filtres morphologiques

Les filtres morphologiques sont des opérateurs garantissant les propriétés de croissance et d'idempotence. Ils peuvent être construits par composition des opérateurs morphologiques élémentaires que sont l'érosion et la dilatation. Ainsi, l'ouverture γ consiste à appliquer une érosion puis une dilatation sur l'image f , tandis que l'opération duale est la fermeture φ qui consiste en une dilatation suivie d'une érosion :

$$\gamma_B(f) = \delta_B(\varepsilon_B(f)) \quad (2.1.17)$$

$$\varphi_B(f) = \varepsilon_B(\delta_B(f)) \quad (2.1.18)$$

Ces deux opérateurs filtres morphologiques sont de plus duaux par rapport à la complémentation, avec la relation $\varphi_{\check{B}} = \gamma_B^*$. En supprimant les maxima locaux, l'ouverture a pour effet de supprimer les détails clairs dans l'image, tandis que la fermeture supprime les détails sombres correspondant aux minima locaux. On a également $\gamma(f) \leq f \leq \varphi(f)$, i.e. l'ouverture est anti-extensive tandis que la fermeture est extensive. Pour éviter ce caractère asymétrique, il est possible d'employer des filtres alternés séquentiels (ou ASF, *alternate sequential filters*). Ceux-ci peuvent être définis par exemple sous la forme d'une application successive d'ouvertures et fermetures avec des éléments structurants de taille croissante. Plus généralement, il est possible de définir un filtre alterné séquentiel avec toute paire de filtres (ζ, ξ) vérifiant les trois propriétés suivantes : ζ est anti-extensif, ξ est extensif, et on a la relation $\zeta_1 < \xi_1$.

Généralisation, filtres algébriques et opérateurs connexes

L'inconvénient principal de l'ouverture, de la fermeture, et des autres filtres morphologiques ou structuraux décrits précédemment est leur forte sensibilité à la forme de l'élément structurant, qui se traduit généralement

par la création de contours artificiels au sein de l'image filtrée. Afin d'éviter ce problème, il est possible de faire appel à différentes techniques telles que les filtres algébriques (une généralisation des opérateurs structuraux), les filtres par reconstruction, ou encore les opérateurs connexes.

Ainsi, tout filtre morphologique (donc opérateur croissant et idempotent) également anti-extensif (respectivement extensif) est une ouverture algébrique (respectivement fermeture algébrique). Une érosion algébrique est quant à elle une transformation ψ qui est distributive avec les infima, i.e. $\psi(\bigwedge \{f\}) = \bigwedge \{\psi(f)\}$ avec $\{f\}$ une famille d'images. De façon similaire, une dilatation algébrique est distributive avec les suprema, i.e. $\psi(\bigvee \{f\}) = \bigvee \{\psi(f)\}$. Une adjonction est une paire (ε, δ) d'érosion et dilatation telle que $f_1 \leq \varepsilon(f_2) \iff \delta(f_1) \leq f_2$ avec $f_1, f_2 \in \mathcal{F}$.

Les filtres par reconstruction sont particulièrement utiles pour préserver les contours initiaux des objets présents dans l'image. Pour les définir, on utilise des opérateurs géodésiques, tels que l'érosion (respectivement la dilatation) géodésique d'un marqueur f sur (respectivement sous) un masque b :

$$\varepsilon_b^{(1)}(f)(\mathbf{p}) = \varepsilon^{(1)}(f)(\mathbf{p}) \vee b(\mathbf{p}) \text{ avec } f \geq b \quad (2.1.19)$$

$$\delta_b^{(1)}(f)(\mathbf{p}) = \delta^{(1)}(f)(\mathbf{p}) \wedge b(\mathbf{p}) \text{ avec } f \leq b \quad (2.1.20)$$

et on définit alors la reconstruction géodésique comme une application successive de ces opérateurs jusqu'à convergence :

$$\rho_b^\varepsilon(f) = \varepsilon_b^{(j)}(f) \text{ avec } j \text{ tel que } \varepsilon_b^{(j)}(f) = \varepsilon_b^{(j-1)}(f) \text{ et } \varepsilon_b^{(n)}(f) = \varepsilon_b^{(1)}(\varepsilon_b^{(n-1)}(f)) \quad (2.1.21)$$

$$\rho_b^\delta(f) = \delta_b^{(j)}(f) \text{ avec } j \text{ tel que } \delta_b^{(j)}(f) = \delta_b^{(j-1)}(f) \text{ et } \delta_b^{(n)}(f) = \delta_b^{(1)}(\delta_b^{(n-1)}(f)) \quad (2.1.22)$$

De ces opérateurs peuvent être formulés les filtres par reconstruction, comme par exemple l'ouverture par reconstruction $\gamma_B^\rho(f) = \rho_f^\delta(\varepsilon_B(f))$ et la fermeture par reconstruction $\varphi_B^\rho(f) = \rho_f^\varepsilon(\delta_B(f))$ où l'image initiale f sert également de masque pour la reconstruction géodésique (et supposant ici B connexe et contenant l'origine).

Une autre possibilité pour définir des filtres algébriques consiste à considérer une famille d'ouvertures (respectivement fermetures) et à calculer leur supremum (respectivement infimum). Chaque ouverture ou fermeture est ici appliquée avec un élément structurant différent. En notant $\mathcal{B} = \{B\}$ la famille des éléments structurants à considérer, on définit l'ouverture et la fermeture algébrique par :

$$\gamma_{\mathcal{B}}^\alpha(f) = \bigvee_{B \in \mathcal{B}} \gamma_B(f) \quad (2.1.23)$$

$$\varphi_{\mathcal{B}}^\alpha(f) = \bigwedge_{B \in \mathcal{B}} \varphi_B(f) \quad (2.1.24)$$

Parmi les principaux filtres algébriques construits à l'aide d'une telle approche, les opérateurs d'aire s'affranchissent complètement de la forme de l'élément structurant. Pour cela, ils considèrent l'ensemble de tous les éléments structurants de cardinalité λ (i.e. ensembles connexes composés de λ pixels). Il s'agit d'un cas particulier des filtres d'attributs ψ^χ où l'attribut ou critère χ à satisfaire est relatif à l'aire (i.e. la fonction booléenne $\chi(B, \lambda) = \{\text{card}(B) = \lambda\}$) et dont la formulation générale est :

$$\gamma_\lambda^\chi(f) = \bigvee_B \{\gamma_B(f) \mid B \text{ est connexe et } \chi(B, \lambda)\} \quad (2.1.25)$$

$$\varphi_\lambda^\chi(f) = \bigwedge_B \{\varphi_B(f) \mid B \text{ est connexe et } \chi(B, \lambda)\} \quad (2.1.26)$$

Une autre approche consiste à définir des opérateurs connexes, en s'appuyant par exemple sur une représentation de l'image sous la forme d'un arbre des coupes [Sal08] (tel que l'arbre des composantes, le *max-tree* [SOG98], l'arbre binaire de partition, etc.). Nous illustrons ici le cas du *max-tree* [SOG98], qui peut être intuitivement construit en appliquant la fonction seuil T_h sur une image f à chacun des niveaux de gris $h \in \mathcal{T}$. L'analyse de chaque image binaire $T_h(f)$ permet de constituer un ensemble de composantes connexes $X_h = \{X_h^i\}$, chacune d'elles représentant un nœud dans l'arbre. La structure arborescente est alors obtenue en

créant des arêtes (ou relations de parenté) entre deux nœuds de niveaux de gris successifs vérifiant une condition d'inclusion, soit $X_h^i \subseteq X_{h+1}^j$, et en définissant les feuilles comme les minima régionaux $\text{MIN}(f)$. Notons que cette procédure n'est pas l'implantation la plus efficace pour obtenir le *max-tree*. Ces représentations, dont un exemple est donné en figure 2.3, permettent de construire nombre d'opérateurs connexes en bénéficiant d'implantations particulièrement efficaces, et elles offrent une alternative élégante aux opérateurs structuraux. Par exemple, en supprimant les branches de l'arbre associées à des composantes de taille inférieure à une valeur prédéfinie, on simplifie l'image sans modifier les contours.



FIG. 2.3 – Représentation d'un arbre des coupes Max-Tree : (de gauche à droite) image à niveaux de gris, composantes connexes aux trois différents niveaux de gris, et arbre correspondant..

2.1.4 Quelques exemples de traitements morphologiques

Mise en correspondance

La mise en correspondance ou reconnaissance de formes (*template matching*) a pour but de vérifier l'occurrence d'un motif prédéfini dans une image. La morphologie mathématique fournit un opérateur non-croissant pour atteindre cet objectif, la transformée en tout ou-rien (ou *hit-or-miss transform*).

Dans le cas des images binaires, la définition de la transformée en tout-ou-rien est relativement aisée. En considérant une paire d'éléments structurants disjoints (B_1, B_2) , cet opérateur peut être formulé comme l'intersection entre l'érosion de l'image par B_1 et de son complémentaire par B_2 . Puisque les érosions et dilatations sont des opérateurs duaux par complémentement, on a :

$$\begin{aligned} \eta_{B_1, B_2}(f) &= \varepsilon_{B_1}(f) \wedge \varepsilon_{B_2}(f^c) \\ &= \varepsilon_{B_1}(f) - \delta_{\tilde{B}_2}(f) \end{aligned} \quad (2.1.27)$$

avec l'opérateur $-$ représentant ici la soustraction booléenne (soit $0 - 1 = 0$).

L'extension aux images à niveaux de gris est plus complexe, et elle nécessite généralement deux étapes [NPR07a]. La première étape est appelée ajustage (ou *fitting*) et consiste à sélectionner les pixels pertinents (en fonction du couple (B_1, B_2) d'éléments structurants utilisés), tandis que la seconde étape intitulée *valuation* a pour but d'assigner une valeur aux pixels. Différentes formulations ont été proposées dans la littérature, citons par exemple celles de Ronse [Ron96] et de Soille [Soi03], respectivement définies par :

$$\eta_{b_1, b_2}(f)(\mathbf{p}) = \begin{cases} \varepsilon_{b_1}(f)(\mathbf{p}) & \text{si } \varepsilon_{b_1}(f)(\mathbf{p}) \geq \delta_{b_2^*}(f)(\mathbf{p}) \\ -\infty & \text{sinon} \end{cases} \quad (2.1.28)$$

$$\eta_{B_1, B_2}(f)(\mathbf{p}) = \begin{cases} \varepsilon_{B_1}(f)(\mathbf{p}) - \delta_{\tilde{B}_2}(f)(\mathbf{p}) & \text{si } \varepsilon_{B_1}(f)(\mathbf{p}) \geq \delta_{\tilde{B}_2}(f)(\mathbf{p}) \\ 0 & \text{sinon} \end{cases} \quad (2.1.29)$$

Rehaussement de contraste

L'utilisation d'opérateurs morphologiques définis par différence, aussi appelés opérateurs résidus, permet de rehausser le contraste présent dans l'image. Ces opérateurs, même s'ils sont caractérisés par des propriétés bien plus limitées que les filtres morphologiques obtenus par composition (par exemple ils ne sont pas croissants), sont particulièrement utiles pour renforcer certains détails présents dans l'image.

Ainsi, la construction d'un opérateur résidu à partir d'une érosion ou d'une dilatation (voire des deux) fournit un opérateur de gradient capable de mettre en évidence les zones d'hétérogénéité, souvent représentatives des contours des objets présents dans les images. Par exemple, le gradient morphologique (ou gradient de

Beucher) est un opérateur auto-complémentaire défini dans le cas d'un élément structurant isotropique par :

$$\nabla_B(f) = \delta_B(f) - \varepsilon_B(f) \quad (2.1.30)$$

mais il ne s'agit que d'un exemple des opérateurs de gradient proposés dans le cadre morphologique [RSB93].

Il est également possible d'utiliser des filtres morphologiques dans la définition de l'opérateur résidu. Ainsi, on définit les chapeaux haut-de-forme (ou *top-hat*) en considérant les opérateurs d'ouverture et de fermeture. L'utilisation conjointe de ces deux opérateurs est à l'origine du chapeau haut-de-forme auto-dual :

$$\kappa_B(f) = \varphi_B(f) - \gamma_B(f) \quad (2.1.31)$$

tandis que le chapeau haut-de-forme blanc (ou par ouverture, ou détecteur de pic) n'utilise que l'ouverture ($\kappa_B^-(f) = f - \gamma_B(f)$), et son complémentaire (chapeau haut-de-forme noir, par fermeture, ou détecteur de vallée) n'utilise que la fermeture ($\kappa_B^+(f) = \varphi_B(f) - f$).

Simplification

La simplification d'une image a pour but de supprimer certains détails de l'image (sans en introduire de nouveaux) en produisant généralement des zones uniformes dans l'image. Elle peut être obtenue à l'aide des nivellements, introduits par Meyer [Mey98], qui sont des opérateurs morphologiques appartenant à la famille des opérateurs connexes (à ce titre ils offrent la garantie de ne pas introduire de nouveaux contours ou détails dans l'image). Ils nécessitent l'utilisation d'une image de référence f et d'une image marqueur g . L'obtention du nivellement de f par g peut s'obtenir en itérant

$$g^{(n)} = \left[f \wedge \delta \left(g^{(n-1)} \right) \right] \vee \varepsilon \left(g^{(n-1)} \right) \quad (2.1.32)$$

jusqu'à convergence, avec la convention $g^{(0)} = g$. Ce schéma de calcul itératif peut être avantageusement remplacé par des algorithmes plus efficaces [Mey04].

Segmentation

La segmentation est une transformation qui consiste en une partition de l'image initiale pour former un ensemble de régions. Elle permet d'identifier les différents objets d'intérêt dans l'image en vue de leur analyse future. L'approche dite de Ligne de Partage des Eaux (LPE) est certainement la méthode de segmentation morphologique la plus connue, et une des méthodes de segmentation parmi les plus utilisées en traitement d'image. En représentant l'image à segmenter f comme un relief, elle identifie dans ce dernier les lignes de partage des eaux et les associe aux frontières des régions ainsi segmentées. Différents paradigmes ont été introduits dans la littérature [RM00], et nous décrivons ici l'approche que nous avons principalement utilisée dans nos travaux (illustrée dans la figure 2.4). Celle-ci consiste à simuler l'inondation du relief et à séparer les différents bassins par des digues (qui déterminent les lignes de partage des eaux ou les frontières des régions). L'implantation algorithmique [Mey91] exploite des structures de files hiérarchiques, ce qui permet une implantation relativement efficace (notons cependant que de nombreuses optimisations ou implantations plus intéressantes ont été proposées depuis). Nous reprenons ci-dessous la formulation mathématique introduite dans [VS91], en supposant que les niveaux de gris de l'image initiale f sont définis dans l'intervalle $[h_{\min}, h_{\max}]$. Les bassins de rétention (c'est-à-dire les régions produites par le processus d'inondation) sont notés $X_{h_{\max}}$ et obtenus par le schéma de récursion suivant :

$$\begin{aligned} X_{h_{\min}} &= T_{h_{\min}}(f) \\ X_{h+1} &= \text{MIN}_{h+1}(f) \cup IZ_{T_{h+1}(f)}(X_h), \quad h_{\min} \leq h < h_{\max} \end{aligned} \quad (2.1.33)$$

où T_h est la fonction seuil de niveau h (c'est-à-dire $T_h(f) = \{\mathbf{p} \mid f(\mathbf{p}) \geq h\}$), MIN_h est l'union de tous les minima régionaux d'altitude h (soit $\text{MIN}_h(f) = \{\mathbf{p} \mid f(\mathbf{p}) = h, \mathbf{p} \in \text{MIN}(f)\}$ avec $\text{MIN}(f) = \rho^\varepsilon(f + 1) - f$ les minima régionaux de f obtenus par reconstruction géodésique [Soi03]), et enfin $IZ_A(B)$ est l'union des

zones d'influence géodésique des composantes connexes de B avec $B \subseteq A$. Nous faisons aussi appel aux définitions suivantes :

$$IZ_A(B) = \bigcup_{j=1}^l iz_A(B_j) \quad (2.1.34)$$

$$iz_A(B_j) = \{p \in A \mid \forall k \in [1, l] \setminus \{j\} : d_A(p, B_j) < d_A(p, B_k)\} \quad (2.1.35)$$

$$d_A(a, B) = \min_{b \in B} d_A(a, b) \quad (2.1.36)$$

où $d_A(a, b)$ représente la distance géodésique entre a et b dans A .

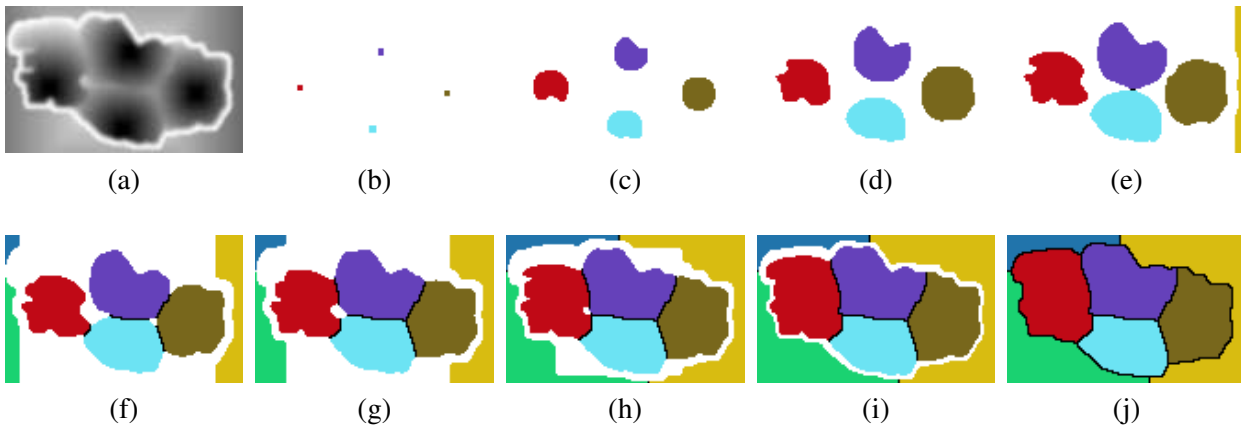


FIG. 2.4 – Les différentes étapes du processus de segmentation par inondation pour obtenir les lignes de partage des eaux : (a) image initiale et (b)-(j) inondation progressive de la surface à partir des minima locaux et construction des digues (en noir). Le résultat de segmentation est l'image finale donnée en (j).

Cette approche n'est généralement pas appliquée directement sur l'image à segmenter, puisqu'elle cherche les frontières d'intensité élevée (lignes de partage des eaux) séparant les zones d'intensité plus faible (bassins de rétention). Ainsi, elle est fréquemment précédée du calcul d'un gradient qui permet de rehausser les valeurs des pixels de frontières tout en limitant les valeurs des pixels dans les zones homogènes. Cependant, comme nous pourrions le constater par la suite, le calcul d'un gradient sur une image produit fréquemment de nombreux minima locaux, qui résultent en autant de bassins de rétention, soit autant de régions produites dans l'algorithme de segmentation. Ce phénomène, appelé *sursegmentation*, peut être évité en employant différentes techniques de réduction de sursegmentation. Une autre approche consiste à fixer a priori le nombre de régions à l'aide de marqueurs qui définissent les bassins de rétention initiaux lors du processus d'inondation du relief. En notant M l'ensemble des marqueurs fournis au titre de l'initialisation spatiale, on définit alors le nouveau relief g comme :

$$g(\mathbf{p}) = \begin{cases} h_{\min-1} & \text{si } \mathbf{p} \in M \\ f(\mathbf{p}) & \text{sinon} \end{cases} \quad (2.1.37)$$

où $h_{\min-1}$ est une nouvelle valeur spécifique aux marqueurs initiaux. La segmentation de l'image correspond à l'ensemble $X_{h_{\max}}$ des bassins de rétention de la fonction g , calculé d'après un schéma de récursion ne nécessitant plus l'utilisation des minima régionaux :

$$\begin{aligned} X_{h_{\min-1}} &= T_{h_{\min-1}}(g) \\ X_{h+1} &= IZ_{T_{h+1}(g)}(X_h), \quad h_{\min-1} \leq h < h_{\max} \end{aligned} \quad (2.1.38)$$

mais notons que la segmentation peut également être obtenue directement par l'équation (2.1.33), en utilisant une technique d'imposition des minima sur la fonction f (cf. [Soi03]) : les minima de f non marqués sont bouchés par une reconstruction par érosion, ou autrement dit, f est remplacée par $\rho_g^\varepsilon(m)$, avec $m(\mathbf{p})$ valant $h_{\min-1}$ si $\mathbf{p} \in M$, et $h_{\max}$ sinon.

2.1.5 Descripteurs morphologiques

Les opérateurs morphologiques peuvent également permettre de décrire des images (ou des régions) via la construction d'une représentation multi-échelle. L'analyse de cette représentation, ou espace-échelle, peut être effectuée au niveau des pixels (description locale) ou des images (description globale).

Représentations multi-échelles

Les filtres morphologiques (tels que l'ouverture et la fermeture) ont pour effet de supprimer des détails de l'image, le niveau de détails supprimés variant en fonction de la taille de l'élément structurant considéré. De ce fait, ils peuvent être utilisés pour construire des représentations multi-échelles aussi appelées espaces-échelle.

Le concept d'espace-échelle [Wit83] est défini comme une famille d'images filtrées $\{\Upsilon_t(f)\}_{t \geq 0}$ (avec la convention $\Upsilon_0(f) = f$) vérifiant certains axiomes. Le plus souvent, cette représentation est construite par des convolutions de l'image f avec un noyau Gaussien $\mathbf{G}$:

$$\Upsilon_t(f)(x, y) = f(x, y) * \mathbf{G}(x, y, t) = \int_{-\infty}^{+\infty} f(u, v) \frac{1}{2\pi t^2} e^{-\frac{(x-u)^2 + (y-v)^2}{2t^2}} du dv \quad (2.1.39)$$

et on parle d'espace-échelle Gaussien. La propriété principale d'un espace-échelle est la causalité, i.e. aucune structure additionnelle n'est créée dans l'image lorsque t augmente (la hauteur et la position des extrema sont préservées). Outre cette propriété fondamentale, on a également la récursivité :

$$\forall t, s \geq 0, \quad \Upsilon_t(\Upsilon_s(f)) = \Upsilon_s(\Upsilon_t(f)) = \Upsilon_{t+s}(f) \quad (2.1.40)$$

mais aussi la croissance, et l'anti-extensivité ou l'extensivité. La propriété de récursivité est vérifiée (sous certaines conditions) par des opérateurs tels que l'érosion et la dilatation, mais ne l'est plus pour les filtres morphologiques. De ce fait, on choisit souvent de définir les espaces-échelle morphologiques en remplaçant cette propriété par la loi d'absorption [Mat75] :

$$\forall t, s \geq 0, \quad \Upsilon_t(\Upsilon_s(f)) = \Upsilon_s(\Upsilon_t(f)) = \Upsilon_{\max(t,s)}(f) \quad (2.1.41)$$

ce qui nécessite que l'élément structurant considéré soit un ensemble convexe compact, et en considérant la propriété d'idempotence $\Upsilon_t(\Upsilon_t(f)) = \Upsilon_t(f)$. L'intérêt des espaces-échelle morphologiques par rapport aux espaces-échelle Gaussiens tient au fait qu'ils préservent bien mieux les contours des objets (si on utilise des opérateurs connexes ou par reconstruction, cf. figure 2.5), permettant ainsi une analyse plus fiable du contenu des images.

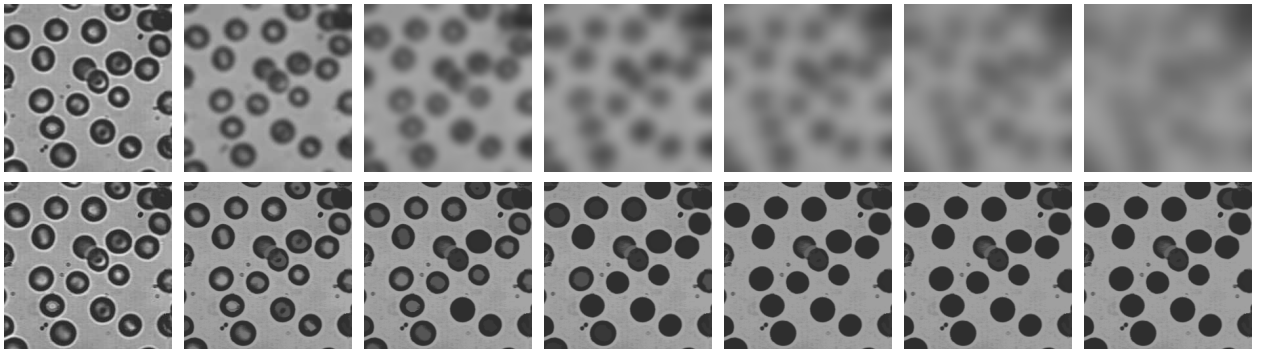


FIG. 2.5 – Espaces-échelle Gaussien (en haut) et morphologique par reconstruction (en bas) (seulement une échelle sur trois est représentée)

Notons $\Pi^\psi(f) = \{\Pi_\lambda^\psi(f)\}_{\lambda \geq 0}$ l'espace-échelle morphologique, i.e. la série d'images successives filtrées en utilisant le filtre ψ avec un élément structurant de taille croissante λ :

$$\Pi^\psi(f) = \left\{ \Pi_\lambda^\psi(f) \mid \Pi_\lambda^\psi(f) = \psi_\lambda(f) \right\}_{0 \leq \lambda \leq n} \quad (2.1.42)$$

où $\psi_0(f) = f$ et $n + 1$ est la longueur de la série (incluant l'image initiale). Il est possible d'utiliser différents filtres pour construire cet espace-échelle, avec naturellement des répercussions sur la série Π générée : ainsi, une série Π^γ sera anti-extensive tandis qu'une série Π^φ sera extensive. Pour éviter leur comportement hautement asymétrique (puisque la série Π^ψ est soit anti-extensive soit extensive), il est possible de regrouper les séries construites par des ouvertures et des fermetures afin d'obtenir une unique série Π de longueur $2n + 1$ (définie par exemple comme $\Pi(f) = \{\Pi_\lambda(f)\}_{-n \leq \lambda \leq n}$, les indices négatifs et positifs étant respectivement associés aux ouvertures et fermetures de taille $|\lambda|$).

Pour décrire le contenu d'une image, cette série est souvent ignorée au profit de sa version différentielle Δ^ψ qui permet de mettre en évidence les détails supprimés pour chaque taille λ :

$$\Delta^\psi(f) = \left\{ \Delta_\lambda^\psi(f) \mid \Delta_\lambda^\psi(f) = \left| \Pi_\lambda^\psi(f) - \Pi_{\lambda-1}^\psi(f) \right| \right\}_{0 \leq \lambda \leq n} \quad (2.1.43)$$

avec l'hypothèse $\Delta_0^\psi = 0$. Dans cette série, un pixel $\mathbf{p}$ apparaîtra (i.e. aura une valeur non nulle) dans $\Delta_\lambda^\psi(f)$ s'il est supprimé par ψ_λ , le filtre morphologique ψ de taille λ (ou autrement dit, s'il était présent dans $\psi_{\lambda-1}(f)$ mais plus dans $\psi_\lambda(f)$). Une unique série $\Delta(f)$ peut aussi être construite en tenant compte des filtres duaux anti-extensifs et extensifs.

Outre les filtres morphologiques, il est également possible de construire des représentations multi-échelles en utilisant n'importe quel opérateur morphologique. Cependant les propriétés de la série obtenue seront plus limitées. Parmi les opérateurs utilisés pour construire de telles représentations, on peut citer l'érosion et la dilatation, mais aussi les opérateurs de rehaussement de contraste (cf. [A2] pour un panorama des espaces-échelle morphologiques construits par des opérateurs structuraux, et leurs diverses applications en analyse d'image). De plus, d'autres opérateurs morphologiques tels que les nivellements peuvent être utilisés pour construire de telles représentations. Dans ce cas [MM00], les choix portent sur les opérateurs extensifs et anti-extensifs à la base du nivellement et la famille d'images marqueurs considérées (cf. équation (2.1.32)). Différents filtres peuvent être utilisés pour construire les images marqueurs : filtres alternés séquentiels, filtres Gaussiens, filtres de diffusion anisotropique, etc.

A partir de ces différentes séries peuvent être formulés l'ensemble des descripteurs morphologiques en considérant soit une analyse locale, soit une analyse globale.

Descriptions locales

L'approche la plus directe pour obtenir une description des images à partir d'une représentation morphologique multi-échelle consiste à utiliser simplement les valeurs des pixels aux différentes échelles. Ainsi, on définit le profil morphologique d'un pixel $\mathbf{p}$ d'une image f par le vecteur $\Pi^\psi(f)(\mathbf{p}) = (\Pi_\lambda^\psi(f)(\mathbf{p}))_{0 \leq \lambda \leq n}$ de taille $n + 1$. Dans le domaine de la télédétection, le profil morphologique différentiel proposé par Pesaresi et Benediktsson dans [PB01, BPA03] est construit sur ce principe en utilisant des séries différentielles basées sur des opérateurs par reconstruction, soit $\text{DMP}(f)(\mathbf{p}) = \Delta^\rho(f)(\mathbf{p})$. Il s'agit d'une caractéristique structurelle et d'une alternative intéressante aux caractéristiques spectrales ou texturales, dont la taille peut être réduite en se limitant aux maxima significatifs obtenus pour chaque pixel.

Dans le cas d'images binaires, le principe précédent est simplifié, et chaque pixel $\mathbf{p}$ de f est associé à l'échelle (ou taille λ du filtre morphologique) à laquelle il disparaît (en cas d'ouvertures) ou apparaît (en cas de fermetures). On parle alors de transformée d'ouverture, avec la définition $\Xi^\psi(f)(\mathbf{p}) = \max\{\lambda \geq 0 \mid \psi_\lambda(f)(\mathbf{p}) > 0\}$ et la convention $\Xi^\psi(f)(\mathbf{p}) = 0$ si $f(\mathbf{p}) = 0$. L'extension de ce principe aux images à niveaux de gris aboutit au concept d'arbre d'ouvertures (ou de fermetures) [Vin00], dont il est également possible de réduire la taille en se limitant aux valeurs les plus significatives.

Descriptions globales

Il est également possible (et en réalité plus fréquent) d'exploiter la représentation multi-échelle à un niveau global, en considérant l'image dans son ensemble. Ainsi, une granulométrie, concept issu des travaux en science des matériaux, a pour objet la mesure de la taille des particules élémentaires qui constituent les ensembles de grains de substances diverses et la définition des fréquences statistiques des différentes tailles de grains dans

l'ensemble étudié². En morphologie mathématique, une granulométrie peut être formulée en utilisant une série d'ouvertures γ :

$$\Omega^\gamma(f) = \left\{ \Omega_\lambda^\gamma(f) \mid \Omega_\lambda^\gamma(f) = \sum_{\mathbf{p} \in \mathcal{D}_f} \Pi_\lambda^\gamma(f)(\mathbf{p}) \right\}_{0 \leq \lambda \leq n} \quad (2.1.44)$$

et, de façon similaire, la formulation d'une anti-granulométrie Ω^φ utilise une série de fermetures φ . Les propriétés des filtres morphologiques permettent de garantir que la granulométrie Ω^γ est décroissante monotone tandis que l'anti-granulométrie Ω^φ est croissante monotone. Une normalisation peut intervenir pour rendre ces mesures invariantes à la taille de l'image, et ainsi obtenir des fonctions de distribution cumulative :

$$\Gamma^\psi(f) = \left\{ \Gamma_\lambda^\psi(f) \mid \Gamma_\lambda^\psi(f) = 1 - \frac{\Omega_\lambda^\psi(f)}{\Omega_0^\psi(f)} \right\}_{0 \leq \lambda \leq n} \quad (2.1.45)$$

Le spectre morphologique [Mar89] peut être vu comme le pendant morphologique de l'histogramme. Au lieu de mesurer la distribution des intensités (par exemple) au sein d'une image, son but est de mesurer la distribution des tailles (et dans une moindre mesure des formes). Il peut être défini par :

$$\Phi(f) = \left\{ \Phi_\lambda(f) \mid \Phi_\lambda(f) = \sum_{\mathbf{p} \in \mathcal{D}_f} \Delta_\lambda(f)(\mathbf{p}) \right\}_{-n \leq \lambda \leq n} \quad (2.1.46)$$

sa version normalisée nécessitant une division de chaque terme $\Phi_\lambda(f)$ par $\Omega_0(f)$.

Enfin, il est possible d'utiliser d'autres opérateurs morphologiques aux propriétés théoriques plus faibles tels que des érosions (on parle alors de pseudo-granulométries), comme dans le cas de la covariance morphologique. On considère ici un élément structurant B composé de deux points $\mathbf{p}_1$ et $\mathbf{p}_2$, et défini par une taille 2λ et une orientation $\vec{v}$. D'après Soille [Soi03], la covariance est alors formulée comme

$$K^{\vec{v}}(f) = \left\{ K_\lambda^{\vec{v}}(f) \mid K_\lambda^{\vec{v}}(f) = \sum_{\mathbf{p} \in \mathcal{D}_f} \Pi_{\lambda, \vec{v}}^\varepsilon(f)(\mathbf{p}) \right\}_{0 \leq \lambda \leq n} \quad (2.1.47)$$

avec $\varepsilon_{\lambda, \vec{v}}(f)(\mathbf{p}) = f(\mathbf{p} - \lambda \vec{v}) \wedge f(\mathbf{p} + \lambda \vec{v})$, tandis que Serra [Ser82] propose quant à lui d'utiliser la fonction d'auto-corrélation, en utilisant un produit plutôt qu'un infimum, soit $\varepsilon'_{\lambda, \vec{v}}(f)(\mathbf{p}) = f(\mathbf{p} - \lambda \vec{v}) \cdot f(\mathbf{p} + \lambda \vec{v})$.

2.1.6 Conclusion

Dans cette section, nous avons donné quelques rappels relatifs à la théorie de la morphologie mathématique et aux opérateurs qu'elle propose. Notre objectif n'était pas de dresser un panorama exhaustif, mais plutôt de montrer les motivations qui ont guidé nos travaux, et de détailler plus particulièrement les outils de base que nous avons utilisés. La morphologie mathématique étant un domaine de recherche très actif, des outils toujours plus poussés sont proposés régulièrement par la communauté (par exemple l'algorithme de ligne de partage des eaux présenté dans cette section peut être avantageusement remplacé par d'autres implantations : hiérarchiques, à base de graphes, modélisées par des équations aux dérivées partielles, etc.). Notons que même si nous n'avons pas nécessairement fait usage des outils les plus récents dans nos travaux, nos contributions nous semblent pouvoir aisément être adaptées à ces méthodes plus avancées.

Pour une présentation bien plus complète de la morphologie mathématique, de ses fondements théoriques à son application en traitement d'image, et des récents développements dans ce domaine, le lecteur pourra consulter le livre *Morphological Image Analysis* de Pierre Soille [Soi03] paru chez Springer en 2003 (deuxième édition), et l'ouvrage en français *Morphologie mathématique (tome 1 : approches déterministes)* édité par Laurent Najman et Hughes Talbot et paru chez Hermès – Lavoisier en 2008 [NT08]. Les actes de la conférence *International Symposium on Mathematical Morphology* (qui se tient tous les deux ans et demi environ) fournissent en outre un panorama représentatif des travaux de la communauté.

²Définition tirée de l'Encyclopédie Universalis (<http://www.universalis.fr>)

2.2 Approches morphologiques pour les images multivaluées

Cette section traite du problème de la prise en compte des images multivaluées en morphologie mathématique. Une partie importante des travaux présentés ici porte sur l'étude des ordres vectoriels (sections 2.2.2 et 2.2.3) : elle a été réalisée avec Erchan Aptoula dans le cadre de sa thèse de doctorat [Apt08], appliquée à l'imagerie couleur (section 3.1). La section 2.2.4 présente des approches alternatives (sans faire explicitement appel à un ordre vectoriel), qui ont été principalement appliquées en télédétection (section 3.2). Ces approches ont motivé directement ou indirectement différentes publications en revues internationales [B3, B4, B5] ou nationale [C4], en conférences internationales [D17, D19, D21, D25] ou nationales [E12, E16], et un chapitre d'ouvrage [A3].

Nous débuterons cette section par une introduction rappelant la problématique abordée et les notions relatives aux ordres vectoriels. Nous dresserons ensuite un bref panorama des ordres vectoriels, avant de nous focaliser sur les ordres totaux dits équilibrés. Nous présenterons alors plusieurs techniques de quantification nous permettant de limiter la dissymétrie induite par l'ordre total le plus connu, l'ordre lexicographique. Nous verrons également quelques approches permettant le traitement morphologique d'images multivaluées ne nécessitant pas d'appel explicite à un ordre vectoriel. Enfin nous donnerons quelques conclusions et les principales pistes de recherche que nous étudions en continuité des travaux présentés ici.

2.2.1 Introduction

Dans la section 2.1, nous avons rappelé les théories utilisées pour formaliser les opérateurs de morphologie mathématique : la théorie des ensembles et la théorie des treillis. Cette dernière est particulièrement adaptée pour définir les opérateurs morphologiques, qui sont formulés à l'aide d'opérations d'infimum $\wedge$ et de supremum $\vee$. Pour cela, il est nécessaire de disposer également d'une relation d'ordre $\leq$ pour ordonner les éléments du treillis.

Dans le cas d'images à une composante où chaque pixel est associé à une valeur numérique, la relation d'ordre $\leq$ est simplement l'ordre naturel entre scalaires. Cependant, dans le cas d'images multivaluées, où à chaque pixel sont associées plusieurs valeurs (numériques), établir un ordre entre deux pixels n'est plus trivial. En effet, comment devons nous ordonner dans l'image f les pixels p et q auxquels sont associées les valeurs $f(p) = (1, 2, 3)$ et $f(q) = (2, 2, 2)$? C'est à cette question que nous tenterons de répondre dans cette section.

Pour décrire les différentes possibilités qui nous sont données d'établir un ordre entre des vecteurs, nous complétons les définitions de réflexivité, d'anti-symétrie, et de transitivité d'une relation binaire $\leq$ données dans la section 2.1.2 (page 37) par celle de totalité. Une relation $\leq$ est :

$$\text{(totale)} \quad \forall a, b \in \mathcal{T}, a \leq b \text{ ou } b \leq a \quad (2.2.1)$$

c'est-à-dire lorsque nous sommes capables de comparer tous les éléments de l'espace $\mathcal{T}$ deux à deux. Une relation binaire $\leq$ qui est réflexive et transitive (mais pas anti-symétrique) est appelée un pré-ordre (ou un quasi-ordre). Si la contrainte d'anti-symétrie est également respectée, il s'agit d'un ordre. Cet ordre ou pré-ordre est dit total s'il garantit la propriété de totalité, partiel sinon. Nous utiliserons cette taxonomie pour présenter les principaux ordres vectoriels dans la section suivante.

2.2.2 Panorama des ordres vectoriels

Savoir ordonner des données multivaluées n'est pas un problème spécifique à la morphologie mathématique. Ce problème a donc été étudié depuis de nombreuses années, et on trouvera très souvent dans la littérature des références aux travaux de Barnett [Bar76]. Celui-ci a proposé une classification en quatre groupes non-exclusifs. Les M-ordres ou ordres marginaux sont obtenus en traitant indépendamment chaque dimension. Les C-ordres ou ordres conditionnels (ou séquentiels) correspondent au cas des vecteurs ordonnés à l'aide de priorités associées à chaque composante. Ainsi (en supposant des niveaux de priorité croissants), la composante de priorité i n'interviendra dans le processus de comparaison que si la composante de priorité $i - 1$ n'a pas permis de départager les valeurs comparées. Les P-ordres ou ordres partiels (même si le terme partiel n'a ici pas le même sens que dans la section précédente) partitionnent l'espace en groupes d'équivalence. Il est alors possible

de comparer des vecteurs de deux groupes différents, mais pas au sein d'un même groupe. Enfin, les R-ordres ou ordres réduits procèdent par réduction des vecteurs à des valeurs scalaires qui sont ensuite facilement comparables. Cette réduction peut être obtenue soit à l'aide d'une projection, soit à l'aide d'une mesure de distance. Cette classification reste d'actualité mais puisque les classes ne sont pas exclusives, nous lui préférons ici celle introduite par Chanussot [Cha98].

Notons qu'il est également possible de différencier les approches morphologiques (ou toute autre approche adaptée au cas d'images multivaluées) selon qu'elles traitent les images de façon marginale ou vectorielle. Dans le premier cas, une image multivaluée f est considérée comme un ensemble d'images monovaluées, soit $f = \{f_1, \dots, f_n\}$. Chacune d'elles est alors traitée de façon similaire à une image à niveaux de gris, sans tenir compte de la corrélation entre les bandes. L'ensemble des résultats obtenus compose l'image multivaluée finale. Dans le cas d'une approche vectorielle, il n'y a pas de décomposition et l'ensemble des composantes est pris en compte simultanément, comme illustré par la figure 2.6. Cependant cette stratégie nécessite d'adapter les algorithmes existants aux données vectorielles.

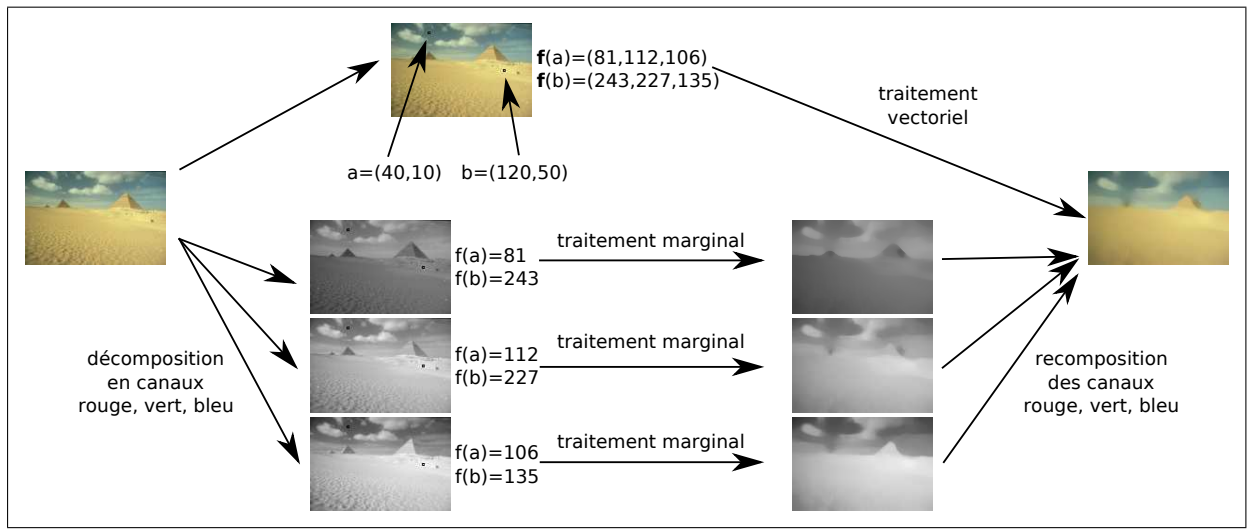


FIG. 2.6 – Comparaison des stratégies marginale (en bas) et vectorielle (en haut) pour le traitement (filtrage médian) d'une image couleur dans l'espace rouge – vert – bleu.

Nous présentons maintenant les principaux ordres vectoriels introduits dans la littérature depuis les travaux précurseurs de Goutsias et al. [GHS95]. Pour un traitement plus approfondi de ce sujet, le lecteur pourra consulter notre publication dans la revue *Pattern Recognition* [B3] .

Ordres partiels

La stratégie marginale décrite dans la figure 2.6 est en réalité un ordre partiel. En effet, dans le cas où $x = (3, 7)$, $y = (5, 9)$, et $z = (5, 5)$, nous avons $x \leq y$ et $z \leq y$, mais ni $x \leq z$ ni $z \leq x$. Cet ordre nous permet néanmoins de calculer les infimum et supremum nécessaires aux opérateurs morphologiques. Ainsi, $x \wedge z = (3, 5)$ et $x \vee z = (5, 7)$. En notant v et v' les vecteurs à ordonner dans un espace $\mathcal{T}$ de dimension n , on a

$$\forall v, v' \in \mathcal{T}, v \leq v' \Leftrightarrow \forall i \in \{1, \dots, b\}, v_i \leq v'_i \quad (2.2.2)$$

Cette stratégie est donc fréquemment employée lors de l'application d'opérations morphologiques à des images multivaluées. Elle a pour avantage une certaine simplicité d'utilisation puisqu'il est possible d'appliquer directement les algorithmes conçus pour les images à niveaux de gris. En contrepartie, elle se caractérise par deux inconvénients principaux.

Tout d'abord, le traitement marginal ou indépendant de chaque composante se traduit par la non-prise en compte de la possible corrélation entre les différentes bandes. Or cette corrélation est fréquemment observée, par exemple dans les images couleur RVB. Pour s'affranchir de ce problème, il est possible d'appliquer une transformation assurant une décorrélation des informations et ainsi obtenir un changement d'espace [GHS95].

La stratégie consistant à décomposer l'image en un ensemble de composantes fera l'objet de la section 2.2.4. Nous avons également étudié l'utilisation d'une étape de fusion des bandes après traitement marginal dans le contexte de la segmentation d'images astronomiques bi-bandes [D15] (cf. section 3.3).

D'autre part, cet ordre n'offre aucune garantie quant à la possible préservation des valeurs initiales dans l'image résultante. Ainsi, nous avons vu dans l'exemple précédent que $x \wedge z$ et $x \vee z$ sont tous deux différents de x et de z . Les valeurs retournées par les opérations d'érosion et de dilatation (ou tout autre opérateur morphologique) peuvent donc n'être pas présentes dans l'image initiale. Cet effet peut se traduire par l'apparition de nouvelles couleurs dans le cas d'images couleur, de nouvelles réponses spectrales dans le cas d'images multispectrales (par exemple, une réponse spectrale correspondant à un matériau non présent à la surface du sol lors d'une observation de la Terre !), etc. La figure 2.7 illustre ce problème. Un des moyens d'éviter l'introduction de nouvelles valeurs consiste à remplacer chaque valeur obtenue après traitement marginal par la valeur la plus proche dans l'espace d'origine (ce qui nécessite le choix d'une mesure de distance pertinente dans cet espace) [AO03].

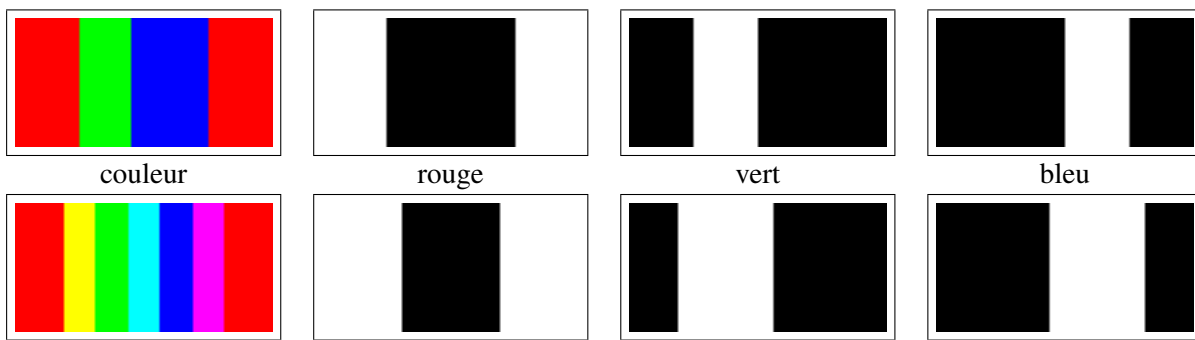


FIG. 2.7 – Apparition de fausses couleurs dans le cas d'une stratégie marginale : (en haut) image en couleurs et décomposition en composantes élémentaires ; (en bas) dilatation de l'image initiale obtenue de façon marginale, en dilatant chaque composante indépendamment.

Pré-ordres totaux

Choisir un ordre garantissant la propriété de totalité permet de s'assurer que tous les vecteurs soient comparables. Ainsi il est certain qu'aucune nouvelle valeur (ou couleur, ou réponse spectrale) ne sera introduite dans le résultat [TEJ98]. Les solutions basées sur des pré-ordres totaux sont néanmoins porteuses d'un autre inconvénient lié au relâchement de la contrainte d'anti-symétrie. Dans ce cas, il est possible d'obtenir plusieurs extrema équivalents. Pour illustrer cette situation, reprenons l'exemple des valeurs $x = (3, 7)$, $y = (5, 9)$, et $z = (5, 5)$. Si nous décidons d'ordonner les vecteurs selon leur norme (en optant pour la norme L_1 ou sommes des valeurs absolues), nous avons $\|x\| = 10$, $\|y\| = 14$, et $\|z\| = 10$, soit un supremum (y) mais deux choix possibles pour l'infimum. Autrement dit, comment choisir $\bigwedge (x, y, z)$ parmi x et z ? Différentes solutions ont été proposées dans la littérature, parmi lesquelles la sélection d'une des valeurs candidates en fonction de la position du pixel dans l'élément structurant [CD99] ou la comparaison des valeurs candidates à l'aide d'un ordre total [Ang07]. Ce dernier choix est sans doute le plus communément adopté.

Pour obtenir un pré-ordre, deux approches sont fréquemment suivies. La première consiste à utiliser une transformation de réduction non-injective h afin d'aboutir à un ordre réduit [GHS95, Cha98]. Cette transformation peut être une mesure de distance, et on peut alors décider d'ordonner deux vecteurs x et y selon leur distance à un vecteur de référence ; il faut donc choisir la mesure de distance d (Manhattan, Euclidienne, etc) et le vecteur de référence v_{ref} :

$$v \leq v' \Leftrightarrow d(v, v_{\text{ref}}) \leq d(v', v_{\text{ref}}) \quad (2.2.3)$$

Si celui-ci est défini comme l'origine de l'espace, la distance obtenue est équivalente à une norme (L_1 , L_2 , etc). Pour éviter le choix explicite d'un vecteur de référence, il est possible d'associer à chaque vecteur la somme

des distances aux autres vecteurs de l'ensemble à ordonner [PMPP04] :

$$\forall \mathbf{v}_k, \mathbf{v}_l \in \{\mathbf{v}_j\}, \quad \mathbf{v}_k \leq \mathbf{v}_l \Leftrightarrow \sum_j d(\mathbf{v}_k, \mathbf{v}_j) \leq \sum_j d(\mathbf{v}_l, \mathbf{v}_j) \quad (2.2.4)$$

avec $\{\mathbf{v}_j\}$ la famille de vecteurs considérés.

Cependant, l'infimum (ou le résultat d'une érosion) est alors représenté par l'élément central (ou médian), tandis que le choix du supremum est très sensible aux valeurs présentes dans l'espace. De plus, l'ordre obtenu n'en est plus réellement un au sens d'une relation binaire : les opérateurs construits à partir de ce type d'ordre seront alors des opérateurs pseudo-morphologiques.

La seconde approche consiste à établir un ordre conditionnel, c'est-à-dire en donnant une priorité à chaque composante. Lors du processus de comparaison, seules certaines composantes sont considérées (sinon il pourrait s'agir d'un ordre et non d'un pré-ordre), par exemple :

$$\mathbf{v} \leq \mathbf{v}' \Leftrightarrow v_1 \leq v'_1 \quad (2.2.5)$$

Appliquons cet ordre aux vecteurs $\mathbf{x}, \mathbf{y}, \mathbf{z}$ précédents. On obtient $\bigwedge(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{x}$ mais une indécision entre $\mathbf{y}$ et $\mathbf{z}$ pour $\bigvee(\mathbf{x}, \mathbf{y}, \mathbf{z})$. Si on ne considère que la seconde composante, les extrema sont ici uniques puisque $z_2 < x_2 < y_2$. Cependant, privilégier certaines composantes par rapport à d'autres (comme dans tous les ordres conditionnels) n'est pas toujours valable (quelle composante privilégier dans l'espace couleur RVB ?), le choix de la priorité associée à chaque composante étant dépendant de l'application considérée.

Ordres totaux

La dernière catégorie présentée ici concerne les ordres totaux. Contrairement aux ordres partiels, la propriété de totalité permet de tenir compte de la possible corrélation entre les bandes et de préserver les vecteurs d'origine. Contrairement aux pré-ordres, la propriété d'anti-symétrie garantit l'unicité des extrema. De nombreux travaux ont donc porté sur ces ordres aux aspects théoriques plus séduisants. En particulier, l'ordre lexicographique est certainement l'ordre total le plus connu. Il consiste à comparer deux vecteurs selon une composante considérée comme prioritaire. Si les deux valeurs comparées sont égales, la comparaison porte sur la composante de priorité directement inférieure. En cas de nouvelle égalité, cette comparaison est répétée jusqu'à étudier les valeurs dans la dernière composante (de plus faible priorité). S'il y a une nouvelle égalité, cela signifie que les vecteurs sont égaux (et la propriété d'anti-symétrie est bien vérifiée). Cet ordre peut être formalisé de la façon suivante :

$$\forall \mathbf{v}, \mathbf{v}' \in \mathcal{T}, \quad \mathbf{v} <_L \mathbf{v}' \Leftrightarrow \exists i \in \{1, \dots, n\}, \quad (\forall j < i, v_j = v'_j) \wedge (v_i < v'_i) \quad (2.2.6)$$

et nous adopterons la notation $V \rightarrow W \rightarrow X$ pour des vecteurs dans un espace VWX à 3 dimensions où la comparaison porte d'abord sur la dimension V , puis W et enfin X .

L'ordre lexicographique, comme tout ordre conditionnel, est adapté aux situations où il existe un ordre d'importance entre les différentes bandes : celles porteuses du plus d'information doivent être les plus prioritaires. Cet ordre peut être naturellement présent au sein de l'image d'origine, soit obtenu artificiellement après changement d'espace (construit par une transformation telle qu'une analyse en composantes principales, une représentation dans des espaces teinte – saturation – luminance dans le cas d'images couleur, etc). Cependant, cette possibilité d'accorder plus d'importance à certaines bandes de l'image est également l'inconvénient majeur de l'ordre lexicographique. Ainsi, la figure 2.8 illustre l'effet d'une dilatation avec un élément structurant de taille 5×5 pixels en utilisant différents ordres lexicographiques dans l'espace RVB.

Limiter l'importance donnée aux premières dimensions est donc primordial pour assurer à l'ordre lexicographique, en complément de ses intéressantes propriétés théoriques, une utilisation pratique réelle. Cette question fait l'objet de la section suivante.

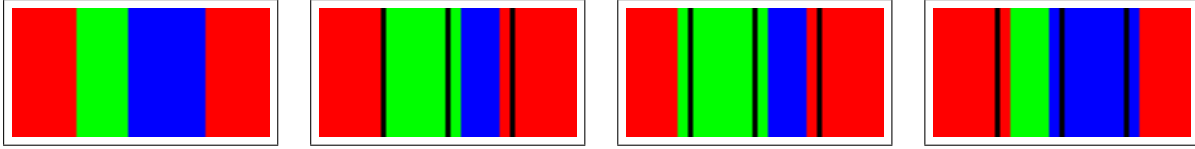


FIG. 2.8 – Importance de l'ordre des composantes dans l'ordre lexicographique : (de gauche à droite) image originale, et dilatation selon les ordres $R \rightarrow V \rightarrow B$, $V \rightarrow R \rightarrow B$ et $B \rightarrow R \rightarrow V$ (avec superposition des contours initiaux).

2.2.3 Ordres totaux équilibrés

Dans la section précédente, nous avons vu les trois principales familles d'ordre vectoriel : les ordres partiels qui ne garantissent ni la prise en compte de la corrélation entre les bandes, ni la préservation des vecteurs d'origine ; les pré-ordres totaux qui, en l'absence d'anti-symétrie, ne peuvent assurer l'unicité des extrema et ainsi aboutir à des indécisions quant au choix des résultats pour les érosions et dilatations par exemple ; et enfin les ordres totaux qui s'affranchissent des limites des précédents ordres, au prix toutefois d'une dissymétrie favorisant extrêmement la ou les premières bandes de l'espace.

Cette situation a motivé nos travaux sur les ordres totaux dits équilibrés, où la priorité accordée aux premières bandes est limitée ou peut être définie à partir de modèles de priorité. Nous débuterons cette section en présentant les solutions proposées dans la littérature et leurs limites, avant de décrire les variantes que nous avons élaborées, basées respectivement sur une quantification spectrale, spatiale, et itérative. Ces travaux ont montré leur intérêt dans différents problèmes de traitement d'image (filtrage, description de texture, classification basée pixels, etc.) et ont fait l'objet de plusieurs publications, notamment deux articles de revue dans *Pattern Recognition Letters* [B4] et *Journal of Visual Communication and Image Representation* [B5].

Le problème de l'ordre lexicographique

Comme indiqué précédemment, l'ordre lexicographique attribue une bien trop grande importance aux premières bandes de l'espace. Nous avons ainsi remarqué [B4] que, pour les images couleur généralement utilisées en traitement d'image (Lenna, Mandrill, etc.), lors d'une dilatation à l'aide d'un élément structurant carré de taille 21×21 pixels, la première composante (le canal rouge) était suffisante pour déterminer le pixel maximum dans 90 à 95 % des cas, tandis que la dernière composante (le canal vert) n'était utilisée dans le processus de comparaison que dans 0,5% des situations.

Différentes solutions ont été proposées dans la littérature pour limiter cette dissymétrie. Une première approche consiste à créer une bande artificielle représentative du contenu du vecteur et à lui affecter la priorité la plus importante. Ainsi, il est possible d'utiliser la norme Euclidienne du vecteur [Riv04] :

$$\mathbf{v} \leq \mathbf{v}' \Leftrightarrow [\|\mathbf{v}\|, v_1, \dots, v_n] \leq_L [\|\mathbf{v}'\|, v'_1, \dots, v'_n] \quad (2.2.7)$$

ou la distance à un vecteur de référence [Ang07]. Cette solution ne permet cependant pas à l'utilisateur de spécifier la priorité qu'il souhaite accorder aux différentes bandes.

Une seconde possibilité consiste à paramétrer (à l'aide d'une valeur notée α) l'importance de la première composante. Les propositions de Ortiz et al. [OTAP01] (ordre α -lexicographique) et d'Angulo et Serra [AS03] (ordre lexicographique modulo- α) respectivement définies par

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{R}^n, \alpha \in \mathbb{R}^+, \mathbf{v} < \mathbf{v}' \Leftrightarrow \begin{cases} v_1 + \alpha < v'_1, \text{ ou} \\ v_1 + \alpha \geq v'_1 \text{ et } [v_2, \dots, v_n] <_L [v'_2, \dots, v'_n] \end{cases} \quad (2.2.8)$$

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \alpha \in \mathbb{N}^+, \mathbf{v} < \mathbf{v}' \Leftrightarrow [\lceil v_1/\alpha \rceil, v_2, \dots, v_n] <_L [\lceil v'_1/\alpha \rceil, v'_2, \dots, v'_n] \quad (2.2.9)$$

augmentent la fréquence des comparaisons portant sur la seconde dimension en formant des groupes d'équivalence pour les valeurs de la première dimension. Notons que ces propositions ne sont pas des ordres totaux : la

première n'est pas transitive (il ne s'agit donc pas d'un ordre d'un point de vue algébrique) tandis que la propriété d'anti-symétrie est absente de la seconde. Ce dernier problème peut néanmoins être contourné en ajoutant la première dimension sous sa forme non-quantifiée à la fin de la cascade lexicographique. Cependant, même dans ce dernier cas, la fonction de quantification reste très simple (division entière) et suppose une partition uniforme de l'espace. De plus, elle ne porte que sur la première dimension. Nous nous sommes donc intéressés à des modèles plus évolués de quantification.

Quantification spectrale

Afin de permettre une modélisation plus complexe de la priorité accordée aux différentes bandes, nous avons généralisé l'ordre lexicographique modulo- α [AS03] en remplaçant la quantification uniforme obtenue avec une division entière par une quantification quelconque, soit :

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \quad \mathbf{v} < \mathbf{v}' \Leftrightarrow [w_1, v_2, \dots, v_n] <_L [w'_1, v'_2, \dots, v'_n] \quad (2.2.10)$$

où w_1 et w'_1 représentent respectivement les groupes d'équivalence de v_1 et v'_1 . Cette quantification peut naturellement porter sur les dimensions suivantes, soit :

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \quad \mathbf{v} < \mathbf{v}' \Leftrightarrow [w_1, w_2, \dots, w_n] <_L [w'_1, w'_2, \dots, w'_n] \quad (2.2.11)$$

Comme dans la section précédente, il faut ajouter à la fin de la cascade lexicographique les dimensions non-quantifiées si on souhaite garantir la propriété d'anti-symétrie.

La taille de chaque groupe d'équivalence, ou le facteur de quantification, est variable (mais limité par une valeur maximale α). Elle est définie précisément selon le contexte à l'aide d'une fonction f fournie par l'utilisateur (ou calculée à partir de l'histogramme de l'image par exemple, résultant en un ordre vectoriel adaptatif). Cette fonction modélise la distribution de priorité souhaitée entre les différentes bandes. La figure 2.9 illustre le cas d'une distribution exponentielle avec $\alpha = 10$, qui permet d'obtenir des groupes d'équivalence de taille variable (et croissante exponentiellement), contrairement à l'ordre lexicographique modulo- α . Notons que ce dernier peut être obtenu en utilisant la fonction constante $\forall n \in \mathbb{N}, f(n) = 1$.

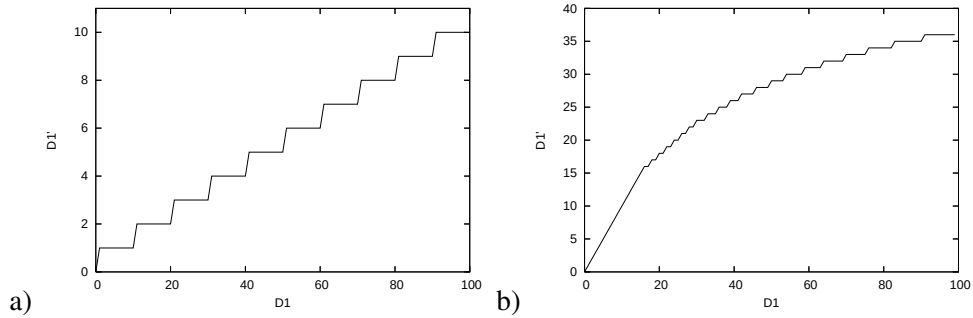


FIG. 2.9 – Exemple d'une quantification appliquée sur la dimension D1 (résultant en la dimension D1') en utilisant $\alpha = 10$ et l'équation (2.2.9) [AS03] (a), et en utilisant la quantification spectrale [B4] avec la même valeur de α et une distribution exponentielle (b).

Cette technique de quantification est particulièrement utile pour le traitement des images couleur dans les espaces perceptuels de type teinte – saturation – luminance. En effet, dans ces espaces, la pertinence de la teinte (ou de la saturation) peut dépendre du niveau de saturation (et/ou de luminance). Cependant, la création des groupes d'équivalence ne tient pas compte de l'information spatiale présente dans les images. Ainsi, les frontières entre les groupes d'équivalence peuvent ne pas correspondre aux frontières présentes dans les images. La figure 2.10 illustre ce problème dans le cas d'un signal discret monodimensionnel. On peut y observer la création de frontières artificielles entre les groupes d'équivalence, correspondant à de faibles variations du signal original.

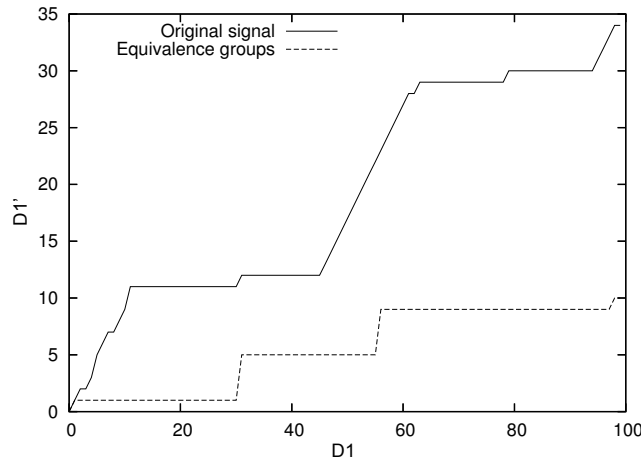


FIG. 2.10 – Exemple d'un signal discret monodimensionnel (dimension D1) et des groupes d'équivalence correspondants (dimension D1') obtenus par l'équation (2.2.9) et $\alpha = 11$.

Quantification spatiale

Pour résoudre le problème induit par le modèle de quantification spectrale, il est possible de créer les groupes d'équivalence en fonction du contenu de l'image à traiter. Plus précisément, la bande de l'image qui doit voir sa priorité réduite fait l'objet d'un prétraitement (un lissage par exemple). L'image obtenue par cette quantification spatiale est appelée un marqueur. En notant celui-ci g , l'ordre par quantification spatiale peut être formulé comme :

$$\forall \mathbf{p}, \mathbf{q} \in \mathcal{D}, \mathbf{f}(\mathbf{p}) < \mathbf{f}(\mathbf{q}) \Leftrightarrow [g(\mathbf{p}), f_2(\mathbf{p}), \dots, f_n(\mathbf{p})] <_L [g(\mathbf{q}), f_2(\mathbf{q}), \dots, f_n(\mathbf{q})] \quad (2.2.12)$$

Autrement dit, la formation des groupes d'équivalence n'est plus contrôlée que par l'image marqueur, qui peut être obtenue de différentes façons (avec un filtre médian, un filtre alterné séquentiel, un nivellement, etc). Les pixels seront donc comparés en premier lieu à l'aide de leurs valeurs dans cette image, et en cas d'égalité la comparaison portera sur les différentes bandes de l'image originale. Cette procédure permet de tenir compte des relations spatiales entre pixels puisque les groupes d'équivalence correspondent ici aux zones homogènes (de niveau constant) de g .

Quantification itérative

Les deux techniques précédentes permettent de quantifier spectralement ou spatialement certaines dimensions de l'image pour limiter leur influence au sein de l'ordre lexicographique. Nous nous sommes également intéressés à une stratégie de quantification itérative, où l'objectif n'est pas nécessairement d'obtenir un ordre au sens de ses propriétés algébriques mais plutôt un moyen pertinent de calculer les extrema d'un ensemble de vecteurs (et ainsi d'identifier les résultats d'une érosion ou d'une dilatation). En l'absence d'une relation d'ordre, on ne dispose pas d'une structure de treillis et les opérateurs qui découlent de cette proposition manquent de validité théorique. Néanmoins, ces opérateurs pseudo-morphologiques offrent un intérêt pratique certain.

Cette proposition suit le principe de l' α -troncature utilisée par exemple dans les filtres moyennés (où la moyenne est calculée à partir d'un ensemble de valeurs sans tenir compte de ses α plus petites et α plus grandes valeurs). En appliquant ce principe à la recherche du minimum ou du maximum d'un ensemble de vecteurs, on peut facilement limiter l'influence des premières dimensions de l'image. La procédure consiste alors à trier les k vecteurs selon leur valeur dans la première dimension à considérer, et à ne conserver que les $\lceil \alpha \times k \rceil$ plus grands ou plus petits vecteurs (selon que l'on recherche respectivement le maximum ou le minimum). Cette procédure est ensuite répétée sur les dimensions suivantes, diminuant à chaque dimension étudiée l'ensemble des vecteurs à comparer (ce qui permet en outre de limiter la complexité algorithmique). La dernière dimension permet d'identifier un unique minimum ou maximum si plusieurs vecteurs ont été conservés à l'issue de cette procédure. La figure 2.11 illustre ce principe dans le cas d'un espace à trois dimensions.

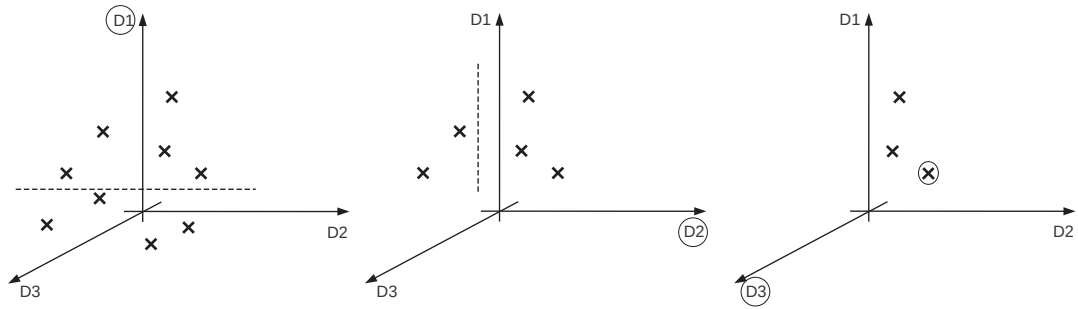


FIG. 2.11 – Un ensemble de vecteurs dans un espace 3-D $D1 \times D2 \times D3$, et les trois itérations (de gauche à droite) du calcul du maximum lexicographique suivant le principe de l' α -troncature, avec $\alpha = 0,5$. Selon l'approche proposée, le vecteur maximal est celui qui a la plus grande valeur dans la dernière dimension (parmi les trois vecteurs restants).

De la valeur de α dépend l'importance accordée aux différentes bandes (une valeur proche de 0 permet d'obtenir un comportement similaire à un ordre lexicographique, tandis qu'une valeur proche de 1 fera porter les comparaisons sur la dernière dimension). Notons qu'il est possible de considérer une quantification différente de chaque bande, ce qui nécessite de disposer de valeurs α_i pour chaque bande f_i de l'image. Ces paramètres α_i peuvent être définis en fonction de la distribution des valeurs des vecteurs dans les différentes bandes (et plus précisément de l'écart-type σ_i des valeurs de f_i) :

$$\forall i \in \{1, \dots, n\}, \quad \alpha_i = 1 - \frac{\sigma_i}{\sum_{j=1}^n \sigma_j} \quad (2.2.13)$$

Une dimension où les données sont très regroupées par rapport aux autres dimensions (donc peu discriminantes dans cette dimension) verra sa priorité diminuée avec une valeur α_i plus importante. Au contraire, une priorité plus grande sera accordée aux dimensions pour lesquelles la dispersion des données est la plus importante et auxquelles est associée une faible valeur α_i . Ce calcul des α_i peut être effectué suivant d'autres stratégies, par exemple en appliquant une analyse en composantes principales sur l'image initiale et en tenant compte des distributions des valeurs selon les différents axes du nouvel espace obtenu.

Différentes variantes peuvent être élaborées à partir du principe précédent [B5]. Ainsi, la sélection des $\lceil \alpha \times k \rceil$ valeurs extrêmes peut être remplacée par celle des valeurs situées à une distance aux extrêmes inférieure à α . De plus, la construction d'un ordre peut être obtenue en recherchant le maximum (ou minimum) d'un ensemble de vecteurs puis en répétant la recherche sur les vecteurs restants jusqu'à avoir ordonné complètement l'ensemble initial (qui peut contenir soit toutes les valeurs réellement présentes dans l'image, soit toutes les valeurs définies dans l'espace considéré).

2.2.4 Décompositions en ensembles d'images

Nous avons vu que la formulation des opérateurs morphologiques suivant la théorie des treillis nécessitait de disposer d'un ordre vectoriel pour comparer les différents vecteurs associés aux pixels d'une image multivaluée. Les ordres totaux équilibrés décrits dans la section précédente offrent des propriétés intéressantes, mais dans certains cas, il est difficile voire impossible de définir la priorité accordée aux différentes bandes. En outre, parmi les nombreuses alternatives proposées dans la littérature (notre état de l'art [B3] comporte une centaine de références), comment choisir l'ordre adéquat pour un problème donné ? Pour éviter de répondre à cette question délicate, on peut s'interroger sur la réelle nécessité de disposer d'un ordre vectoriel pour élaborer des approches morphologiques multivaluées, et chercher à traiter de telles images sans faire explicitement appel à un ordre vectoriel.

Nous présentons ici quelques tentatives que nous avons effectuées pour éviter l'utilisation explicite d'un ordre vectoriel lors du traitement d'images multivaluées. Dans ce cas, il faut disposer d'images binaires ou à niveaux de gris pour lesquelles sont déjà définis les différents opérateurs morphologiques. Pour cela, nous

considérons une décomposition de l'image initiale en ensembles d'images (binaires ou à niveaux de gris). Ces images sont alors traitées soit simultanément (ou conjointement) soit indépendamment les unes des autres. En outre, les traitements appliqués peuvent dépendre de chaque image ou être identiques pour toutes les images. Dans son application la plus simple, la stratégie marginale décrite en section 2.2.2 correspond au cas d'une image multivaluée décomposée en images à niveaux de gris, sur lesquelles sont appliqués de façon indépendante des traitements identiques.

Décomposition en images binaires

Décomposer une image multivaluée en un ensemble d'images binaires nous permet d'utiliser les opérateurs morphologiques formulés à l'aide de la théorie des ensembles. C'est cette approche qui avait été initialement suggérée pour traiter les images à niveaux de gris, au travers d'une décomposition en pile (cf. figure 2.2). Une telle représentation s'obtient à partir d'une image f à k niveaux de gris en appliquant k seuillages successifs avec des seuils croissants T_i , afin de construire les k images binaires F_i . Autrement dit, la définition $F_i = \{\mathbf{p} : f(\mathbf{p}) \geq T_i\}$ permet d'assurer la propriété $\forall i > j, F_i \subseteq F_j$. Traiter une image à niveaux de gris à partir de cette représentation est cependant coûteux (il faut appliquer k traitements identiques, un sur chaque image binaire). En outre, comment construire une pile d'images binaires à partir d'une image multivaluée ? La construction de n piles indépendantes (une pour chaque composante de l'image) correspond à l'application d'une stratégie marginale, tandis que la construction d'une seule pile induit l'utilisation de seuils croissants T_i multivalués : il faut alors savoir ordonner les vecteurs de l'espace multidimensionnel pour définir les valeurs T_i , et comparer ensuite chaque pixel $\mathbf{f}(\mathbf{p})$ avec les seuils T_i au moyen d'un ordre vectoriel !

Nous avons donc étudié une autre décomposition, où la relation d'ordre induite par la représentation en pile (soit $\leq$ entre les seuils T_i , ou $\subseteq$ entre les images binaires F_i) n'est pas nécessaire. La représentation utilisée ici est une partition, un pixel $\mathbf{p}$ ne pouvant appartenir qu'à une seule image binaire : $\forall i \neq j, F_i \cap F_j = \emptyset$. Bien qu'il soit possible de construire une image binaire pour chaque valeur différente de l'espace (c'est-à-dire $F_i = \{\mathbf{p} : \mathbf{f}(\mathbf{p}) = T_i\}$), nous préférons réduire le nombre d'images en définissant chaque image binaire comme un groupe de pixels similaires (au sens d'un certain critère). Ces groupes ou classes peuvent être obtenus à l'aide d'une procédure de classification supervisée ou non (on obtient alors des approches morphologiques supervisées qui feront l'objet de la prochaine section). Une fois les différents groupes obtenus, l'application des opérateurs morphologiques est réalisée de façon marginale sur chaque image F_i et résulte en la création d'autant d'images F'_i . Il est en outre possible d'utiliser des traitements ou des paramètres variables pour chaque image. Enfin, la construction du résultat $\mathbf{f}'$ s'obtient en combinant les différentes images F'_i , par exemple en remplaçant chaque pixel $\mathbf{p}$ de F'_i par sa valeur initiale $\mathbf{f}(\mathbf{p})$ ou une valeur représentative de sa classe F_i .

Nous avons appliqué cette stratégie dans [C4,D17] pour extraire les structures bâties d'une image de télédétection, en construisant les images binaires F_i par regroupement de pixels aux signatures spectrales homogènes, et en recherchant des formes prédéfinies dans ces différentes images (cf. section 3.2). Les formes ou les paramètres utilisés peuvent varier d'une image à l'autre, ce qui permet d'adapter (de façon automatique dans notre cas) le traitement à chacune des classes spectrales. Dans cette application, l'objectif était d'identifier les zones de bâti sous la forme d'une carte binaire : nous avons donc choisi de procéder à une combinaison des différents résultats F'_i à l'aide d'une opération d'union. Lors de cette expérimentation, nous avons observé les limites d'une telle décomposition : dans certains cas, les pixels d'un objet sont répartis dans différentes images binaires par l'effet de la classification. Pour résoudre ce problème, nous avons étendu l'ensemble d'images binaires $\{F_i\}$ en appliquant une opération d'union entre les images binaires pour en produire de nouvelles, soit $\forall i \neq j, F_{i,j} = F_i \cup F_j$. Ce principe peut naturellement être appliqué à des n -uplets d'images binaires, permettant alors la construction d'un treillis d'images binaires, tel qu'illustré par la figure 2.12. Dans cette représentation, nous pouvons faire l'hypothèse qu'un objet soit intégralement représenté dans au moins une des images binaires. En outre, la structure obtenue peut être un arbre si les n -uplets sont disjoints, ou si on applique directement une procédure de classification hiérarchique sur l'image multivaluée initiale.

Notons que nous n'avons pas réalisé d'étude théorique poussée sur ce type de représentation, qui peut être également assimilée à des images d'étiquettes, chaque étiquette étant associée à une des images binaires. L'application de la morphologie mathématique à des images d'étiquettes a fait l'objet de travaux bien plus approfondis par Ronse et Agnus [RA05,RA08], menés de façon indépendante à la nôtre. Leur démarche consiste

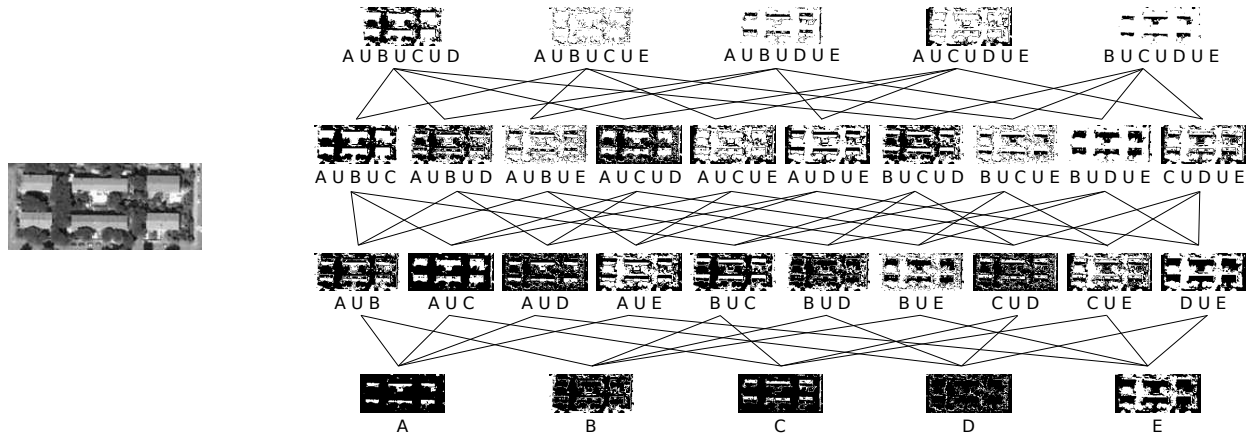


FIG. 2.12 – Illustration d'un treillis d'images binaires (à droite) construit par regroupements successifs (unions) à partir d'une décomposition obtenue par classification basée pixels d'une image à niveaux de gris (gauche).

à imposer un treillis sur les étiquettes, avec les plus petit et plus grand éléments définis respectivement comme l'absence d'étiquette ou un conflit entre plusieurs étiquettes, et sans autre relation d'ordre entre les étiquettes. Cette structure est intéressante mais ne semble cependant pas adaptée au cas d'une hiérarchie (arborescente ou non) d'images binaires.

Décomposition en images à niveaux de gris

Outre la représentation en ensemble d'images binaires, nous pouvons décomposer une image multivaluée f à n composantes en n images à niveaux de gris f_i . L'application d'un traitement morphologique (noté ψ_i) est alors effectuée sur chaque image f_i , résultant en une image $f'_i = \psi_i(f_i)$. Le résultat global f' est alors construit à partir des différentes images f'_i . La stratégie marginale (cf. figure 2.6) utilise ce type de décomposition, avec la particularité d'un traitement identique pour chaque image (soit $\forall i, \psi_i = \psi$) et l'utilisation de chaque résultat f'_i pour représenter la i ème composante de f' .

Dans [D25], nous avons mis en œuvre une approche similaire pour la détection d'objets dans des images multispectrales. Nous utilisons à cet effet l'opérateur de transformée en tout-ou-rien η , dont la définition a été donnée dans la littérature pour les images binaires et à niveaux de gris (cf. section 2.1.4). Dans le cas d'images multivaluées, l'application de cet opérateur procède aussi en deux étapes : l'ajustage ou *fitting* qui identifie les pixels correspondant au motif recherché, et la *valuation* qui associe à chaque pixel de l'image sa valeur définitive en fonction du résultat de l'ajustage et des valeurs de l'image initiale. Cet opérateur peut être défini à l'aide d'ordres vectoriels (nous avons proposé une telle définition dans [B7] que nous avons utilisée dans le contexte de l'imagerie couleur, cf. section 3.1.2) mais il est aussi possible d'éviter l'usage explicite d'un ordre vectoriel. Ainsi, en décomposant f en ses composantes f_i , chaque composante peut être traitée indépendamment, soit $f'_i = \eta_i(f_i)$. Notons cependant que lors de l'étape d'ajustage les différences entre les traitements η_i portent ici sur les éléments structurants utilisés et non sur les opérations morphologiques effectuées (il s'agit toujours de comparaisons entre érosions et dilations). Cette étape peut donc être assimilée à une stratégie marginale, puisqu'un résultat similaire serait obtenu en appliquant un unique traitement η avec des éléments structurants multivalués. L'emploi d'une stratégie marginale pourrait être étendu à l'étape suivante de valuation : il produirait alors une image f' où le résultat de la transformée en tout-ou-rien (représentant un score de mise en correspondance) serait multivalué et défini pour chaque composante f'_i en fonction de la qualité de l'appariement dans la composante f_i . Cependant, dans le contexte de la détection d'objets, l'usage d'une telle stratégie nous semble inapproprié : en effet, comment interpréter une qualité d'appariement exprimée de façon vectorielle ? Il est bien plus aisé d'évaluer en chaque pixel la qualité d'une détection à partir d'une unique valeur scalaire. En outre, l'application complète d'une stratégie marginale se traduit par l'absence totale de prise en compte de la corrélation entre les bandes. Nous avons donc choisi d'unifier (ou fusionner) les différents résultats f'_i en une seule image à niveaux de gris f' : cette image est facilement interprétable et combine

les informations présentes dans les différentes bandes [D25].

Afin d'éviter de ne tenir compte de la possible corrélation entre les bandes que lors de l'étape finale d'unification ou de fusion (dont la définition peut s'avérer particulièrement délicate), le traitement des différentes images à niveaux de gris obtenues par décomposition d'une image multivaluée peut également s'opérer de façon simultanée ou conjointe. Nous avons employé une telle approche pour la segmentation d'images multivaluées par ligne de partage des eaux [D19] (notons que cette approche est également représentative de nos travaux en morphologie mathématique supervisée, et sera donc aussi discutée dans la section suivante). Comme décrit dans la section 2.1.4, ce paradigme de segmentation consiste à simuler l'inondation d'un relief (représenté par une image à niveaux de gris) à partir de bassins de rétention initiaux pour produire la carte de segmentation. L'application de cette méthode de segmentation à des images multivaluées s'opère généralement en construisant le relief comme une image à niveaux de gris alors que l'image initiale est multivaluée, plutôt que de procéder directement à l'inondation d'un relief multivalué. Pourtant, cette dernière solution présente l'avantage de pouvoir conserver le plus longtemps possible l'information complète (i.e. multidimensionnelle). Pour pouvoir être appliquée, elle nécessite néanmoins l'utilisation d'un ordre vectoriel total. Pour éviter le choix d'un tel ordre, nous avons proposé une autre approche où le relief multivalué est décomposé en différents reliefs définis comme des images à niveaux de gris. Chaque bassin de rétention est initialement défini au sein d'une image à niveaux de gris, chacune de ces images ou reliefs pouvant contenir un nombre quelconque de bassins (en s'assurant du non-chevauchement spatial de l'ensemble des bassins initiaux définis dans les différentes images). Chaque bassin s'étend ensuite suivant l'application du processus d'inondation au relief sur lequel il a été défini, et en fonction d'un unique niveau d'eau commun à tous les reliefs (ce qui nécessite une normalisation préalable des différents reliefs). L'information qu'un pixel est atteint par l'extension d'un bassin (et qu'il sera donc affecté à la région correspondante) est propagée dans les différentes images, au contraire d'une inondation indépendante de chaque relief qui aurait produit une carte de segmentation par relief (cas d'une stratégie marginale). Ici nous obtenons une seule segmentation, résultat du traitement conjoint des différents reliefs. En reprenant les notations introduites page 43, nous formulons les reliefs g_i par :

$$g_i(\mathbf{p}) = \begin{cases} h_{\min-1} & \text{si } \mathbf{p} \in M_i \\ f_i(\mathbf{p}) & \text{sinon} \end{cases} \quad (2.2.14)$$

avec $M = \{M_i\}_{1 \leq i \leq c}$ la collection des marqueurs associés aux c reliefs. Les composantes connexes ou bassins de rétention $X = \{X^i\}_{1 \leq i \leq c}$ sont obtenus par le schéma de récursion suivant :

$$\begin{aligned} X_{h_{\min-1}}^i &= T_{h_{\min-1}}(g_i) \\ X_{h+1}^i &= IZ_{\{T_{h+1}(g_i)\}_{1 \leq i \leq c}}^i(X_h), \quad h_{\min-1} \leq h < h_{\max} \end{aligned} \quad (2.2.15)$$

en considérant une définition adaptée des zones d'influence :

$$IZ_A^i(B) = \bigcup_{j=1}^l iz_A^i(B_j^i) \quad (2.2.16)$$

$$\begin{aligned} iz_A^i(B_j^i) &= iz_A(B_j^i) \cup \bigcap_{m=1}^c \{p \in A^i \cap A^m \mid \\ &\forall k \in [1, l] : d_{A^i}(p, B_j^i) < d_{A^m}(p, B_k^m)\} \end{aligned} \quad (2.2.17)$$

2.2.5 Conclusion

Alors que les opérateurs morphologiques sont bien établis dans le cas des images binaires ou à niveaux de gris, il n'existe pas de définition unique pour les images multivaluées. Pourtant ces images sont présentes dans de nombreux domaines dans lesquels la morphologie mathématique pourrait être appliquée avec succès (notamment ceux auxquels nous nous sommes intéressés, et où une information spectrale est associée à chaque pixel : imagerie couleur, télédétection, imagerie astronomique – cf. sections 3.1 à 3.3). Dans cette section, nous avons donc abordé la question du traitement des images multivaluées à l'aide des opérateurs morphologiques.

Deux principales stratégies peuvent être suivies lors de l'adaptation des techniques morphologiques au cas multivalué. La stratégie marginale consiste à traiter indépendamment chaque composante ou bande de l'image multivaluée sous la forme d'une image monovaluée (telle qu'une image à niveaux de gris) pour créer la composante correspondante dans l'image résultat. Elle offre l'avantage d'une certaine simplicité puisqu'il n'est pas nécessaire d'adapter les traitements existants, déjà définis pour des images à niveaux de gris. Néanmoins elle présente deux inconvénients majeurs. D'une part, elle ignore la corrélation qui peut exister entre les différentes bandes de l'image. Ainsi, dans le cas d'une image multispectrale, les traitements effectués ne pourront pas s'appuyer sur les réponses ou signatures spectrales de chaque pixel. D'autre part, si l'objectif est de filtrer ou restaurer une image multivaluée, l'application d'une stratégie marginale peut résulter en l'apparition de nouveaux vecteurs dans l'image finale (c'est-à-dire des valeurs qui n'étaient pas présentes dans l'image initiale). Ce phénomène se traduit par la création de nouvelles couleurs dans le cas d'une image couleur, ou de nouvelles signatures spectrales dans le cas d'une image multispectrale ! Si l'objectif est différent (segmentation ou détection d'objets par exemple), le résultat n'est généralement plus une image multivaluée et il n'est nul besoin de conserver les valeurs d'origine. A l'opposé, la stratégie vectorielle considère l'image multivaluée comme l'élément de base du traitement, chaque pixel étant alors représenté non pas par une seule valeur scalaire mais par un vecteur (ses valeurs dans les différentes composantes de l'image). On respecte ainsi la nature intrinsèque de l'image en tenant compte de toute l'information disponible lors des traitements qui y sont appliqués (et on s'assure également de la préservation des vecteurs d'origine, évitant l'apparition de fausses couleurs ou réponses spectrales). Cependant, il est nécessaire d'adapter les traitements existants au cas multivalué. En morphologie mathématique, ces traitements s'appuient sur la recherche des minimum et maximum locaux ou plus généralement sur l'existence d'une relation d'ordre entre les pixels. L'application d'une stratégie vectorielle nécessite donc la définition d'une relation d'ordre vectoriel. Or, contrairement au cas des valeurs scalaires, il n'existe pas de relation d'ordre universelle entre vecteurs.

Nous avons donc mené une étude des ordres vectoriels et de leurs propriétés. Si la relation d'ordre n'est pas totale (i.e. certains vecteurs ne sont pas comparables), l'ordre est dit partiel. Cet ordre ne semble pas adapté aux opérateurs morphologiques puisqu'il ne garantit pas de pouvoir comparer tous les pixels pour y rechercher des extrema. Néanmoins ceux-ci peuvent être calculés de façon marginale (la stratégie marginale pouvant être formulée comme un ordre partiel) ou indépendante dans chaque composante de l'image. Il en résulte les inconvénients propres à la stratégie marginale : ni préservation des vecteurs d'origine, ni prise en compte de la corrélation entre les bandes. L'élaboration des opérateurs morphologiques dans le cas multivalué nécessite en réalité un ordre total [TEJ98]. Si celui-ci ne garantit pas la propriété d'anti-symétrie (i.e. les extrema d'un ensemble de vecteurs peuvent alors ne pas être uniques), l'ordre total est en fait un pré-ordre (ou quasi-ordre). Dans ce cas, il faut le plus souvent introduire des critères additionnels (ad hoc) pour départager les vecteurs identifiés comme extrema, de façon à pouvoir obtenir un unique minimum ou maximum (utilisé par exemple comme le résultat d'une érosion ou d'une dilatation). La solution la plus élégante d'un point de vue théorique consiste donc à utiliser un ordre qui soit à la fois total et anti-symétrique. L'ordre le plus représentatif de cette catégorie est certainement l'ordre lexicographique, fréquemment utilisé pour l'élaboration d'approches morphologiques multivaluées s'assurant de leur validité théorique. Il présente cependant un inconvénient majeur, qui consiste en la très (trop) forte importance accordée à la première dimension considérée lors du processus de comparaison des vecteurs puisque celui-ci ne portera sur les dimensions suivantes qu'en cas d'égalité sur la première dimension. En pratique, son application correspond fréquemment à l'utilisation exclusive de la première composante de l'image multivaluée !

Disposer d'une solution théoriquement élégante pour le traitement des images multivaluées n'est pertinent que si celle-ci est aussi intéressante pour résoudre des problèmes concrets posés en traitement d'image. En nous focalisant sur l'ordre lexicographique du fait de sa justesse théorique, nous avons cherché à accroître son intérêt pratique en tentant de limiter l'importance excessive accordée à la première dimension. Les différentes variantes que nous avons élaborées, regroupées ici sous le terme d'ordres totaux équilibrés, s'appuient sur un principe de quantification pouvant être appliqué aux différentes dimensions. Plus précisément, la quantification d'une dimension permet de créer des groupes d'équivalence entre pixels, et ainsi d'accroître la fréquence des comparaisons effectuées sur les dimensions suivantes. Différentes techniques de quantification ont été présentées : la quantification spectrale permet la prise en compte d'un modèle complexe de distribution de priorité

en fonction du contenu de l'image ou de l'espace vectoriel considéré (par exemple une représentation en teinte – saturation – luminance) ; la quantification spatiale intègre l'information spatiale (les valeurs des pixels et de leur voisinage) dans le processus de création de groupes d'équivalence ; enfin, la quantification itérative procède par élagages successifs dans les différentes composantes de façon à réduire itérativement le nombre de pixels candidats jusqu'à obtenir un seul minimum ou maximum. Outre les mêmes propriétés théoriques que l'ordre lexicographique, les ordres obtenus sont de plus facilement paramétrables et d'un intérêt pratique plus important.

Cependant, même si elle présente la plus grande validité théorique, l'emploi d'une stratégie vectorielle n'est pas toujours aisé : en effet, il faut alors choisir un ordre vectoriel parmi les multiples propositions introduites dans la littérature. En complément de nos travaux sur les ordres vectoriels, nous avons donc étudié s'il était possible d'appliquer les traitements morphologiques aux images multivaluées sans faire explicitement appel à un ordre vectoriel. Pour cela, nous avons procédé à une décomposition de l'image multivaluée en un ensemble d'images, binaires ou à niveaux de gris. La décomposition en images binaires peut être effectuée à l'aide d'une structure de pile, comme lors des premières tentatives d'application de la morphologie mathématique aux images à niveaux de gris. Cependant elle nécessiterait alors l'emploi d'un ordre vectoriel pour ordonner les différentes valeurs de seuils ou les images binaires correspondantes. Nous lui avons donc préféré une représentation sans ordre intrinsèque, où les images binaires sont construites comme une partition de l'image multivaluée initiale (par exemple à l'aide de techniques de classification auxquelles nous nous intéresserons également dans la prochaine section). Chaque image fait l'objet d'un traitement indépendant (voire spécifique) et une fusion ou unification des images traitées est effectuée pour obtenir le résultat final. Notons qu'en cas de traitement similaire appliqué sur chacune des images, cette décomposition pourrait bénéficier du formalisme développé pour l'application de la morphologie mathématique aux images d'étiquettes. Enfin, pour assurer une certaine robustesse à la qualité de la partition, nous avons proposé de créer de nouvelles images binaires par regroupement des images existantes, ou d'employer une classification hiérarchique pour obtenir une décomposition arborescente. La décomposition en images à niveaux de gris pourrait quant à elle s'apparenter à une stratégie marginale. Cette dernière n'en est en réalité qu'un cas particulier, avec un traitement identique appliqué indépendamment sur chaque composante, et la production d'un résultat consistant en une image multivaluée, où chaque composante est définie par le résultat du traitement marginal de la composante correspondante dans l'image initiale. Nous avons mis en œuvre différentes stratégies pour exploiter une telle décomposition en images à niveaux de gris. Ainsi il est possible d'appliquer des traitements indépendants sur chacune des images produites par la décomposition : ces traitements ne sont pas nécessairement identiques et peuvent être adaptés à chacune des images analysée. L'introduction d'une étape de fusion ou d'unification permet en outre d'assurer une certaine prise en compte de la corrélation entre les différentes composantes de l'image multivaluée, au contraire d'une stratégie purement marginale. Cette étape peut néanmoins être évitée si les images issues de la décomposition sont traitées dans leur ensemble, c'est-à-dire de façon conjointe ou simultanée. Nous avons illustré un tel scénario en présentant un nouvel algorithme de segmentation morphologique adapté au cas multivalué, tandis que l'application de traitements indépendants à des décompositions en images binaires ou à niveaux de gris a été considérée dans le contexte de la détection d'objets dans des images multispectrales.

Dans la figure 2.13, nous présentons un récapitulatif des différentes approches pouvant être employées pour le traitement morphologique des images multivaluées, qu'il s'agisse d'approches que nous avons étudiées ou non. En résumé, les conclusions que nous pouvons tirer de nos travaux menés dans ce domaine sont les suivantes :

- puisque la morphologie mathématique est généralement formulée dans le cadre de la théorie des treillis, son extension aux images multivaluées nécessite l'existence d'une structure de treillis entre les vecteurs ou autrement dit le choix d'un ordre vectoriel [B3] ; or, les ordres partiels n'assurent ni prise en compte de la corrélation entre les bandes, ni préservation des vecteurs d'origine, tandis que les pré-ordres ne garantissent pas l'unicité des extrema : il est donc préférable d'utiliser un ordre total tel que l'ordre lexicographique ;
- en dépit de ses intrinsèques qualités théoriques, l'ordre lexicographique présente l'inconvénient d'une forte dissymétrie entre les différentes composantes d'une image multivaluée (il accorde une extrême

approche vectorielle	approche par décomposition $\mathbf{f} \xrightarrow{f=\{f_i\}} \dots$							approche par fusion anticipée
totalité ?	puis analyse identique		puis analyse indépendante		puis analyse conjointe			$\mathbf{f} \xrightarrow{\prod_i f_i} f$
anti-symétrie ?	$\dots f_i \xrightarrow{\psi} f'_i \dots$		$\dots f_i \xrightarrow{\psi_i} f'_i \dots$		$\dots \{f_i\} \xrightarrow{\psi} \dots$			et analyse monovaluée
équilibrage ?	stratégie marginale	avec fusion	sans fusion	avec fusion	sans fusion	fusion explicite	fusion implicite	
$\mathbf{f} \xrightarrow{\psi} \mathbf{f}'$	$\dots \{f'_i\} \xrightarrow{\psi} \mathbf{f}'$	$\dots \prod_i f'_i \xrightarrow{\psi} f'$	$\dots \{f'_i\} \xrightarrow{\psi_i} \mathbf{f}'$	$\dots \prod_i f'_i \xrightarrow{\psi_i} f'$	$\dots \{f'_i\} \xrightarrow{\psi_i} \mathbf{f}'$	$\dots \{f'_i\} \xrightarrow{\prod_i f'_i} f'$	$\dots f'$	$f \xrightarrow{\psi} f'$

FIG. 2.13 – Synopsis des différentes approches multivaluées en morphologie mathématique. Les images multivaluées sont notées en gras, les opérations de composition/décomposition et de fusion sont respectivement représentées par les symboles $\{ \}$ et $\prod$, et les pointillés sont utilisés pour alléger l'écriture mathématique.

importance à la première composante) : il est néanmoins possible d'élaborer des variantes équilibrées de cet ordre total basées sur un principe de quantification ; cette quantification peut être spectrale ou spatiale [B4], mais également itérative suivant le principe du filtre par α -troncature [D21, B5] ;

- pour éviter l'emploi d'une stratégie vectorielle et le nécessaire choix d'un ordre vectoriel qui en découle, il est possible de procéder à une décomposition de l'image multivaluée en ensembles d'images à niveaux de gris et d'effectuer un traitement indépendant sur chacune de ces images, de façon similaire à une stratégie marginale ; ces traitements peuvent être spécifiques à chaque image, et suivis par une étape de fusion pour combiner les informations distribuées dans les différentes composantes [D25] ; pour éviter cette étape de fusion et tenir compte de la corrélation entre les bandes au plus tôt lors du processus d'analyse d'image, les différentes composantes peuvent également être traitées conjointement, notamment dans le cas d'une segmentation d'image multivaluée [D19] ; enfin, la décomposition peut aussi produire un ensemble d'images binaires, traitées ensuite de façon indépendante (voire spécifique) [D17, C4]. Cette dernière approche présente des similarités avec l'usage de la morphologie mathématique sur des images d'étiquettes (chaque image binaire représentant une étiquette).

2.2.6 Perspectives

Images hypervaluées

Les différentes approches présentées dans cette section ont été appliquées soit en imagerie couleur, soit en imagerie multispectrale. Dans les deux cas, le nombre de dimensions des images à traiter restant faible (3 pour les images couleur, jusqu'à 6 pour les images multispectrales), nous n'avons pas pu vérifier leur comportement lorsque le nombre de dimensions augmente considérablement. Nous souhaitons donc nous intéresser au cas des images caractérisées par un grand nombre de dimensions, comme les images hyperspectrales en télédétection ou en imagerie astronomique. Ces images *hypervaluées*³ doivent faire l'objet d'une attention particulière. En effet, il n'est pas évident que les solutions discutées dans cette section (qu'elles soient issues de la littérature ou le fruit de nos travaux) soit adaptées à de telles images.

Ainsi, l'emploi d'une stratégie marginale doit être définitivement proscrit puisque la corrélation entre les différentes composantes est souvent plus fréquente dans ces images et ne peut pas être ignorée. L'utilisation d'une décomposition en images binaires ou à niveaux de gris n'est sans doute pertinente en pratique que si le nombre d'images produites reste réduit. De plus le mécanisme d'unification, s'il est présent, peut s'avérer

³ce terme est certainement un abus de langage mais nous semble porteur de sens, le parallèle étant aisé avec les images multi- et hyperspectrales.

particulièrement complexe pour être adapté au nombre important d'images à fusionner. L'introduction d'une technique de réduction de dimension, basée par exemple sur une analyse en composantes principales (ACP), semble donc nécessaire pour appliquer ces approches à des images hypervaluées. Cependant cette assertion reste à vérifier puisque nous n'avons pas expérimenté ces approches sur de telles images, que ce soit dans un but de détection d'objets ou de segmentation.

Dans le cas d'une stratégie vectorielle, l'utilisation d'un ordre total devient délicat : en effet, l'augmentation du nombre de dimensions accroît le besoin d'un équilibre entre les différentes composantes. L'ordre lexicographique, même appliqué à l'issue d'une transformation de type ACP, ignorera généralement la majeure partie des informations contenues dans une image hypervaluée. Nos propositions, plus équilibrées, pourraient mieux exploiter des transformations de type ACP en adaptant leur comportement en fonction des informations retournées par ce type de transformation (valeurs propres par exemple). Elles n'ont cependant pas été conçues pour traiter un si grand nombre de dimensions. Face à cette situation, les ordres vectoriels intrinsèquement équilibrés comme les ordres réduits pourrait apporter une alternative intéressante, mais il faudrait leur assurer la propriété d'anti-symétrie pour pouvoir identifier les extrema de façon unique. En outre, il serait intéressant d'étudier la possibilité d'appliquer une technique de réduction de dimension, non pas globalement sur l'image lors d'un prétraitement, mais localement lors de chaque recherche de minimum ou maximum réalisée à l'aide d'un ordre vectoriel. Cependant, il n'est alors pas certain que ces opérateurs adaptatifs, construits par réduction locale de dimension, conservent les propriétés théoriques inhérentes au cadre morphologique.

Une image hypervaluée à 2 dimensions peut également être appréhendée comme une image monovaluée à 3 dimensions (ce raisonnement étant bien sûr également valable pour un plus grand nombre de dimensions) : chaque pixel de l'image hypervaluée devient alors une ligne de voxels dans le volume monovalué. Cette représentation suppose que la proximité des dimensions d'une image hypervaluée ait un sens (ce qui est généralement le cas, par exemple avec les images hyperspectrales) puisqu'elle va se traduire par un voisinage spatial créé artificiellement dans la troisième dimension. Dans une telle représentation, la valeur produite par un traitement morphologique pour un pixel et une composante donnée sera basée sur les valeurs de ce pixel et de ses voisins dans la composante considérée, mais aussi dans les composantes voisines. Notons que ce principe de construction d'un espace artificiel par augmentation du nombre de dimensions a déjà été employé dans différents domaines, notamment pour la classification par SVM (*support vector machines*). Il serait intéressant d'étudier comment les avancées réalisées dans ces domaines pourraient être appliquées au traitement morphologique d'images hypervaluées.

Dans le cas d'images hypervaluées, la multitude des informations à traiter par les méthodes morphologiques pourrait se traduire par des coûts de calcul rédhibitoires. Aussi, nous étudions avec Stéphane Genaud de l'équipe ICPS⁴ du LSIIT les implantations parallèles des approches morphologiques afin de garantir un niveau d'efficacité élevé. Les possibilités offertes par les implantations parallèles ou distribuées sont nombreuses : exécution sur des grilles de calcul ou sur des architectures multicœurs.

Séquences d'images

Les images multivaluées peuvent en outre contenir une information temporelle. Chaque pixel est alors décrit par un vecteur représentant son évolution au cours du temps. En télédétection, de telles images sont appelées images multitemporelles : elles peuvent être traitées avec les approches décrites dans cette section (que ce soit celles utilisant un ordre vectoriel ou celles procédant par décomposition d'images), les traitements étant basés sur la signature temporelle de chaque pixel (et non plus sa couleur ou réponse spectrale).

Plus généralement, ces images contenant une information spatio-temporelle sont des séquences d'images (ou séquences vidéo). Notons que chacune des images de la séquence peut elle-même être de nature multivaluée (séquence d'images couleur par exemple). Dans une séquence d'images, lorsque l'objectif est de décrire chaque pixel et son évolution au cours du temps (comme dans le cas des images multitemporelles), ou de regrouper les pixels en zones homogènes au sens de cette évolution, les approches introduites dans cette section restent pertinentes. Cependant il est plus fréquent de s'intéresser au contenu d'une séquence vidéo au travers des objets qu'elle contient. Les éléments à extraire et à décrire sont alors des régions de l'image qui apparaissent ou

⁴Image et Calcul Parallèle Scientifique, <http://icps.u-strasbg.fr>

disparaissent au cours de la séquence, et dont la forme et la position évoluent. Les approches morphologiques discutées dans cette section semblent ici moins adaptées et le traitement des séquences vidéo par la morphologie mathématique nécessite une étude spécifique.

Pour analyser une séquence d'images, trois principales stratégies peuvent être employées. La première consiste à traiter de façon marginale chaque image de la séquence, puis à effectuer un post-traitement pour combiner les résultats obtenus sur les différentes images et ainsi tenir compte de la dimension temporelle. Le résultat obtenu est donc très sensible à la qualité du post-traitement dont le rôle est ici crucial, ce qui limite grandement l'intérêt de cette approche. La seconde stratégie ($2D + t$) s'approche également d'une stratégie marginale : on traite indépendamment chaque image de la séquence, mais le traitement d'une image donnée tient compte des résultats obtenus sur l'image précédente (ou sur les k dernières images analysées). Ainsi, l'utilisation d'un post-traitement devient inutile, l'information temporelle étant utilisée pour initialiser chaque traitement d'image fixe. Enfin, il est également possible d'opter pour une représentation de la séquence sous forme volumique ou $3D$, telle que suggérée précédemment dans le cas des images hypervaluées. Une telle représentation est le plus souvent pertinente, puisque la fréquence d'acquisition des images garantit généralement l'existence d'une connexité temporelle des occurrences d'un même objet au cours du temps. Chaque objet correspond alors à une région spatio-temporelle de la séquence. Notons que dans ce dernier cas, l'information temporelle peut être exploitée au sein même de la définition des opérations morphologiques. Ainsi, Laveau [Lav05] propose d'utiliser des éléments structurants adaptatifs où, pour un pixel donné, la forme de l'élément structurant spatio-temporel dépendra des vecteurs de mouvement associés aux pixels voisins du pixel étudié.

Nous nous intéressons à ces deux dernières stratégies ($2D + t$ et $3D$), notamment pour la segmentation spatio-temporelle de séquences vidéo dans le cadre de la thèse de doctorat Jonathan Weber dont l'objectif est la structuration ou classification automatique d'ensembles de séquences d'images. Afin de privilégier une approche *objet* (c'est-à-dire de baser le processus de classification sur les objets contenus dans les séquences), une étape de segmentation est nécessaire. Cependant les contraintes associées à une telle segmentation sont nombreuses : le volume des données à traiter est important, ce qui nécessite des algorithmes efficaces, exploitant par exemple une représentation multi-échelle et/ou l'utilisation de graphes d'adjacence de régions ; la segmentation ne doit pas produire une multitude de régions définies comme des zones spatio-temporelles de contenu homogène mais plutôt un ensemble d'objets sémantiques sur lesquels sera basée l'étape ultérieure de classification : une des difficultés est alors l'introduction de cette information sémantique dans les traitements (nous aborderons cette question dans la prochaine section).

La segmentation spatio-temporelle est aussi une des solutions que nous souhaitons mettre en œuvre pour effectuer le suivi d'écoulements gravitaires, dans le cadre d'un projet BQR soutenu par le conseil scientifique de l'Université de Strasbourg et porté par Alexandre Remaître de l'IPGS⁵.

Images d'étiquettes

Dans le cas des images hypervaluées ou des séquences d'image, comme dans celui des images multivaluées discuté dans cette section, chaque pixel est représenté par une information numérique. Mais il existe également des images où la nature de l'information est symbolique. Ainsi, dans les images résultant d'une procédure de classification (supervisée ou non, cf. section suivante), chaque pixel peut être associé à une classe et ainsi représenté par une étiquette ou un identifiant de classe. Nous avons mentionné les travaux de Ronse et Agnus [RA05, RA08] sur la morphologie mathématique appliquée à des images d'étiquettes, qui pourrait être appliquée à des images issues d'une classification. Il nous semble intéressant de vérifier la mise en œuvre réelle d'approches morphologiques sur des images d'étiquettes, notamment à partir de cette formalisation, et d'identifier les traitements morphologiques pertinents pour ce type d'image.

En outre, il existe de nombreux paradigmes de classification et les images qui en résultent sont de nature très différente : ainsi, un pixel peut ne se voir attribuer aucune étiquette (classification partielle) ou au contraire plusieurs étiquettes (classification douce), un degré d'appartenance à chacune des classes ou autrement dit un vecteur de nombre réels (classification floue), une étiquette issue d'une représentation arborescente (classi-

⁵Institut de Physique du Globe de Strasbourg — UMR7516 CNRS-UDS, <http://eost.u-strasbg.fr/IPGS>

fication hiérarchique), etc. Ces images, où la représentation de chaque pixel se compose d'une ou plusieurs valeurs numériques et/ou symboliques, doivent faire l'objet d'une attention particulière. Ainsi, il n'est pas certain qu'elles puissent être correctement analysées par les approches morphologiques existantes, que ce soit celles définies pour les images multivaluées ou les images d'étiquettes. De nouvelles méthodes, de nouvelles formalisations semblent nécessaires.

Réussir à traiter correctement des images de nature très variée est un problème délicat à résoudre. Néanmoins il ouvre des perspectives particulièrement intéressantes pour l'élargissement des champs d'application de la morphologie mathématique, vers de nouveaux types d'images mais aussi hors du traitement d'image, comme nous l'indiquons dans notre programme de recherche (section 4.3).

2.3 Intégration de connaissances en morphologie mathématique

Cette section traite de l'intégration de connaissances en morphologie mathématique, qui est généralement effectuée de façon implicite par le concepteur de solutions morphologiques, répondant à des problèmes d'analyse et traitement d'image posés dans un contexte applicatif particulier : les connaissances portent alors sur le choix des opérateurs morphologiques et des éléments structurants. Le problème de l'intégration de connaissances nous a été posé dans le cadre de la thèse de doctorat de Sébastien Derivaux [Der09], où nous avons notamment proposé d'introduire une classification supervisée floue pour guider un processus de segmentation d'image en télédétection. Nous nous sommes également intéressés, dans le cadre d'applications en télédétection et imagerie astronomique, aux connaissances pouvant être introduites au sein des éléments structurants. Enfin, nous avons étudié les apports réciproques de la classification et de la segmentation, en proposant d'une part une méthode de segmentation guidée plus efficacement par des marqueurs, et d'autre part une reformulation de la classification sous la forme d'une segmentation itérative.

Les travaux concernant la classification ont fait l'objet de publications en conférence internationale [D14, D19, D20] ou nationale [E10, E11, E13, E15, E16], et d'un chapitre de livre à paraître [A4]. L'intégration de connaissances au travers des éléments structurants a pour sa part été discutée dans différentes publications, sans en être cependant le sujet principal, en revues internationales [B7, B9] ou nationale [C4], et en conférences internationales [D17, D25] ou nationale [E9].

Nous débuterons cette section en rappelant les principales connaissances pouvant être introduites en traitement des images dans le contexte des approches morphologiques. Nous nous intéresserons ensuite à l'intégration de connaissances au travers des éléments structurants, en précisant les différents types de connaissances que ces derniers peuvent contenir. Outre l'utilisation d'éléments structurants, nous verrons comment les techniques de classification, qu'elles soient supervisées ou non, dures ou floues, peuvent guider des approches morphologiques, notamment en modifiant l'espace de représentation des pixels sur lequel sont appliqués les traitements morphologiques. En complément d'un changement d'espace réalisé en amont des approches morphologiques, nous présenterons également quelques liens entre classification et segmentation morphologique : la reformulation de la classification non-supervisée comme une méthode de segmentation itérative, et l'exploitation des marqueurs au travers d'une classification supervisée pour guider plus efficacement le processus de segmentation par ligne de partage des eaux. Nous terminerons cette section par quelques conclusions et dresserons une liste de perspectives aux travaux qui y sont décrits.

2.3.1 Introduction

Pour résoudre un problème d'analyse d'image posé dans un contexte particulier, il est souvent utile, parfois primordial, de tenir compte d'informations disponibles a priori ou de connaissances du domaine. Ces connaissances peuvent être introduites de façon implicite ou explicite par le traiteur d'image et/ou l'expert du domaine.

Dans le premier cas, elles se traduisent par des choix quant aux algorithmes à utiliser ou à certains de leurs éléments (critères, mesures, etc.). Les méthodes élaborées sont alors difficilement réutilisables dans d'autres contextes, qui nécessiteront plutôt la définition de nouvelles solutions ad hoc. En morphologie mathématique, cette approche est fréquemment rencontrée puisque le spécialiste sait comment choisir, combiner et paramétrer les opérateurs morphologiques à bon escient pour résoudre un problème donné. En général, la définition des élé-

ments structurants (forme et taille), le choix des opérateurs ou algorithmes, et enfin le réglage de paramètres ou de seuils sont les principaux éléments sur lesquels le spécialiste agit pour adapter le traitement morphologique au contexte applicatif.

Au contraire, les connaissances formulées explicitement, sous la forme de paramètres (sémantiques ou aisément compréhensibles et manipulables par l'utilisateur), de données d'apprentissage, voire de représentations plus complexes exploitant des ontologies, permettent une certaine indépendance des algorithmes utilisés au contexte considéré : les méthodes développées dans un domaine donné, tenant compte de connaissances associées à ce domaine, peuvent ainsi être réutilisées dans un autre cadre applicatif, en modifiant les connaissances mises en œuvre [A1]. Nous nous intéressons à cette stratégie et discutons ici les principaux types de connaissances pouvant être introduits dans les approches morphologiques.

Différents types de connaissance

En morphologie mathématique, comme en analyse et en traitement d'image dans le cas général, des connaissances de nature variée peuvent être exploitées pour offrir des réponses plus adaptées à des problèmes donnés. La nature de ces connaissances peut dépendre du type de problème considéré, et nous nous intéressons aux cas de la détection d'objet et de la segmentation en régions, auxquels nous avons été confrontés dans plusieurs contextes applicatifs (cf. sections 3.1 à 3.3). Nous dressons ici une liste non-exhaustive des connaissances pouvant être intégrées dans les approches morphologiques (en nous focalisant en particulier sur les connaissances dont nous avons fait usage dans les contextes applicatifs auxquels nous nous sommes intéressés).

La description des objets ou régions à extraire d'une image, lorsqu'elle est (partiellement) effectuée a priori, peut ainsi être liée à la forme ou à la taille des objets. Ces informations, qu'il est relativement aisé d'intégrer dans les éléments structurants [BZQ95, BV00, NPR07b], peuvent également être exploitées au sein de méthodes de segmentation (par exemple sous la forme de critères pour la fusion de régions, ou de contraintes dans le processus de segmentation [Bea06]).

La taille des objets peut également être appréhendée au travers de l'échelle à laquelle les images sont analysées (nous avons discuté des descripteurs multi-échelles dans la section 2.1.5), tandis que leur position peut faire l'objet de marqueurs dans le cadre de la segmentation par ligne de partage des eaux [RBD92]. La définition des marqueurs est une des finalités les plus fréquemment rencontrées en intégration de connaissances en morphologie mathématique (cf. par exemple [LC02]).

Les objets ou régions peuvent aussi être caractérisés par leur composition spectrale ou texturale. Ainsi, les éléments structurants peuvent contenir une information spectrale (en associant une fonction structurante à chaque bande spectrale de l'image), mais plus difficilement une information texturale [NPR07b]. Cette dernière information est plus aisément exploitable au travers d'indices de texture, pouvant être extraits d'une image initiale pour produire une nouvelle image sur laquelle les algorithmes seront appliqués (cette démarche étant également pertinente dans le cas d'une information spectrale [LZ07]).

Les critères précédents sont des exemples représentatifs des éléments de connaissance pouvant être intégrés dans une modélisation des classes d'intérêt dans une image (ainsi Grau et al. proposent une modélisation bayésienne de chaque région [GMA⁺04]). La connaissance des modèles de classes et/ou du nombre de classes peut être exploitée au travers de processus de classification réalisés en amont des approches morphologiques (comme des changements d'espace). Nous verrons par la suite que la connaissance peut porter non pas sur le nombre de classes d'intérêt mais sur le nombre d'objets d'intérêt (ou occurrences de ces classes) présents dans l'image.

En complément des connaissances relatives aux objets ou classes d'intérêt (i.e. leurs caractéristiques, leurs modèles ou leur nombre), il est possible de tenir compte d'informations disponibles a priori et relatives à la nature des images analysées. Ainsi, les caractéristiques du bruit (niveau, type) présent dans l'image, la distribution des valeurs (niveaux de gris, couleurs, etc.), même s'ils ne sont pas associés aux éléments à extraire, apportent une information supplémentaire pouvant aider à résoudre le problème d'analyse d'image considéré.

Enfin, indépendamment de leur nature, on distingue deux types de connaissances : les connaissances expertes et les données d'apprentissage. Les premières peuvent consister en une base de connaissances ou plus simplement un jeu de règles ou de paramètres fourni par l'expert suite à une démarche de formalisation de ses connaissances. Les secondes peuvent prendre la forme d'un ensemble d'exemples visuels auxquels l'expert est

en mesure d'associer une sémantique particulière, ces exemples étant alors utilisés comme un jeu d'apprentissage dans un processus de classification supervisée. Notons que ces deux types de connaissances peuvent être exploitées conjointement [Mai05], même si nous n'avons pas observé de telle combinaison dans le cadre de la morphologie mathématique.

Par la suite, nous nous intéresserons à ces deux types de connaissances en discutant la nature des connaissances formalisées au moyen des éléments structurants, l'apport d'une classification non-supervisée où seul le nombre de classes ou d'objets est connu, et enfin l'utilisation d'exemples au travers de l'intégration d'un processus de classification supervisée dans la segmentation morphologique.

2.3.2 Formalisation par les éléments structurants

En morphologie mathématique, l'intégration de connaissances prend souvent la forme d'une spécification d'éléments structurants adaptés au problème considéré (par exemple, aux objets recherchés dans le cas de la transformée en tout-ou-rien). Différents types d'éléments structurants peuvent être utilisés en fonction des connaissances disponibles.

Dans sa formulation la plus simple et la plus courante, l'élément structurant est défini comme un ensemble B . Il apporte alors seulement une information de taille et de forme. Son utilisation la plus fréquente concerne les images binaires où seules ces informations sont disponibles, mais il est également à la base des opérateurs de la morphologie mathématique plate dans les images à niveaux de gris.

Dans les images à niveaux de gris, l'élément structurant peut également être représenté comme une fonction b en introduisant un profil d'intensité, en complément des informations précédentes qui sont utilisées pour définir le support ou domaine de définition de la fonction structurante.

L'information spectrale est exploitée pour sa part à l'aide des éléments structurants vectoriels $\mathbf{b}$, qui permettent d'adapter la spécification des connaissances précédentes (forme, taille, profil d'intensité) à chaque composante spectrale de l'image. Dans certains cas, il est plus judicieux de s'intéresser au profil d'intensité spectrale plutôt que spatiale : on peut alors utiliser une fonction structurante constante spatialement (mais pas spectralement) sur son support. Notons néanmoins que la plupart des approches morphologiques multivaluées de la littérature, qu'elles soient marginales ou vectorielles, n'utilisent pas de tels éléments structurants : elles supposent des éléments structurants identiques (de même forme et de même taille) dans chaque bande, et peu d'entre elles utilisent des fonctions structurantes (pourtant utiles pour spécifier les motifs recherchés [B7]).

Dans certains cas, il peut aussi être pertinent de disposer d'éléments structurants composites, représentant différents ensembles ou fonctions (à valeurs scalaires ou vectorielles) élémentaires. Ainsi, la transformée en tout-ou-rien utilise une composition de deux éléments structurants (un pour la forme ou l'image, un pour le fond ou le complémentaire de l'image). Même si les types précédents (B , b , $\mathbf{b}$) permettent la définition des éléments structurants composites, il est souvent plus aisé de formuler ces derniers à partir de leurs composants élémentaires, comme c'est le cas pour la transformée en tout-ou-rien et ses deux éléments structurants (il s'agit ici d'un cas particulier, puisqu'il est possible de définir des opérateurs morphologiques tenant compte de plus de deux composants élémentaires, comme nous l'avons fait pour la détection d'objets en télédétection en utilisant des connaissances spatiales et spectrales [D25]).

Notons que si l'élément structurant est anisotrope, il contient naturellement une information directionnelle liée à son orientation dans l'image. Si une telle information n'est pas souhaitée, il faut élaborer des opérateurs robustes, à la manière des filtres radiaux qui s'appuient sur un ensemble d'éléments structurants orientés (et retournent l'infimum ou supremum des résultats fournis en utilisant les différents éléments structurants). Dans certains cas, ces filtres peuvent être implantés efficacement en exploitant tout d'abord une version isotrope de l'élément structurant, puis en limitant l'utilisation des éléments structurants orientés aux résultats obtenus [B9].

En outre, il est possible que les connaissances disponibles soient moins précises et qu'elles ne soient pas suffisantes pour la définition correcte d'un élément structurant unique. De façon similaire au cas de l'optimisation des traitements exploitant des éléments structurants anisotropes, il faut alors considérer différents éléments structurants. L'application des opérateurs morphologiques avec chacun de ces éléments structurants nécessite une étape de fusion pour produire le résultat final. Ainsi, lorsque seule une forme ou une taille approximative des objets recherchés est connue, des résultats propres à chaque taille ou forme d'élément structurant sont calculés (nous avons mis en œuvre une telle approche pour la détection des bâtiments en télédétection [D17]).

Le profil d'intensité peut également varier, accroissant encore le nombre d'éléments structurants utilisés (nous avons été confronté à une telle situation en détection des galaxies [B9], pour laquelle il a été également nécessaire d'exploiter l'information relative au niveau de bruit). Lorsque seul le nombre de classes d'intérêt est connu, la fusion peut s'opérer sur des résultats spécifiques à chaque classe d'intérêt, et nous traitons dans les paragraphes suivants de cette dernière possibilité.

2.3.3 Connaissance limitée au nombre de classes d'intérêt

Il n'est pas toujours possible de formaliser les connaissances au travers des éléments structurants. Ainsi, nous pouvons être confrontés à des situations extrêmes où seul le nombre de classes d'intérêt est connu. Ces classes peuvent être de différentes natures, et nous avons par exemple traité le cas de classes spectrales en télédétection. Notons que les considérations et méthodes discutées dans la suite de cette section ne sont en aucun cas spécifiques à une nature de classe particulière.

Exploiter l'information relative au nombre de classes d'intérêt s'effectue généralement au travers d'une classification non-supervisée (ou *clustering*), dont nous rappelons ici le principe avant de préciser comment nous l'avons utilisée en morphologie mathématique. Nous présentons également le cas plus original où les classes d'intérêt sont directement assimilées aux régions de l'image : la classification est alors une segmentation.

Rappel du principe de la classification non-supervisée

La classification non-supervisée (ou *clustering*) consiste à regrouper des données d'un ensemble $\mathcal{X}$ en un ensemble $\mathcal{C}$ d'amas homogènes c appelés classes (ou *clusters*). Contrairement au cas supervisé, elle ne s'appuie pas sur des connaissances a priori relatives aux classes à construire (modèles de classes, exemples d'apprentissage), et elle peut procéder par partitionnement ou par hiérarchie au travers de nombreuses méthodes proposées dans la littérature. Pour plus d'informations sur ce sujet, on pourra consulter l'article très complet de Jain et al. [JMF99], ou sa version actualisée [Jai09].

Nous nous sommes intéressés essentiellement aux approches par partitionnement qui produisent une unique partition ω des données analysées, qui peut être de nature dure ou floue. Dans le premier cas, chaque élément $\mathbf{x}$ se voit affecté à une et une seule classe c , et on a alors $\omega : \mathcal{X} \times \mathcal{C} \rightarrow \{0, 1\}$, avec $\forall \mathbf{x} \exists c, \omega(\mathbf{x}, c) = 1$ et $\forall c' \neq c, \omega(\mathbf{x}, c') = 0$. Le second cas fournit, pour chaque élément à classer, son degré d'appartenance aux différentes classes, soit $\omega : \mathcal{X} \times \mathcal{C} \rightarrow [0, 1]$ avec $\forall \mathbf{x}, \sum_{c \in \mathcal{C}} \omega(\mathbf{x}, c) = 1$.

Nos contributions aux approches morphologiques basées sur une classification non-supervisée ont fait principalement l'objet d'expérimentations avec l'algorithme *k*-means, certainement le plus connu des algorithmes de *clustering* [Jai09]. Elles sont néanmoins complètement indépendantes de l'algorithme utilisé (sauf pour l'approche introduite dans [E15] qui est une reformulation du *k*-means pour la segmentation).

L'algorithme *k*-means, qui s'appuie sur un nombre de classes connu a priori, représente chaque classe c par son centroïde $\mathbf{x}_c$ initialisé aléatoirement ou non. Il procède de façon itérative (et jusqu'à convergence) en deux étapes : l'affectation de chaque élément $\mathbf{x}$ à la classe c qui minimise une distance $d(\mathbf{x}, \mathbf{x}_c)$, puis la mise à jour de chaque centroïde $\mathbf{x}_c$ à partir des éléments $\mathbf{x}$ tels que $\omega(\mathbf{x}, c) = 1$ (par exemple en définissant $\mathbf{x}_c$ comme le centre de gravité de ces éléments).

Changement d'espace de représentation

L'utilisation la plus directe des techniques de classification non-supervisée consiste à modifier l'espace de représentation des pixels. Leurs valeurs, qu'elles soient scalaires $f(\mathbf{p})$ ou vectorielles $\mathbf{f}(\mathbf{p})$, sont alors remplacées par leurs classifications ω . Dans le cas d'une classification dure, l'image obtenue peut être une image d'étiquettes (ou de classes) $f' : \mathcal{D} \rightarrow \mathcal{C} : \mathbf{p} \mapsto f'(\mathbf{p}) = \{c : \omega(f(\mathbf{p}), c) = 1\}$, et être analysée à l'aide du cadre développé par Ronse et Agnus [RA05, RA08]. Elle peut également être représentée sous la forme d'un ensemble d'images binaires $\{F'_c\}$, chacune d'elles étant associée à une classe donnée, soit $\forall c \in \mathcal{C}, F'_c = \{\mathbf{p} : \omega(f(\mathbf{p}), c) = 1\}$. Une telle décomposition a été discutée dans la section 2.2.4 et nous l'avons utilisée pour la détection de bâtiments dans des images de télédétection, en faisant l'hypothèse que le

nombre de classes spectrales était connu a priori [C4, D17]. Nous avons ainsi transformé l'image initiale, à niveaux de gris ou multispectrale, en un ensemble d'images binaires (représentant les différents matériaux ou types de sol) sur lesquelles ont été appliqués différents opérateurs morphologiques tels qu'une granulométrie bidimensionnelle, une transformée en tout-ou-rien à différentes échelles (différentes tailles et formes d'éléments structurants), et une reconstruction géodésique.

Si la classification est floue, le résultat obtenu peut être représenté sous forme d'une image multivaluée f' : on associe alors à chaque pixel ses degrés d'appartenance aux différentes classes (chaque classe correspondant à une bande de l'image), soit $f'(\mathbf{p}) = \{f'_c(\mathbf{p})\}$ avec $f'_c(\mathbf{p}) = \omega(f(\mathbf{p}), c)$. Pour disposer de valeurs à niveaux de gris, une étape de discrétisation peut être nécessaire (puisque initialement les degrés d'appartenance sont des réels dans $[0, 1]$). Cette image multivaluée peut être analysée selon une stratégie marginale ou vectorielle (cf. section 2.2), afin d'atteindre différents objectifs : filtrage, détection, segmentation, description, etc. Ainsi nous avons procédé à la segmentation d'images satellite en appliquant la technique de ligne de partage des eaux sur une carte de classification floue (ou plus précisément, sur la norme euclidienne du gradient morphologique calculé marginalement sur chaque bande de l'image, ou carte des degrés d'appartenance associée à chacune des classes) [E10]. L'intérêt d'une telle approche est de modifier le critère de regroupement des pixels en régions : ces dernières ne sont plus constituées de pixels aux réponses spectrales fortement similaires, mais de pixels aux classifications spectrales proches. Ce critère plus général permet de réduire la sursegmentation mais présente un intérêt pratique relativement limité (du moins d'après nos observations). Il peut néanmoins être avantageusement remplacé par l'utilisation de classes supervisées, que nous décrirons dans la section 2.3.4.

Lorsque les classes sont les régions

Dans les approches précédentes, nous avons considéré que les classes d'intérêt étaient connues. Nous pouvons également faire l'hypothèse que la connaissance porte plutôt sur le nombre d'objets (ou d'occurrences des classes d'intérêt) présents dans l'image. Dans ce cas, la classification non-supervisée peut être couplée au cadre morphologique pour identifier ces objets, c'est-à-dire pour segmenter l'image en régions. En effet, puisque la segmentation (morphologique) consiste à regrouper les pixels en groupes homogènes, elle peut être assimilée à une classification non-supervisée. Nous avons étudié le lien existant entre segmentation et classification, en nous focalisant sur l'algorithme bien connu des k -means.

L'utilisation classique de la classification en segmentation consiste à partitionner l'espace $\mathcal{T}$ des valeurs des pixels. L'obtention des régions nécessite alors de les construire à partir des composantes connexes dans la carte de classification (chaque région étant constituée de pixels adjacents affectés à une même classe) [HS85]. Nous avons proposé dans [E15] d'appliquer la classification directement dans l'espace des pixels $\mathcal{D}$ pour produire la segmentation. Chaque classe obtenue est alors un amas de pixels ou autrement dit une région. Pour que ce principe présente un intérêt, les critères utilisés pour l'association d'un élément (pixel) à une classe (région) et pour la mise à jour des centres des classes doivent être adaptés au problème de la segmentation qui suppose la constitution de classes connexes.

L'emploi d'une distance purement spatiale (telle que la distance euclidienne calculée sur les coordonnées des pixels) est à proscrire puisque celle-ci résulterait en une classification ne tenant pas compte du contenu de l'image. Pour combiner ces deux informations (positions spatiales et valeurs des pixels), nous avons choisi d'utiliser une distance topographique, par exemple une version simplifiée de la mesure de Philipp-Foliguet [PF00] :

$$d(\mathbf{p}, \mathbf{q}) = |f(\mathbf{p}) - f(\mathbf{q})| + \varepsilon, \quad \mathbf{q} \in \mathcal{N}_8(\mathbf{p}) \quad (2.3.1)$$

avec $\mathcal{N}_8(\mathbf{p})$ le 8-voisinage du pixel $\mathbf{p}$, et ε une constante très faible permettant de s'assurer que la mesure proposée soit bien une distance. Notons cependant que des mesures plus évoluées (basées sur la couleur, la texture, le gradient, etc.) pourraient être utilisées. A chaque itération de l'algorithme k -means, la distance est calculée entre chaque pixel et les différents centres de classe afin d'associer le pixel à la classe la plus proche. La mise à jour des centres exploite également cette distance : cependant, le centre d'une classe n'est plus choisi comme son centre de gravité (ce qui nécessiterait la recherche du point minimisant la somme des distances topographiques aux autres points de la classe), mais comme son médian. Ce dernier peut être obtenu par élagage de la classe, en supprimant itérativement les points extrêmes (sur le bord) et ne conserver au final

que l'élément central. Une implantation efficace des deux étapes d'affectation et de mise à jour peut être assurée en utilisant des cartes de distance topographiques construites à l'aide d'algorithmes à base de file d'attente, de façon similaire à la segmentation par ligne de partage des eaux. L'approche proposée pourrait ainsi être assimilée à une segmentation morphologique itérative.

Les figures 2.14 et 2.15 illustrent quelques caractéristiques de l'approche proposée. A l'instar d'autres méthodes de classification, il faut connaître le nombre de classes ou régions présentes dans l'image, au risque de produire une sur- ou une sous-segmentation. Cependant, contrairement aux usages actuels de la classification en segmentation, aucun post-traitement n'est nécessaire pour obtenir la segmentation.



FIG. 2.14 – Formalisation de la segmentation comme une classification et influence du nombre de classes (de gauche à droite) : sous-segmentation (6 classes), segmentation correcte (11 classes) et sur-segmentation (16 classes).

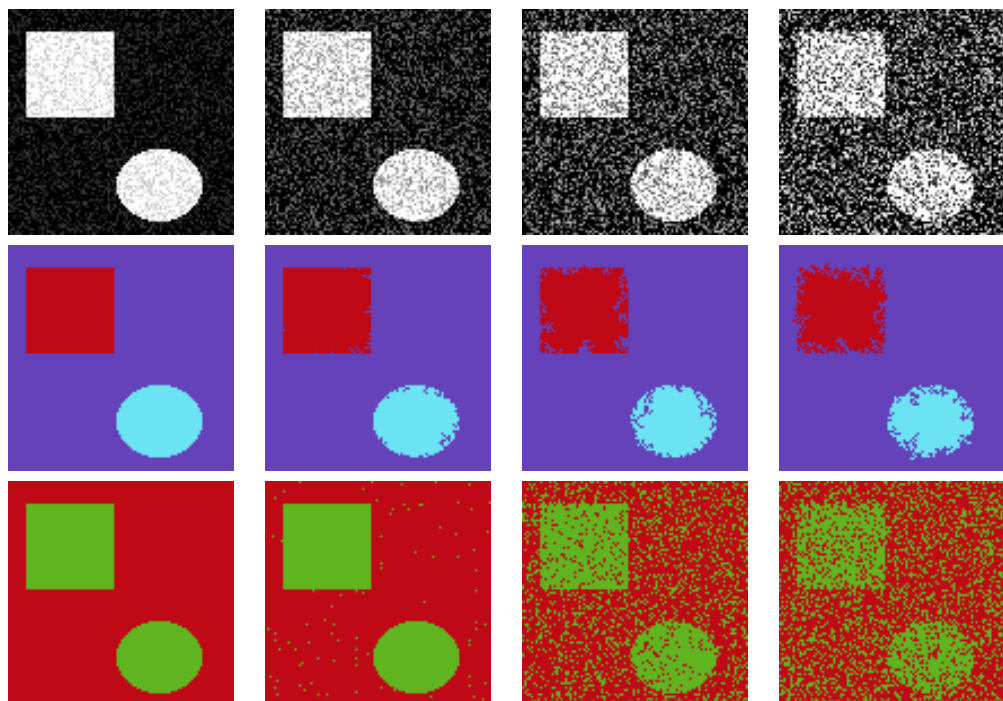


FIG. 2.15 – Comparaison entre segmentation et classification, pour différents niveaux de bruit (de gauche à droite). De haut en bas : image dégradée par un bruit Gaussien, classification proposée en 3 classes (régions), classification par k -means en 2 classes d'intensité (où les régions sont les composantes connexes de chaque classe).

2.3.4 Utilisation de régions de référence

Nous nous sommes intéressés précédemment à l'apport de la classification non-supervisée pour l'intégration de connaissances relatives au nombre de classes ou d'objets d'intérêt. Naturellement il est aussi possible d'exploiter des techniques de classification supervisée, dont nous rappelons le principe ici. Nous présentons ensuite différentes approches que nous avons élaborées pour la segmentation supervisée, avec ou sans prise en compte de l'information spatiale, et intégrant la phase de classification à différentes étapes du processus de segmentation (avant, pendant, ou après la segmentation). Notons que l'utilisation de la classification supervisée en segmentation morphologique n'est pas nouvelle, des approches similaires ayant été déjà introduites dans la littérature (par exemple [LC02, LZ07, GMA⁺04]) : cependant, celles-ci limitent généralement l'usage de la classification à la définition des marqueurs, à une séparation en deux classes (objets et fond), ou encore supposent une correspondance un-pour-un entre les marqueurs et les classes d'intérêt.

Rappel du principe de la classification supervisée

Contrairement au cas non-supervisé, la classification supervisée suppose que des exemples des différentes classes à construire soient déjà disponibles. Une phase d'apprentissage permet d'exploiter ces exemples pour produire des modèles de classe, qui sont ensuite utilisés lors de la phase de reconnaissance (ou classification) dont le but est d'associer à chaque nouvel élément la classe qui lui correspond.

Dans le cadre d'un processus d'analyse d'image, l'utilisation d'une classification supervisée se traduit souvent par l'implication de l'expert ou de l'utilisateur pour fournir les exemples d'apprentissage, par exemple des pixels ou des zones de l'image. Nous avons fait ici l'hypothèse que l'expert fournissait aux approches morphologiques supervisées des zones de l'image étiquetées, sans doute plus aisées à saisir que les pixels pris individuellement.

De nombreuses méthodes de classification supervisée, fournissant des partitions dures ou floues, ont été introduites dans la littérature [DHS01]. Les contributions que nous présenterons par la suite s'appuient sur l'utilisation d'une classification supervisée floue, mais sont indépendantes de l'algorithme de classification utilisé. Précisons néanmoins que les expérimentations que nous avons menées ont exploité une version floue de l'algorithme des k plus proches voisins [KGG85].

Dans cet algorithme, on recherche, pour chaque élément $\mathbf{x}$ à classer, ses k plus proches voisins $\mathbf{y}_i$ parmi les exemples d'apprentissage. Connaissant la classe c de chaque voisin $\mathbf{y}_i$ (celle telle que $\omega(\mathbf{y}_i, c) = 1$), on tient compte de la distance entre $\mathbf{x}$ et $\mathbf{y}_i$ dans le calcul du degré d'appartenance de $\mathbf{x}$ à c : plus $\mathbf{x}$ est proche de $\mathbf{y}_i$ (relativement aux autres plus proches voisins), plus il est probable que $\mathbf{x}$ appartienne à la classe c . Autrement dit, on définit la fonction de décision par $\omega(\mathbf{x}, c) = \sum_{i=1}^k \omega(\mathbf{y}_i, c) d(\mathbf{x}, \mathbf{y}_i)^{-1} / \sum_{i=1}^k d(\mathbf{x}, \mathbf{y}_i)^{-1}$.

Sans prise en compte de l'information spatiale

Nous avons proposé différentes méthodes pour l'intégration de connaissances en morphologie mathématique au travers de l'utilisation d'un processus de classification supervisée. Les exemples d'apprentissage étant définis comme des zones de l'image étiquetées par l'expert, il est possible d'exploiter uniquement leur contenu (les valeurs des pixels composant ces zones) ou également leur position. Nous nous intéressons ici à la première possibilité, que nous avons expérimenté dans le contexte de la télédétection (figure 3.11 page 93). Dans ce domaine, il est rarement fait usage de la position des exemples d'apprentissage dans l'image lors d'un processus de classification supervisée (fréquemment employé en interprétation d'image à basse résolution, cf. section 3.2).

De façon similaire au cas non-supervisé, la classification peut être utilisée en amont de la segmentation pour transformer l'espace de représentation des pixels. La méthode de segmentation, telle que la ligne de partage des eaux, opère alors non pas sur l'image initiale mais sur l'image des degrés d'appartenance aux classes. L'intervention de l'utilisateur, au travers des exemples d'apprentissage qu'il fournit, guide le processus de segmentation en modifiant le critère de similarité des pixels (qui doivent être homogènes au sens de leurs degrés d'appartenance aux classes) utilisé pour la construction des régions. Les attributs utilisés pour représenter les pixels lors de la classification peuvent être de différente nature, et nous avons employé par exemple leurs réponses spectrales [D14] ou des indices de texture calculés dans leurs voisinages [E13].

Il est aussi possible d'effectuer la classification supervisée après la segmentation [E11, D20]. Dans ce cas le processus de classification ne porte plus sur les pixels mais sur les régions produites par la segmentation. Les données d'apprentissage permettent de faciliter l'évaluation de la qualité de la classification obtenue, et ainsi de corriger si besoin le résultat de segmentation, ou autrement dit les paramètres utilisés pour fournir celui-ci. Dans le cas de la technique de ligne de partage des eaux, de nombreuses techniques ont été introduites dans la littérature pour réduire la sursegmentation et améliorer la qualité du résultat obtenu. Chacune de ces techniques nécessite de fixer un ou plusieurs paramètres, ces derniers ayant une influence majeure sur la qualité de la segmentation produite. Nous avons donc proposé de fixer ces paramètres en s'appuyant sur un processus d'optimisation par algorithmes génétiques [Gol89] : chaque individu représente un ensemble de valeurs associées aux paramètres de réduction de sursegmentation, et son évaluation s'appuie sur la qualité de la classification appliquée à la segmentation produite avec le jeu de paramètres qu'il représente.

Les deux approches précédentes sont indépendantes et peuvent être combinées afin d'exploiter au mieux la connaissance de l'expert introduite au moyen des exemples d'apprentissage. La figure 2.16 résume les différentes alternatives auxquelles nous nous sommes intéressés.

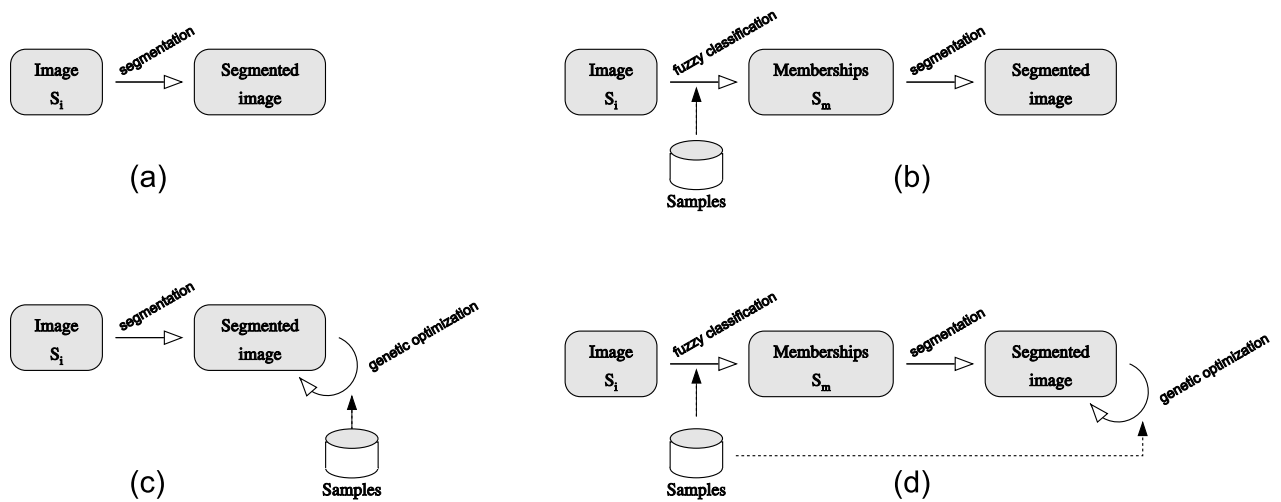


FIG. 2.16 – Approches morphologiques supervisées sans prise en compte de l'information spatiale : (a) segmentation non-supervisée, (b) segmentation par transformation d'espace, (c) optimisation de segmentation, (d) segmentation supervisée hybride.

Avec prise en compte de l'information spatiale

Les exemples d'apprentissage, lorsqu'ils sont fournis par l'utilisateur sous la forme de zones étiquetées, portent également une information spatiale qu'il peut être pertinent d'exploiter. Nous avons ainsi proposé dans [D19, E16] d'assimiler ces zones à des marqueurs, utilisés dans un processus de segmentation par ligne de partage des eaux. Chaque marqueur est dans un premier temps considéré comme une zone d'apprentissage associée à une classe donnée. Le processus de classification supervisée floue s'appuie alors sur l'ensemble des marqueurs étiquetés pour produire les cartes des degrés d'appartenance relatives aux différentes classes. Dans un second temps, les marqueurs permettent l'initialisation spatiale de l'algorithme de ligne de partage des eaux en définissant les positions initiales des bassins de rétention. Cependant, contrairement à la formulation originale de cette méthode, nous utilisons ici un relief différent par classe, et l'extension de chaque bassin (ou marqueur) s'effectuera en considérant le relief associé à sa classe (naturellement les différents reliefs sont inondés de façon synchrone, afin d'éviter de produire plusieurs segmentations au final). Pour construire ce relief dit supervisé, la carte des degrés d'appartenance propre à la classe considérée est fusionnée avec le relief initial, commun à toutes les classes (généralement un gradient calculé sur l'image). L'objectif ici est de limiter l'altitude des pixels de forte similarité avec un bassin lors de l'extension de celui-ci, même si leur altitude initiale était importante (par exemple en cas de gradient élevé lié à la présence d'une texture, d'un bruit, etc.).

Pour un pixel donné p , son altitude dans le relief supervisé lié à la classe c peut, par exemple, être calculé comme $f'(p, c) = f(p) \cdot (1 - \omega(p, c))$, avec f et f' les reliefs initial et supervisé, et ω la carte des degrés d'appartenance (toutes ces images étant normalisées pour permettre leur comparaison). La figure 2.17 illustre ce processus de segmentation supervisée, en supposant que les marqueurs soient fournis par l'utilisateur (notons qu'ils pourraient aussi être extraits automatiquement par analyse de l'image initiale). Comme nous le verrons dans le chapitre suivant, cette méthode a été appliquée avec succès en imagerie couleur et en télédétection (notons par ailleurs que son application en imagerie sous-marine est à l'étude).

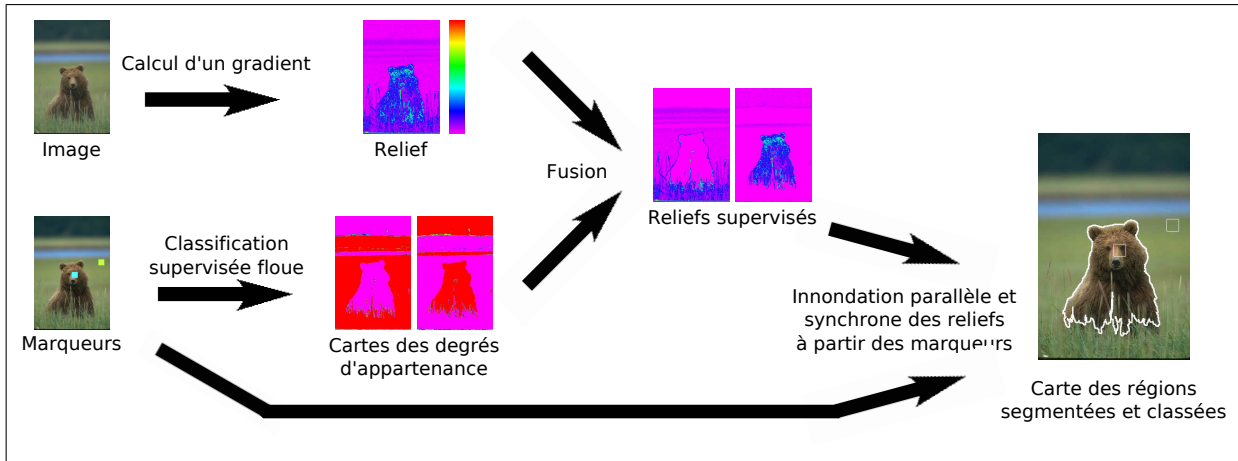


FIG. 2.17 – Processus de segmentation par marqueurs contextuels

Nous pouvons observer la complémentarité des informations utilisées au travers des marqueurs (leur position et leur contenu) dans la figure 2.18, où nous pouvons constater le faible nombre de marqueurs nécessaires pour produire un résultat pertinent. Dans cette figure, notons que le contenu de l'image n'est pas exploité de façon optimale, puisque seules les valeurs brutes des pixels (dans l'espace rouge – vert – bleu) sont utilisées pour calculer le gradient nécessaire à l'algorithme de ligne de partage des eaux et pour produire la classification supervisée (des meilleurs résultats auraient ainsi pu être obtenus si nous avions utilisé, comme dans [MH05], une information de couleur et de texture).

2.3.5 Conclusion

La résolution d'un problème d'analyse ou de traitement d'image, posé dans un contexte applicatif particulier, peut (et doit) bénéficier de connaissances du domaine pour conduire à une solution la plus pertinente possible. En morphologie mathématique, ces connaissances sont mises en œuvre le plus souvent de façon implicite, par les choix des opérateurs morphologiques et éléments structurants effectués lors de la conception de la solution. Il est alors délicat de reproduire les résultats sur d'autres images ou dans d'autres contextes sans faire appel de nouveau au spécialiste de la morphologie mathématique. Au contraire, en explicitant les connaissances utilisées, on favorise la réutilisation des méthodes élaborées et leur adaptation à de nouvelles problématiques, puisqu'il suffit de modifier les connaissances du domaine fournies aux algorithmes morphologiques. Dans cette section, nous avons discuté le problème de l'intégration de connaissances en morphologie mathématique et avons proposé différentes stratégies pour élaborer des approches supervisées ou guidées par l'utilisateur.

Puisque la connaissance est fréquemment introduite dans les solutions morphologiques au moyen des éléments structurants utilisés, nous avons étudié la nature des connaissances pouvant être apportées par les éléments structurants. En plus des informations de taille et de forme, traditionnellement utilisées en morphologie mathématique ensembliste (images binaires) ou plate (images à niveaux de gris), on peut en effet considérer d'autres connaissances telles que le profil d'intensité des pixels, que ce soit dans des images monovaluées [B9], couleur [B7], ou multispectrales [D25]. La prise en compte de plusieurs éléments structurants, soit de façon conjonctive [D25], soit de façon disjonctive [D17], permet d'appréhender des objets ou des connaissances de nature complexe ou imprécise, mais nécessite de savoir fusionner les résultats obtenus en utilisant les différents éléments structurants.

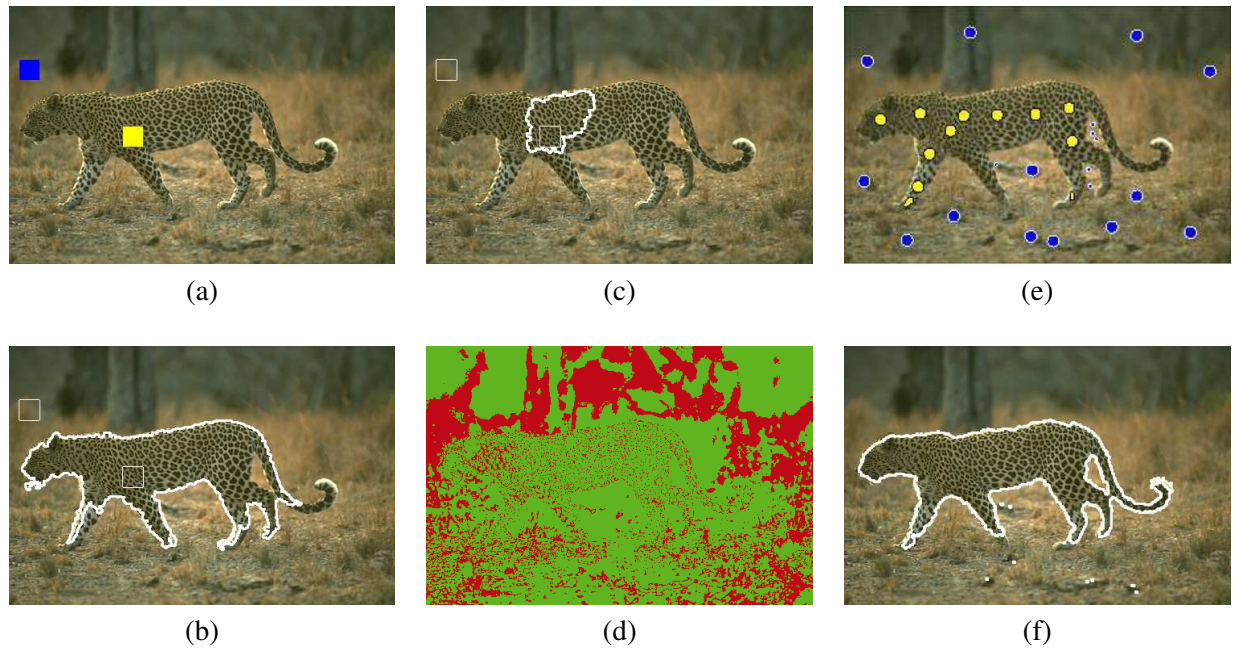


FIG. 2.18 – Segmentation d’une image couleur guidée par des marqueurs contextuels : (a) image initiale avec marqueurs, (b) segmentation avec notre algorithme [D19, E16], (c) segmentation par LPE avec marqueurs, (d) classification supervisée, (e) marqueurs et (f) résultats présentés dans [MH05].

Outre l’usage des éléments structurants, nous nous sommes également intéressés à l’apport des techniques de classification. Nous avons considéré le cas de classifications supervisée ou non-supervisée, dure ou floue. L’utilisation la plus directe de ces techniques consiste en un prétraitement qui a pour but de modifier l’espace de représentation des pixels sur lequel seront basés les algorithmes morphologiques appliqués par la suite (par exemple, segmentation par ligne de partage des eaux, détection d’objet par transformée en tout-ou-rien). La classification dure fournit une image de classes ou un ensemble d’images binaires [D17] tandis que le cadre flou produit une image multivaluée mesurant en chaque bande le degré d’appartenance des pixels à une classe donnée. Cette dernière peut-être obtenue par une approche non-supervisée [E10] ou une approche supervisée [D14] et basée sur les réponses spectrales des pixels ou d’autres attributs tels que la texture [E13].

En complément d’un changement d’espace effectué en prétraitement, la classification supervisée peut également être utile pour aider à déterminer les paramètres optimaux d’un algorithme morphologique. Ainsi, dans le cas de la segmentation par ligne de partage des eaux et des techniques de réduction de sursegmentation qui y sont associées, nous avons proposé d’appliquer une classification supervisée pour évaluer la pertinence de la segmentation (au sens de la qualité de la classification des régions qu’elle produit), et d’optimiser les différents paramètres mis en jeu à l’aide d’un algorithme génétique [E11, D20].

Les zones d’apprentissage, nécessaires à la classification supervisée, peuvent être par ailleurs considérées comme des marqueurs dans un processus de segmentation par ligne de partage des eaux [D19, E16]. Le relief initial sur lequel doit s’appliquer la segmentation est alors remplacé par des reliefs supervisés obtenus en fusionnant celui-ci avec les cartes des degrés d’appartenance aux classes produites par la classification supervisée. En associant chaque marqueur à une classe donnée et au relief correspondant, la segmentation peut exploiter en parallèle et de façon synchrone les différents reliefs pour ne produire qu’une seule partition.

Enfin, nous avons étudié le lien entre la segmentation et classification non-supervisée. Tandis que la première produit une partition en amas connexes de points homogènes appelés régions, la seconde produit une partition en amas d’éléments homogènes appelés classes (ou *clusters*). En appliquant la classification sur l’espace des pixels et en supposant que les classes produites sont connexes, on obtient alors directement une segmentation [E15]. Nous avons proposé une telle formulation en l’illustrant par l’algorithme k -means, sans doute le plus connu des algorithmes de classification non-supervisée. La morphologie mathématique apporte ici une mesure de distance qui permet de garantir la connexité des classes produites, et une implantation rapide de l’algorithme. La méthode obtenue peut tout autant être assimilée à un k -means topographique ou à une segmentation

morphologique itérative.

Les différents travaux que nous avons menés sur l'intégration de connaissances en morphologie mathématique nous amènent à dresser les constats suivants :

- bien que fondamentale, l'intégration de connaissances dans les approches morphologiques (comme dans d'autres approches en analyse et traitement d'image) est confrontée à la difficulté que peuvent rencontrer les experts pour formaliser leurs connaissances du domaine ; il est souvent plus aisé d'introduire ces connaissances sous la forme d'ensembles d'apprentissage dans un processus de classification supervisée que de les spécifier au travers des éléments structurants ;
- la forme que peut prendre l'intégration de connaissances en morphologie mathématique dépend néanmoins du contexte applicatif et du problème d'analyse ou de traitement d'image à résoudre ; ainsi, nous avons étudié les cas de la détection d'objet et de la segmentation en régions, pour lesquels il est plus pertinent de formuler la connaissance respectivement au travers des éléments structurants et d'une classification floue ; cette dernière peut être supervisée ou non, selon la capacité qu'a l'utilisateur à fournir des exemples d'apprentissage ; notons que face à d'autres problèmes, les connaissances introduites dans les approches morphologiques supervisées pourraient être de nature différente à celles discutées dans cette section ;
- nous avons proposé des stratégies générales pour l'élaboration d'approches morphologiques supervisées exploitant un processus de classification (supervisée ou non) ; mais les résultats obtenus dans un contexte applicatif donné dépendent fort logiquement de deux paramètres principaux : l'espace des attributs dans lequel est effectuée la classification, et l'algorithme de classification considéré ; ces deux paramètres, même s'ils ne sont pas de nature morphologique, doivent être étudiés attentivement.

2.3.6 Perspectives

Modélisation avancée des connaissances et nouveaux paradigmes en apprentissage et fouille de données

Les connaissances que nous avons exploitées dans cette section ont pris la forme soit d'éléments structurants, soit d'éléments relatifs à un processus de classification (nombre de classes dans le cas non-supervisé, ensembles d'apprentissage dans le cas supervisé). L'intégration de connaissances en morphologie mathématique pourrait néanmoins être effectuée en s'appuyant sur d'autres modélisations des connaissances ou d'autres paradigmes de classification.

Ainsi, la démarche de formalisation des connaissances de l'expert peut aboutir à la constitution d'une base de connaissances, voire à la création d'une ontologie permettant la définition des concepts d'un domaine. Dans le cadre de la thèse de Sébastien Derivaux et du projet ACI Masse de Données FoDoMuST (2004–2007), nous avons proposé d'exploiter une telle base de connaissances dans le domaine de la télédétection urbaine pour construire de façon automatique des méthodes dédiées à la détection d'objets d'une classe d'intérêt donnée (et formalisée dans la base de connaissance) [E14]. Ces détecteurs sont composés de filtres appliqués de façon séquentielle, chaque filtre étant relatif à un critère donné (information spectrale, spatiale, etc.). La morphologie mathématique n'a ici été utilisée que dans le cas binaire pour définir un filtre de forme (rectangulaire). Ces travaux préliminaires demandent à être poursuivis, en utilisant par exemple des éléments structurants à la capacité de description plus riche (cf. section 2.3.2), permettant une meilleure prise en compte des différentes informations (forme, taille, mais aussi spectre ou texture) et un certain niveau de robustesse, nécessaire pour s'assurer de l'intérêt pratique d'une telle approche. Une étape supplémentaire serait d'exploiter une ontologie pour inférer des méthodes d'analyse d'image aptes à résoudre un problème donné [Ren07]. Dans le contexte de la morphologie mathématique, cette approche consisterait à composer et paramétrer automatiquement (à partir des connaissances disponibles) des opérateurs ou algorithmes morphologiques élémentaires pour obtenir une solution morphologique à un problème donné.

Outre la perspective d'exploiter en morphologie mathématique des connaissances expertes au travers de modèles de représentation plus évolués, il serait également intéressant de réfléchir à l'apport que pourraient fournir différents paradigmes introduits récemment en apprentissage automatique et fouille de données. En effet, les techniques relatives à ces domaines sur lesquelles nous avons basé nos approches morphologiques supervisées se sont pour l'instant limitées à des méthodes de classification (supervisée ou non, dure ou floue)

déjà anciennes. Ainsi, lorsque les attributs décrivant les pixels ou les objets d'une image produisent une représentation creuse des données, il peut être intéressant de faire appel aux techniques basées sur la réduction de données ou l'apprentissage topologique. La classification collaborative s'appuie sur différents classifieurs dont la combinaison permet de produire le résultat final : appliquée à la morphologie mathématique et notamment au problème de segmentation, cette approche pourrait servir à fusionner différents critères dans le processus de segmentation, ou autrement dit à tenir compte de connaissances de différentes natures fournies par l'utilisateur. La classification semi-supervisée limite la connaissance utilisée en tenant essentiellement compte de contraintes relationnelles entre éléments à classer (l'utilisateur peut ainsi spécifier que deux éléments X et Y doivent, ou au contraire ne doivent pas, appartenir à la même classe). Appliquée au principe de segmentation, cette stratégie permettrait de faciliter l'interaction avec l'utilisateur qui guiderait le processus en associant ou en séparant deux éléments de l'image (pixels ou régions). L'apprentissage actif limite pour sa part l'intervention de l'utilisateur en lui demandant d'étiqueter des éléments particuliers (c'est-à-dire d'indiquer leur classe) afin de permettre au processus de produire une classification pertinente en s'appuyant sur un minimum de données d'apprentissage. Là encore, ce paradigme présente un intérêt en traitement d'image (notamment en segmentation), puisqu'il offre un nouveau mode d'interaction avec l'utilisateur, lui permettant de guider le processus. Enfin, notons qu'il pourrait être judicieux d'étudier le lien entre les travaux en morphologie mathématique appliquée sur les graphes d'une part, et les méthodes de classification spectrale d'autre part : ces dernières représentent les données sous forme de graphes, en définissant les sommets comme les éléments à classer (par exemple les pixels, les régions, etc.) et les arêtes comme les similarités entre éléments. La classification est alors obtenue en cherchant la partition optimale (au sens d'un certain critère) du graphe.

Conception itérative d'une solution morphologique

En complément d'une formalisation plus poussée des connaissances, ou de l'utilisation de paradigmes récents en classification (supervisée ou non), nous pensons que la recherche d'une solution morphologique à un problème donné peut être assimilée à un problème d'optimisation. L'utilisation de techniques itératives et robustes d'optimisation semble effectivement un moyen pertinent pour favoriser l'adaptation des opérateurs morphologiques à un objectif particulier tout en limitant (voire en évitant) l'intervention de l'expert.

Ainsi, en collaboration avec Alexandre Blansché et Pierre Gançarski de l'équipe FDBT, nous étudions l'adaptation de la méthode d'optimisation génétique MACLAW [BGK06] au cadre de la morphologie mathématique. Notre objectif est de proposer une méthode de classification morphologique, où chaque classe est associée à un détecteur un-contre-tous. Un détecteur est construit comme une séquence d'opérations morphologiques (paramétrés par des éléments structurants) qui fait l'objet d'une optimisation par une stratégie évolutionnaire.

Dans le cadre de la thèse de doctorat de Jonathan Weber, toujours menée en collaboration avec Pierre Gançarski de l'équipe FDBT, nous nous intéressons à la remise en cause, à l'aide de techniques de fouille de données multistratégies, de segmentations spatio-temporelles de séquences vidéo et de descriptions d'objets extraits de ces séquences. L'implication de l'expert dans le processus, au moyen d'une évaluation des résultats fournis par le système d'analyse vidéo, semble être une solution efficace pour faciliter l'introduction d'informations sémantiques dans la classification de séquences vidéo en fonction des objets qu'elles contiennent.

Enfin, dans le contexte de la modélisation bayésienne hiérarchique de structures astronomiques, la morphologie mathématique peut être utilisée pour construire des filtres dont le but est de supprimer les structures qui ne doivent pas apparaître dans les modèles. La méthodologie choisie, réalisée dans le cadre de la thèse de doctorat de Benjamin Perret en collaboration avec des spécialistes de la modélisation bayésienne et des approches statistiques (Vincent Mazet et Christophe Collet, équipe MIV), s'appuie sur des algorithmes de Monte Carlo par chaînes de Markov, où les paramètres des filtres morphologiques sont des éléments du modèle à estimer.

Morphologie mathématique par l'exemple

L'intégration de connaissances en morphologie mathématique pourrait également prendre la forme, pour un problème d'analyse d'image donné, de résultats de référence fournis par l'utilisateur. Autrement dit, étant donné un (ou plusieurs) exemple de résultat, est-il possible de définir automatiquement le traitement morphologique

permettant d'y aboutir ? Les résultats obtenus sont-ils alors reproductibles (assurant ainsi une robustesse à la méthodologie) ?

Les solutions mises en œuvre pour répondre à cette question dépendent bien naturellement du problème étudié. Si l'objectif est d'effectuer la détection d'objet (processus intrinsèquement supervisé), l'identification par l'utilisateur des objets devant être détectés dans une image fournit non seulement des motifs à l'opérateur de reconnaissance de formes (connaissances endogènes) mais également des informations contextuelles (connaissances exogènes) obtenues par analyse du reste de l'image annotée par l'utilisateur. La segmentation d'image, même si elle suscite toujours de nombreux travaux dans un cadre non-supervisé, a récemment été étudiée comme un problème d'apprentissage, notamment suite aux contributions de l'équipe de Berkeley qui a fourni une base d'images segmentées manuellement [MFTM01] et pouvant être utilisées comme ensemble d'apprentissage. Dans ce contexte, l'utilisation d'images exemples pour construire un gradient supervisé [MFM04] peut bénéficier aux approches morphologiques telles que la technique de ligne de partage des eaux [HM09]. Outre la détection et la segmentation, d'autres problèmes pour lesquels la morphologie mathématique offre des solutions (détection de contours, description de texture, etc.) pourraient également être abordés selon cette stratégie "par l'exemple". Notons que la détermination de la solution morphologique pourrait être assimilée à un problème d'optimisation et résolu par les méthodes que nous avons discutées dans la perspective précédente, et que l'apprentissage de solutions morphologiques a déjà fait l'objet de quelques études [TZR95,DB02,HHM03] (ces classifieurs morphologiques étant le plus souvent dédiés aux images binaires) qu'il nous semble pertinent d'approfondir.

Chapitre 3

Applications

Dans ce chapitre, nous présentons les différents cadres applicatifs dans lesquels nos travaux se sont inscrits : analyse d'image couleur pour l'annotation et la recherche par le contenu (section 3.1), détection et segmentation d'objets géographiques en télédétection (section 3.2), et détection, segmentation et classification de galaxies en imagerie astronomique (section 3.3). Ce chapitre contient également une description de la plate-forme PELI-CAN (section 3.4) que nous utilisons au quotidien pour résoudre les problèmes d'analyse d'image qui nous sont posés.

Notons que nous avons fait le choix de ne pas présenter d'évaluation quantitative des résultats expérimentaux que nous avons obtenus (ils auraient nécessité la présence de nombreux tableaux de mesures statistiques, et nous renvoyons le lecteur aux publications issues de ces travaux). La présentation se veut ici plus pédagogique, et nous avons préféré inclure des exemples visuels de résultats pour illustrer le comportement et l'intérêt de nos approches.

3.1 Analyse d'images couleur

Dans cette première section applicative, nous nous intéressons à l'analyse des images couleur. Les travaux que nous présentons ici ont été en grande partie réalisés avec Erchan Aptoula dans le cadre de sa thèse de doctorat [Apt08], en application des travaux fondamentaux sur les approches morphologiques multivaluées (vectorielles) décrites dans la section 2.2. Dans le cadre d'un projet soutenu par Oséo, notre objectif principal était de concevoir un système d'annotation et de recherche par le contenu dans des bases d'images couleur. Nous nous sommes donc intéressés aux différents éléments constitutifs d'un tel système qui pouvaient être élaborés à l'aide d'approches morphologiques. Nos principales réalisations ont porté sur la détection d'objets d'intérêt, la segmentation en régions, et la description du contenu des images. Elles ont donné lieu à plusieurs articles parus dans des revues internationales [B8, B7, B6, B10], des chapitres d'ouvrage [A2, A3], et à des communications en conférence internationale [D16, D18, D19, D22, D23] ou nationale [E16].

Nous débuterons cette section en rappelant les spécificités des images couleur, qui nécessitent un espace de représentation particulier. Nous discuterons donc les modalités d'application de la morphologie mathématique aux images couleur, en particulier dans le cas où les images sont représentées dans un espace teinte – saturation – luminance. Puis nous décrirons les solutions morphologiques que nous avons apportées aux différents problèmes auxquels nous nous sommes intéressés (détection d'objet et segmentation en régions, description du contenu des images). Notre système d'annotation et de recherche par le contenu sera présenté, avant de conclure sur les travaux que nous avons menés en imagerie couleur et de dresser quelques perspectives.

3.1.1 Introduction

L'homme voit la vie en couleur. Il est donc fort logique que, dès l'arrivée des capteurs 3-CCD, les travaux menés en analyse et traitement des images couleur aient été très nombreux. Le traitement de telles images nécessite d'adapter les techniques élaborées pour les images à niveaux de gris, voire de développer de nouvelles approches dédiées à ces images, mais aussi de choisir un espace de représentation adapté [GBL04, TFMB04].

L'adaptation des techniques morphologiques aux images multivaluées telles que les images couleur a fait l'objet de la section 2.2. Le choix d'un espace de représentation est pour sa part plus spécifique au cas des images couleur et, face à la grande variété des alternatives introduites dans la littérature, il doit être effectué avec attention.

Espaces couleur

L'espace rouge – vert – bleu, celui utilisé pour afficher à l'écran les images numériques, est souvent utilisé puisqu'il est accessible sans transformation. Néanmoins il présente certains inconvénients, notamment : les couleurs sont difficilement exprimables en termes de valeurs dans cet espace (excepté pour les couleurs les plus saturées où la valeur dans chaque composante de l'espace est soit la valeur minimale, soit la valeur maximale) ; les distances euclidiennes entre les couleurs ne correspondent pas aux différences visuellement perçues par l'homme ; enfin, il existe une très forte corrélation entre les bandes. Cet espace n'est donc pas le plus adapté pour traiter des images couleur (même si en réalité la littérature comprend de nombreux exemples d'une telle utilisation).

Les espaces perceptuels (ou perceptuellement uniformes) tels que $L^*a^*b^*$ représentent les similarités entre les couleurs de façon plus fidèle à la vision humaine. Ils séparent la composante de luminance des composantes chromatiques, ces dernières étant représentées sur des axes opposant deux couleurs (par exemple, a^* est négatif pour des couleurs proches du vert, positif pour des couleurs proches du rouge, tandis que b^* est négatif pour des couleurs proches du bleu et positif pour des couleurs proches du jaune). Bien qu'ils offrent ces deux propriétés intéressantes, ces espaces nécessitent néanmoins de connaître la composition du blanc de référence, et les formules de calcul depuis des valeurs définies dans l'espace rouge – vert – bleu sont calculatoirement complexes.

Les espaces couleurs polaires, de type teinte – saturation – luminance, différencient l'intensité d'une couleur, sa pureté, et sa longueur d'onde principale. Ils offrent l'intérêt de représenter plus fidèlement le système de la vision humaine, mais ne garantissent pas d'uniformité perceptuelle, et la teinte définie comme une valeur angulaire doit être traitée avec attention (nous traiterons plus en détail ce problème par la suite). De surcroît, il a été montré récemment que les définitions les plus utilisées de ces espaces présentaient des inconsistances importantes [Han02, Han08] liées à leur formation cylindrique ou à la dépendance imposée entre luminance et saturation. Il en résulte la proposition d'un espace amélioré de forme bi-conique et noté LSH [AS07] qui assure l'indépendance de la saturation vis-à-vis de la luminance, l'existence de normes pour les notions de luminance et de saturation, et l'existence d'une transformation inverse permettant de revenir dans l'espace initial rouge – vert – bleu. Par la suite, nous avons utilisé le plus souvent cet espace pour nos applications de la morphologie mathématique aux images couleur, et nous nous sommes en particulier intéressés à l'importance accordée aux différentes composantes dans cet espace [B8].

Morphologie mathématique et images couleur

L'application d'un ordre vectoriel (par exemple un ordre total équilibré – cf. section 2.2.3) sur un espace couleur donné permet de traiter les images couleur au moyen de la morphologie mathématique. Néanmoins, si cet espace de représentation est polaire (de type teinte – saturation – luminance), et que la teinte est utilisée dans l'ordonnement des couleurs des pixels, il faut disposer d'une relation d'ordre entre valeurs angulaires (ici les teintes).

Ainsi, quelques solutions ont été proposées dans la littérature, et la plus pertinente est sans doute celle proposée par Hanbury et Serra [HS01] qui consiste à choisir une teinte de référence et à ordonner les teintes par rapport à leur distance angulaire à la référence. Cette méthode n'est cependant pas adaptée dans le cas où l'image est caractérisée par plusieurs teintes principales (ce qui est souvent le cas en réalité) et nous avons donc suggéré d'utiliser plusieurs teintes de référence [B8], voire de pondérer dans la relation d'ordre l'importance donnée à chaque teinte de référence en fonction de sa représentativité dans l'image. La figure 3.1 illustre l'intérêt d'utiliser plusieurs teintes de référence au lieu d'une seule, pour des opérations simples telles qu'érosion et dilatation.

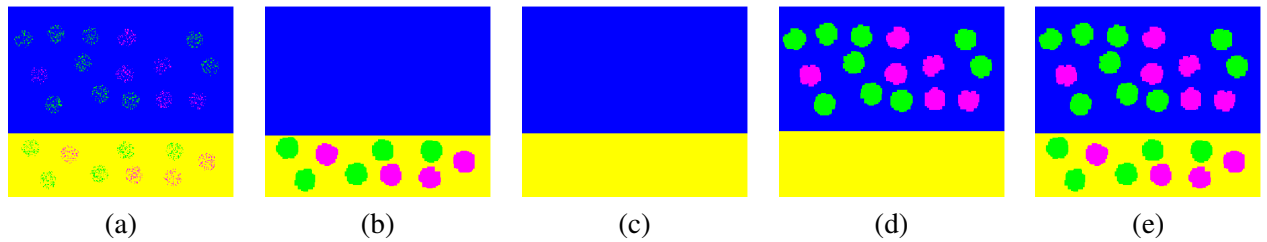


FIG. 3.1 – Traitement de la teinte en morphologie mathématique : image initiale (a), dilatation avec une seule teinte (b) ou deux teintes de référence (c), et érosion avec une seule teinte (d) ou deux teintes (e) de référence. Les teintes de référence sont automatiquement calculées comme étant les plus représentatives dans l'image initiale (a).

3.1.2 Détection d'objet et segmentation en régions

Comme nous pourrions le constater dans les sections suivantes relatives à des applications de la morphologie mathématique dans d'autres domaines, la détection d'objet et la segmentation en régions sont deux approches alternatives permettant l'interprétation du contenu des images. Tandis que la première se focalise sur un objet donné et qu'elle consiste en la recherche des occurrences de cet objet dans les images, la seconde procède par simplification de l'image pour produire un ensemble de régions représenté sous forme de carte de segmentation (on suppose alors que chaque objet d'intérêt est composé d'une ou plusieurs régions, la carte de segmentation contenant généralement aussi des zones sans intérêt pour l'utilisateur). Ces deux stratégies sont discutées dans la suite de cette section, où nous décrirons les solutions mises en œuvre dans le contexte de l'imagerie couleur.

Détection

Nous nous sommes intéressés à la détection d'objets prédéfinis dans les images couleur à l'aide de méthodes morphologiques, en proposant d'appliquer sur de telles images l'opérateur de recherche de motifs connu sous le nom de transformée en tout-ou-rien (cf. section 2.1.4). Pour cela, nous avons étendu les principales définitions introduites pour les images à niveaux de gris [NPR07a] au cas multivalué et proposé une version vectorielle de cet opérateur [B7] (la spécificité par rapport aux autres opérateurs morphologiques vectoriels discutés dans la section 2.2 est qu'ici l'ordre vectoriel utilisé doit assurer la propriété d'invariance à la translation dans l'espace $\mathcal{T}$).

L'application de cet opérateur aux images couleur pour y détecter des objets nécessite cependant quelques aménagements, puisque les images ne contiennent généralement pas un motif exactement identique à celui recherché. Nous avons ainsi adapté au cas des images couleur les techniques de filtre de rang et de fonctions structurantes artificielles introduites dans la littérature pour accroître la robustesse de la transformée en tout-ou-rien dans les images binaires ou à niveaux de gris.

Néanmoins, même avec ces adaptations, l'utilisation de la transformée en tout-ou-rien ne présente un intérêt que si l'objet d'intérêt peut être représenté par un ou plusieurs motifs caractéristiques. En outre, nous n'avons ici pas tenu compte des transformations pouvant être appliqués à ces motifs (rotation ou changement d'échelle par exemple, puisque la translation est intrinsèquement prise en compte par les opérateurs morphologiques). Dans le cas de concepts plus généraux, cette approche ne pourra pas être utilisée et les techniques de segmentation (en préalable à une description et reconnaissance des régions obtenues) pourraient être une alternative plus intéressante.

Segmentation

Nous avons brièvement décrit dans la section 2.1.4 la technique dite de ligne de partage des eaux qui est certainement la méthode de segmentation morphologique la plus connue. Elle considère l'image à segmenter comme un relief, et identifie les frontières des objets comme les lignes de partage des eaux. Il faut donc s'assurer que dans l'image fournie à l'algorithme, les bassins de rétention correspondront à l'intérieur des régions et les

crêtes aux frontières entre les objets. C'est pourquoi cette technique est le plus souvent appliquée non pas sur l'image à segmenter mais sur son gradient qui permet de rehausser les contours entre les régions. La complexité intrinsèque des images est néanmoins à l'origine du principal défaut de cette technique, qui produit le plus souvent une sursegmentation (contenant des frontières artificielles et un trop grand nombre de régions). Pour pallier à ce problème, les deux principales approches consistent à introduire des techniques de réduction de sursegmentation ou de fusion de régions, ou à utiliser des marqueurs définis automatiquement ou saisis par l'utilisateur.

Dans le cas des images couleur, de nombreuses méthodes de segmentation utilisant la ligne de partage des eaux ont été proposées (par exemple [LC02, VPS03, AS07]). Nous nous sommes également intéressés à ce problème et avons proposé deux solutions inspirées de nos travaux fondamentaux présentés dans les sections 2.2 et 2.3.

La première de ces approches n'implique pas d'intervention de l'utilisateur. Pour éviter le problème de sursegmentation, nous faisons ici appel à différentes techniques qui sont appliquées en amont et/ou en aval de l'algorithme LPE. Plus précisément, l'image est filtrée au moyen d'un nivellement (cf. page 43). Le relief est alors construit en recherchant en chaque pixel le maximum des gradients calculés de façon marginale sur chaque composante (en tenant compte de pondérations, puisque par exemple la teinte n'est pas significative si la saturation est faible). Une première segmentation est obtenue en identifiant les lignes de partage des eaux sur ce relief, permettant ainsi de produire un ensemble de régions ou bassins de rétention, représenté sous la forme d'un graphe d'adjacence de régions où chaque noeud du graphe représente un bassin. Nous opérons alors une fusion de bassins en appliquant sur ce graphe des opérations morphologiques qui tiennent compte des attributs des noeuds (couleur moyenne, texture) et des arêtes (contours). Finalement, les noeuds restants sont utilisés comme des marqueurs et une nouvelle segmentation permet d'affiner les contours des régions. Cette approche fournit des résultats intéressants mais nécessite de nombreuses étapes (cf. figure 3.2) pour éviter le phénomène de sursegmentation [D23].

La seconde approche limite pour sa part le phénomène de sursegmentation en exploitant des marqueurs fournis par l'utilisateur. Cette approche a déjà été décrite dans la section précédente (cf. section 2.3.4) et son application à des images couleur nécessite uniquement d'effectuer la classification supervisée floue dans un espace adapté (de préférence un espace teinte – saturation – luminance plutôt que l'espace rouge – vert – bleu) [D19, E16]. Ainsi on évite l'introduction de nombreuses étapes telles que des fusions de région, puisque le nombre de régions correspondra au nombre de marqueurs saisis par l'utilisateur (il pourra même être inférieur si on choisit de fusionner les régions adjacentes produites à partir de marqueurs d'une même classe). La pertinence d'une telle approche est illustrée en figure 3.3.

3.1.3 Description d'image

La description d'une image couleur est une étape importante dans le processus d'analyse ou d'interprétation d'image. En effet, une méthode de classification ou de reconnaissance de formes, quelle qu'elle soit, ne sera d'aucune utilité si les descripteurs utilisés en amont pour caractériser l'image ne sont pas pertinents. Dans la section introductive à la morphologie mathématique (et plus précisément en section 2.1.5), nous avons rappelé les principaux descripteurs morphologiques permettant la caractérisation globale d'une image selon une représentation dans un espace-échelle [A2] : granulométrie, spectre morphologique, covariance. Ces descripteurs ont été initialement définis pour des images binaires et à niveaux de gris, et offrent une alternative intéressante aux descripteurs relatifs à la distribution des intensités des pixels (au travers d'histogrammes par exemple) plus fréquemment utilisés pour décrire le contenu d'une image (notamment dans un but d'indexation et recherche par le contenu). Nous avons étudié leur adaptation aux images couleur, et plus généralement l'intérêt des approches morphologiques dans le contexte de la description et la reconnaissance des images couleur.

L'adaptation la plus directe consiste à calculer ces mesures morphologiques de façon marginale ou vectorielle sur l'image couleur, résultant ainsi en une signature multidimensionnelle [D18]. Mais la couleur offre d'autres possibilités et nous avons ainsi proposé une mesure granulométrique spécifique à certaines couleurs [B10]. Ces couleurs peuvent être par exemple obtenues après une étape de quantification, quel que soit l'espace couleur considéré, et utilisées pour construire autant d'images de distance où chaque pixel se voit attribuer sa

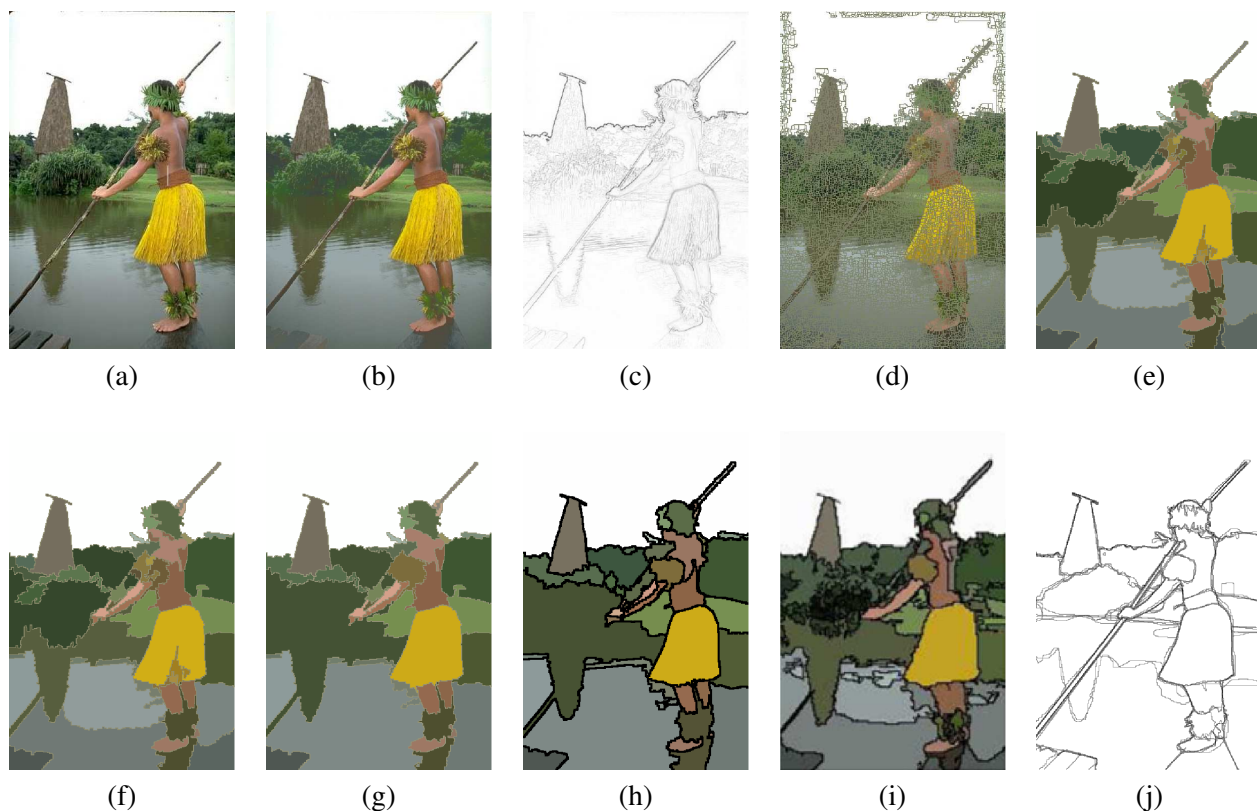


FIG. 3.2 – Segmentation non-supervisée d’images couleur : (a) image initiale, (b) simplification, (c) gradient, (d) sursegmentation, (e)-(g) fusions successives, et (h) résultat final, comparé à (i) la méthode d’Angulo et Serra [AS07] et (j) des segmentations manuelles [MFTM01].

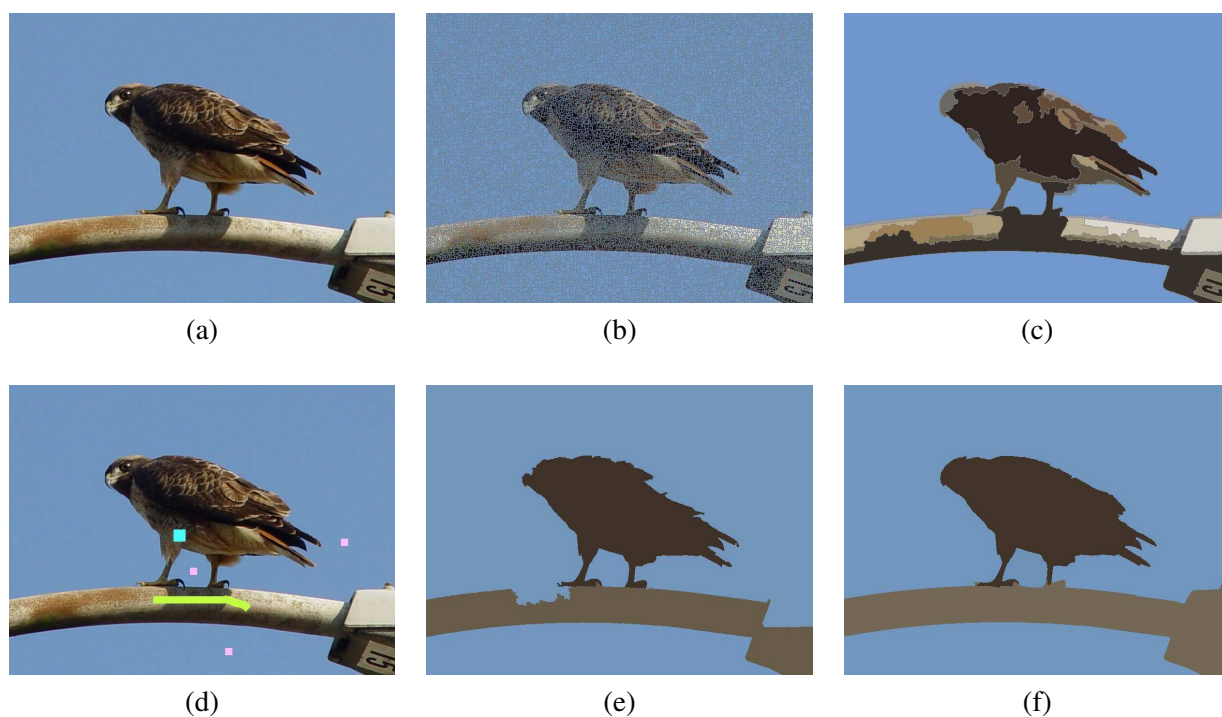


FIG. 3.3 – Segmentation guidée par des marqueurs spatiaux et colorimétriques : (a) image initiale, (b) approche classique de ligne de partage des eaux (sursegmentation), (c) segmentation non-supervisée [D23], (d) marqueurs (et classes correspondantes) fournis par l’utilisateur, (e) approche classique de ligne de partage des eaux avec marqueurs (purement spatiaux), (f) résultat obtenu par l’approche proposée dans [D19, E16].

distance (dans l'espace couleur choisi) avec la couleur en question. Les granulométries sont finalement calculées sur ces images de distance, comme illustré dans la figure 3.4 pour l'espace rouge – vert – bleu.

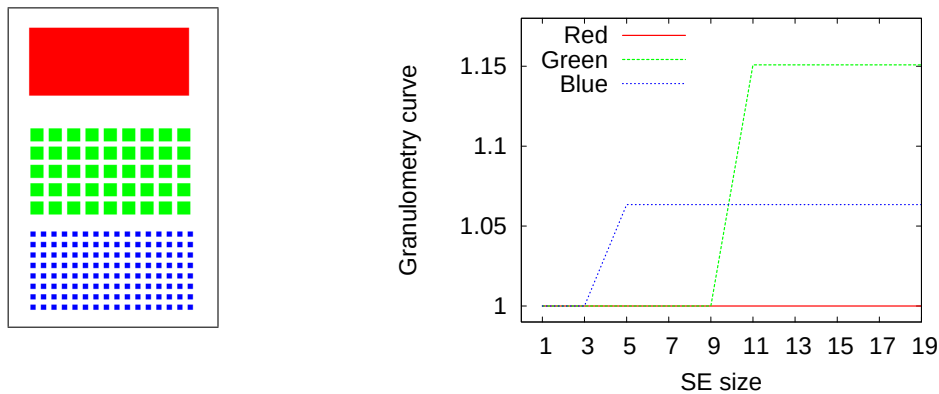


FIG. 3.4 – Image couleur (gauche) et courbes granulométriques spécifiques (aux couleurs rouge, vert, et bleu) associées (droite).

Puisque les histogrammes de distribution des intensités (ou des couleurs) sont un des outils les plus fréquemment utilisés pour caractériser une image, nous avons également étudié quel pouvait être l'apport des opérateurs morphologiques à ces outils. Ainsi, nous avons proposé de calculer des histogrammes multirésolutions, non pas sur des espaces-échelle Gaussiens [HGN04] comme c'est généralement le cas, mais sur des espaces-échelle morphologiques construits à l'aide de nivellements ou de segmentations hiérarchiques obtenues par ligne de partage des eaux [B10]. Une autre approche consiste à combiner les distributions d'intensité (mesurées avec les histogrammes) et les distributions de taille (mesurées avec les granulométries). Cette démarche est également valable pour les images à niveaux de gris et fournit par exemple le résultat donné en figure 3.5

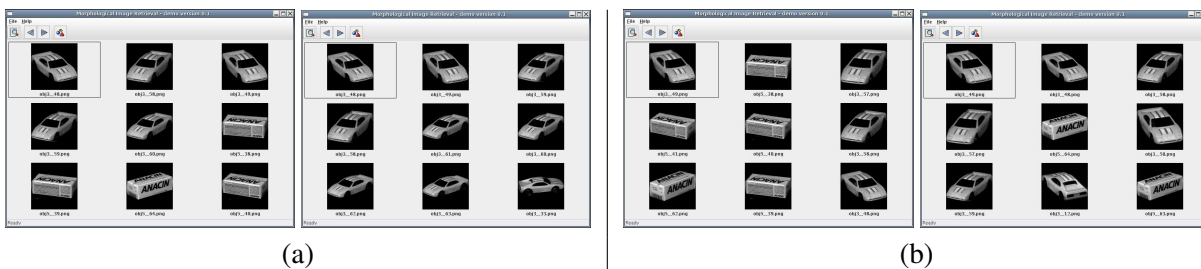


FIG. 3.5 – Deux recherches (a) et (b) dans la base COIL-20 de l'Université de Columbia, avec l'image requête (encadrée) et les résultats obtenus en utilisant des histogrammes seuls (gauche) ou couplés à des granulométries (droite).

Coupler l'information portée par une granulométrie avec un autre type d'information (relatif à la forme, l'orientation, la position, la couleur, la distribution des intensités, etc.) est une piste intéressante pour produire des descripteurs plus pertinents. Ainsi, nous avons étudié l'intérêt de ces descripteurs morphologiques multidimensionnels [B6] et présentons dans la figure 3.6 quelques exemples représentatifs. Nos contributions ont notamment porté sur des granulométries bidimensionnelles à base de rectangles [D17], des covariances tenant compte de l'information spatiale à l'aide de moments [D16], ou encore la combinaison des principes de covariance et de granulométrie au sein d'un unique descripteur [D22].

Les différents descripteurs que nous venons de présenter peuvent être utilisés pour caractériser le contenu des images, une tâche essentielle dans tout système d'indexation et de recherche par le contenu. Nous décrivons maintenant notre approche pour concevoir un tel système.

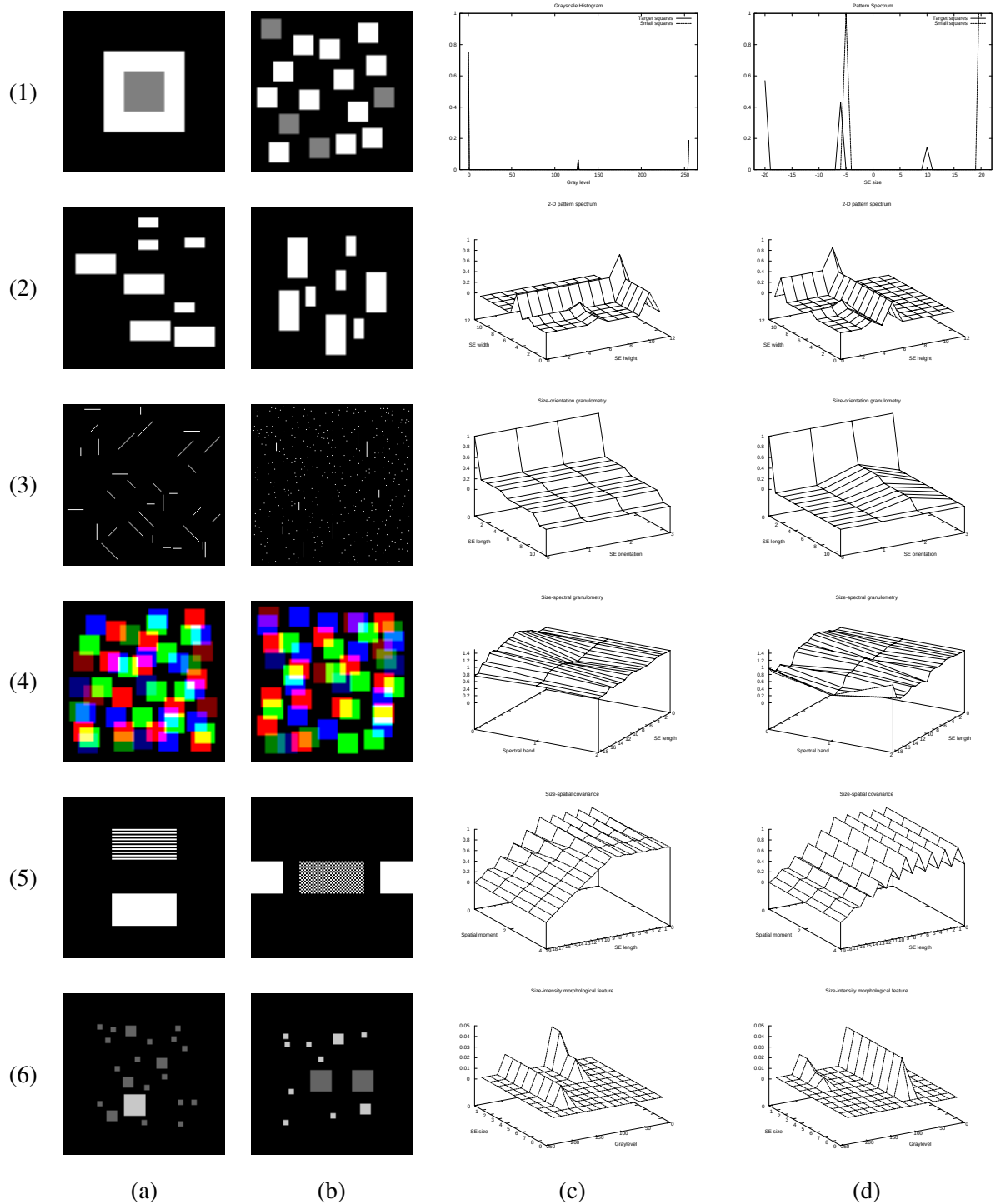


FIG. 3.6 – Intérêt des descripteurs morphologiques multidimensionnels. Sur chaque ligne, deux images (a) et (b) aux descriptions identiques selon leurs histogrammes (1), granulométries à l'aide d'un carré (2), granulométries à l'aide de lignes horizontales (3), granulométries multivaluées marginales (4), covariances (5), et à la fois leurs histogrammes et leurs granulométries (6). Ces images ont cependant un contenu différent si on les décrit au moyen de leur granulométrie (1) ou si on combine cette mesure morphologique avec une information de forme (2), une information directionnelle (3), une approche vectorielle (4), des moments spatiaux (5), ou encore un histogramme (6).

3.1.4 Annotation semi-automatique et recherche d'image couleur

À l'ère des technologies de l'information et de la communication, les informations accessibles sous forme numérique sont de différentes natures : texte le plus souvent, mais également images, sons, vidéos, etc. Les images (et par extension les séquences d'images ou séquences vidéo) sont actuellement sous-exploitées du fait de l'absence de méthodes de recherche et d'exploration appropriées. En effet, la solution mise en oeuvre historiquement par les systèmes d'information ou les moteurs de recherche, qui consiste en une recherche par mots-clé, nécessite l'annotation manuelle de tout le corpus (soit des millions voire des milliards d'images en se limitant aux données disponibles sur le réseau Internet). Elle ne peut donc être appliquée de manière pertinente à grande échelle.

La solution alternative préconisée par la communauté scientifique [Gro07, DJLW08] et suivie par quelques systèmes commerciaux, connue sous le nom de recherche par le contenu (CBIR : Content-Based Image Retrieval), s'appuie sur une définition de la part de l'utilisateur du contenu à chercher, et ce de deux façons possibles : soit en précisant les caractéristiques de l'image recherchée (l'exemple le plus courant est la couleur principale de l'image recherchée), soit en fournissant un ou des exemples (voire également des contre-exemples) de l'image recherchée. Le résultat consiste alors en une liste d'images de pertinence décroissante vis-à-vis de la requête de l'utilisateur. Ce type de système, même s'il a l'avantage sur une solution manuelle d'annotation de permettre une prise en compte d'un nombre important d'images, présente cependant certains inconvénients, les deux principaux étant les suivants : d'une part, il est nécessaire que l'utilisateur fournisse un exemple de l'image qu'il recherche (dans le cas de recherche par l'exemple) et donc qu'il dispose d'une telle image ; d'autre part, les résultats sont retournés de manière complètement automatique (la requête étant affinée itérativement) et contiennent fréquemment des erreurs (généralement des faux positifs, c'est-à-dire des images ne correspondant pas au résultat attendu par l'utilisateur). De nombreux travaux sont donc menés actuellement pour tenter d'élaborer des systèmes plus performants et les efforts sont notamment menés sur la prise en compte de l'utilisateur dans le mécanisme de recherche [Fer05], la combinaison d'attributs locaux et globaux dans la description du contenu [Jol05], l'optimisation du processus de recherche de descripteurs dans des grandes bases de données [Ber04]. Cependant, il semble toujours aussi délicat de combler le *semantic gap*, c'est-à-dire la différence entre les descriptions numériques des images et leur signification sémantique.

Dans le cadre d'un projet soutenu par Oséo, nous avons choisi d'adopter une solution par annotation préalable du corpus, permettant ainsi de garantir une fiabilité élevée et une plus grande facilité d'utilisation. Il était toutefois nécessaire d'automatiser le processus d'annotation pour pouvoir considérer un grand nombre d'images. Ce faisant, le système bénéficie de tous les avantages d'un système par mots-clé tout en s'affranchissant de ses principales limites. Pour aider à l'automatisation du processus d'annotation, il est possible de considérer un ensemble restreint d'exemples et de contre-exemples qui permettent d'illustrer ce qui représente (et ne représente pas) un mot-clé donné. L'objectif consiste alors à définir des descripteurs d'image pertinents et discriminants, communs aux images exemples et absents des images contre-exemples. Une fois ces descripteurs élaborés pour un ensemble d'images (exemples et contre-exemples), il est possible de parcourir le corpus complet à la recherche d'images répondant à la description élaborée précédemment. Les images retenues peuvent alors être proposées à l'utilisateur pour validation de l'annotation définie par le système. Finalement, la base d'images peut être interrogée de manière traditionnelle à l'aide des mots-clé prédéfinis, l'ajout d'un mot-clé à la liste ne nécessitant que la fourniture au système d'un ensemble d'exemples et de contre-exemples puis la validation des réponses proposées par le système. La plus grande difficulté dans le système proposé réside dans l'élaboration de descripteurs robustes d'images couleur. Ces descripteurs peuvent être génériques et définis a priori, ou spécifiques à chaque problème considéré (ici le mot-clé en cours d'annotation) et choisis (voire construits, mais il s'agit là d'une perspective) en se basant sur les images fournies (exemples et contre-exemples).

La figure 3.7 présente un schéma du système MIMAR (Morphological Image Annotation and Retrieval) élaboré lors de ce projet avec l'aide d'Erchan Aptoula puis de Régis Witz (ingénieur recruté quelques mois sur contrat). Il se compose d'environ 85 classes Java pour 23000 lignes de code (sans compter les méthodes de traitement d'image, intégrées à la plate-forme PELICAN, qui fera l'objet de la section 3.4), et présente la caractéristique d'être facilement extensible : en effet, bien qu'il s'appuie sur les différentes approches morphologiques décrites dans cette section, ce système n'est en rien limité à ces approches et utilise également d'autres

techniques pour la segmentation ou la description des images. L'ajout de nouvelles techniques se fait au travers d'un fichier de configuration, sans modifier le code source existant.

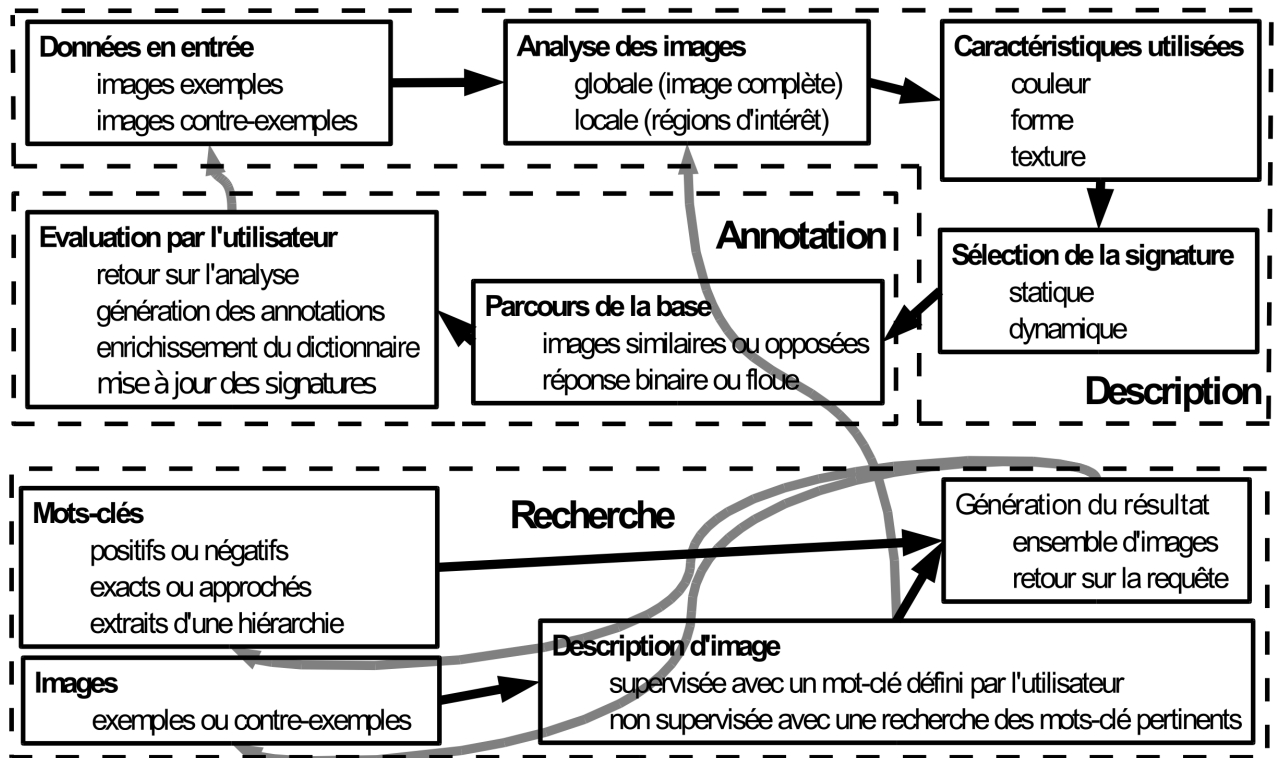


FIG. 3.7 – Schéma du système conçu dans le cadre d'un projet soutenu par Oséo, et ayant pour objectif l'annotation et la recherche d'images couleur par le contenu.

Notre objectif était principalement de fournir une preuve de concept, et le système développé nécessiterait quelques améliorations pour être industrialisable. Néanmoins, il offre la plupart des fonctionnalités présentes dans tout système de recherche par le contenu (composé généralement d'un mécanisme d'extraction de caractéristiques, d'un outil de visualisation, d'un outil graphique de requête, d'une structure d'index, d'un moteur de recherche, et de capacités d'interactivité). En outre, il propose des solutions originales à certains problèmes encore ouverts dans le domaine : il peut s'adapter à tout type de base d'image (généraliste ou spécifique) en garantissant un bon niveau de performances ; il permet l'intégration d'informations sémantiques via l'implication de l'utilisateur dans la définition des mots-clé, la suggestion d'exemples et contre-exemples, et l'analyse des résultats obtenus ; il assure un temps de réponse très faible lors des requêtes, l'annotation étant préalablement effectuée (avec un coût calculatoire bien plus important) ; enfin, il est capable d'évaluer l'intérêt et la pertinence des différents descripteurs d'images qui sont impliqués dans le processus d'annotation.

3.1.5 Conclusion

Dans cette section, nous nous sommes intéressés à l'application de la morphologie mathématique à des images couleur et avons présenté nos travaux dans ce domaine. Les approches mises en œuvre sont directement issues de nos travaux fondamentaux sur les images multivaluées (section 2.2), et exploitent pour certaines nos travaux sur l'intégration de connaissance (section 2.3). Une partie importante d'entre elles a été le fruit du travail de thèse d'Erchan Aptoula.

En premier lieu, il est nécessaire de choisir un espace dans lequel les images couleur vont être traitées et analysées (cette question n'est bien sûr pas spécifique à l'utilisation d'approches morphologiques). Il n'existe pas de position unanime quant à l'espace couleur à utiliser (parmi rouge – vert – bleu, espace perceptuel comme $L^*a^*b^*$, espace polaire de type teinte – saturation – luminance, etc.). Notre choix s'est porté sur un espace polaire récemment défini [AS07, Han08] et nous avons étudié les problèmes relatifs à la définition d'opérateurs

morphologiques dans un tel espace [B8], dont la spécificité principale est apportée par l'information de teinte. Plus précisément, nous avons introduit un nouveau modèle de pondération de la teinte permettant de mieux refléter la pertinence de celle-ci en fonction de niveaux variables de saturation et de luminance. Nous avons également traité du problème de l'ordonnancement de valeurs angulaires : après avoir montré les limites des approches existantes basées sur une unique teinte de référence, nous avons suggéré l'utilisation de multiples teintes de référence, assurant ainsi un meilleur comportement des opérateurs morphologiques.

Nous nous sommes ensuite intéressés à différents problèmes en analyse et traitement des images couleur. Nous avons en particulier étudié le cas de la détection d'objets, assimilée à de la reconnaissance de motifs. Pour cela, nous avons proposé une solution [B7] basée sur la transformée en tout-ou-rien qui a nécessité d'étendre aux images multivaluées cet opérateur ainsi que différentes stratégies lui permettant d'accroître sa robustesse.

Nous avons abordé le problème de la segmentation en proposant deux méthodes différentes utilisant la technique de ligne de partage des eaux. La première est totalement non-supervisée [D23], elle nécessite donc d'effectuer des réductions de sursegmentation ou fusion de régions pour compenser l'absence de l'utilisateur. Les résultats obtenus, quoiqu'intéressants, sont limités par le caractère non-supervisé de l'approche. La seconde approche [D19, E16], au contraire, est supervisée et elle a été précédemment présentée dans un cadre général (section 2.3) : elle s'affranchit d'une possible sursegmentation par l'utilisation de marqueurs fournis par l'utilisateur. En exploitant le contenu couleur des marqueurs, elle offre des résultats d'une grande pertinence.

Outre les problèmes de détection et de segmentation, nous avons également étudié les solutions apportées par la morphologie mathématique pour décrire le contenu des images couleur. Ainsi, il est possible de construire des espaces-échelle morphologiques [A2] à partir desquels pourront être calculés des descripteurs classiques tels que les histogrammes multi-échelles [B10]. Nous avons aussi proposé des mesures granulométriques spécifiques aux images couleur [B10], ou plus générales combinant l'information de distribution de taille portée par la granulométrie avec d'autres informations : la distribution des niveaux de gris [D18], une information de forme [D17], une information spatiale [D16], ou encore une autre mesure morphologique comme la covariance [D22]. Ces descripteurs multidimensionnels offrent un pouvoir de discrimination plus important que leurs versions initiales monodimensionnelles [B6].

Les différents éléments morphologiques que nous avons introduits pour le traitement des images couleur, ont finalement été utilisés au sein d'un système d'annotation et de recherche d'images par le contenu. Ce système, élaboré dans le cadre d'un projet soutenu par Oséo, propose de combiner l'utilisation de mots-clé et d'images pour offrir une bonne capacité d'adaptation aux bases d'images considérées et favoriser son usage par les détenteurs de fonds d'images et leurs clients. La présence d'une phase d'annotation réalisée hors ligne permet d'assurer un temps de réponse minimal lors des requêtes formulées lors des recherches dans la base, mais aussi d'adapter la description des images au contexte, c'est-à-dire aux mots-clé du système et à leur représentation visuelle sous forme d'images exemples et contre-exemples.

Les conclusions que nous pouvons tirer de nos travaux en imagerie couleur sont les suivantes :

- l'analyse d'images couleur au moyen des méthodes issues de la morphologie mathématique nécessite de disposer d'opérateurs aptes à traiter de telles images ; il est donc nécessaire de disposer d'une part d'un espace de représentation de la couleur approprié (par exemple un espace polaire de type teinte – saturation – luminance) et d'une relation d'ordre pour comparer les couleurs de chaque pixel représentées sous forme vectorielle ;
- outre le choix d'un espace couleur et d'un ordre vectoriel, l'analyse morphologique des images couleur (lorsqu'elle est effectuée dans un espace polaire) nécessite également de disposer d'un moyen pour ordonner des valeurs angulaires, voire de pondérer l'importance accordée aux différentes composantes ; les méthodes faisant appel à une unique teinte de référence n'apportent pas des solutions pleinement satisfaisantes et il est préférable d'utiliser de multiples teintes de référence pour assurer un comportement des opérateurs morphologiques plus fidèle à leur définition et au contenu des images ;
- lorsque tous les éléments nécessaires à l'analyse morphologique d'images couleur sont réunis, une grande variété de problèmes peuvent être traités : la détection des objets à l'aide de la transformée en tout-ou-rien, la segmentation en régions à l'aide de la ligne de partage des eaux, la description du contenu à l'aide des descripteurs morphologiques en espace-échelle (comme la granulométrie, la covariance, etc.) ; les expérimentations que nous avons menées nous ont confirmé l'intérêt des approches morphologiques

- pour l'analyse d'image couleur (et notamment la recherche d'image par le contenu), et nous pensons que de telles approches devraient avoir une place de choix dans les alternatives offertes aux traiteurs d'image.
- concernant le problème plus spécifique de la recherche d'image par le contenu, même si la direction prise par les équipes de recherche depuis plusieurs années consiste à élaborer des systèmes où l'utilisateur fournit des exemples visuels, nous restons persuadés de l'intérêt des mots-clé pour aider l'utilisateur à formuler ses requêtes ; en l'absence de données multimédia (texte et image) telles que des pages web ou des images étiquetées (avec des tags, comme dans Flickr), l'utilisation de mots-clé nécessite l'introduction d'une phase préalable d'annotation ; l'annotation par le contenu (plutôt que la recherche par le contenu) permet en outre d'adapter les descriptions des images au contexte, c'est-à-dire aux mots-clé introduits dans le système.

3.1.6 Perspectives

Indexation et recherche dans de grandes bases d'images

Dans le cadre du projet soutenu par Oséo, nous avons conçu un nouveau système pour l'annotation et la recherche par le contenu (cf. section 3.1.4). Ce système nous a permis d'étudier l'intérêt des opérateurs morphologiques pour la description du contenu des images, en complément des approches existantes. Il a également montré comment l'utilisation de mots-clé pouvait être couplée à une recherche par le contenu, et qu'il était aisé de ne pas dépendre d'une liste de descripteurs prédéfinis. Les résultats obtenus sur des bases génériques ou spécifiques de plusieurs milliers d'images sont prometteurs. Néanmoins, le chemin est sans doute encore long pour arriver au niveau des systèmes académiques ou commerciaux les plus performants (ce qui n'était pas notre objectif), qui savent traiter 10 ou 100 fois plus d'images.

Aussi, ce projet présente différentes opportunités d'amélioration. Tout d'abord, le temps de calcul des descripteurs morphologiques reste rédhibitoire pour une utilisation en ligne (c'est également une des raisons qui nous ont poussé à séparer la partie annotation réalisée hors ligne de la partie recherche effectuée en ligne). Nous n'avons pas tenu compte systématiquement des implantations algorithmiques les plus efficaces, et il serait intéressant d'améliorer les performances des éléments morphologiques utilisés dans le système, voire d'étudier la mise en œuvre de ces éléments sur des architectures émergentes (en particulier GPGPU).

Un de nos objectifs initiaux était de concevoir des descripteurs adaptés à chaque mot-clé, par exemple en formulant le descripteur comme une séquence d'opérations (morphologiques) et en cherchant la séquence permettant la description la plus pertinente du mot-clé (c'est-à-dire la meilleure séparation entre les exemples et les contre-exemples, même si la méthode de classification utilisée joue également un rôle important). Cette recherche pourrait être effectuée par des algorithmes génétiques, et nous avons, dans un autre contexte, débuté des travaux dans ce sens avec des chercheurs de l'équipe FDBT (Alexandre Blansché et Pierre Gançarski). A défaut d'une telle possibilité, notre système est aujourd'hui capable de choisir, lors de la définition d'un mot-clé, les descripteurs qui sont les plus adaptés à la description des images associées à ce mot-clé, parmi la liste de descripteurs disponibles.

Un autre élément qu'il nous reste à étudier concerne l'échelle à laquelle l'annotation est effectuée. Dans son état actuel, notre système permet d'annoter chaque image par un ou plusieurs mots-clé mais il serait envisageable d'annoter également des régions d'une image. Cependant, cette approche nécessite de disposer d'une segmentation pertinente, du point de vue de l'utilisateur qui effectue l'annotation, de toutes les images de la base. Autant il est aisé d'effectuer une telle segmentation de façon semi-automatique pour les images fournies comme exemple ou contre-exemple (cette possibilité est offerte dans le système pour calculer des descriptions sur les zones d'intérêt de ces images), autant cette tâche devient ardue voire impossible à réaliser lorsqu'elle est effectuée de façon totalement non-supervisée. Deux solutions peuvent être envisagées : la segmentation multi-échelle et l'apprentissage de segmentations. Dans le premier cas, il s'agit d'effectuer plusieurs segmentations à différentes échelles, sous l'hypothèse que les régions d'intérêt soient présentes à une des échelles. Dans le second cas, il s'agit d'apprendre comment l'utilisateur a sélectionné les zones d'intérêt dans les images exemples et contre-exemples : les segmentations de ces images fournissent un certain point de vue, qui sera ensuite utilisé pour segmenter automatiquement toutes les images de la base à annoter.

Points d'intérêt et descripteurs locaux

La recherche d'image par le contenu a connu des progrès récents, permis par l'arrivée d'outils dédiés à la caractérisation locale des images. Ainsi, les descripteurs de type SIFT [Low99] sont très fréquemment utilisés aujourd'hui pour comparer des images ou parties d'image. Ils s'appuient sur deux étapes principales : la détection dans l'image de points caractéristiques (appelés points d'intérêt) assurant certaines propriétés d'invariance (à la rotation, à l'échelle), puis la description texturale du voisinage de ces points (en étudiant par exemple l'information de contour dans toutes les orientations). Notons qu'il est aussi possible d'extraire non pas des points d'intérêt mais des zones d'intérêt (ou patches), la description d'une image portant alors sur leur contenu et/ou voisinage. Ces descripteurs locaux assurent dans de nombreux cas une meilleure performance que les approches globales classiquement utilisées en recherche par le contenu et pour lesquelles nous avons proposé des alternatives morphologiques. Nous pourrions donc nous interroger sur l'apport de la morphologie mathématique quant à l'extraction de points ou régions d'intérêt, et à leur caractérisation pour produire des descriptions locales d'image.

Ainsi, la détection des points d'intérêt pourrait être effectuée en utilisant une transformée en tout-ou-rien, dans sa définition initiale si l'image est à niveaux de gris, ou à l'aide d'une des extensions que nous avons proposées si l'image à analyser est multivaluée (image couleur, ou image d'attributs relatifs à la texture ou au gradient par exemple). Nous avons montré que cet opérateur morphologique pouvait être utilisé pour la détection d'objet. Encore faut-il lui assurer une grande robustesse, les propositions que nous avons formulées dans ce but n'étant certainement pas suffisantes si la représentation de l'objet recherché est trop variable (elles le sont néanmoins si on suppose que les motifs recherchés sont bien définis, ce qui est le cas des points ou régions d'intérêt). De plus, l'invariance à différentes transformations (rotation, changement d'échelle ou de contraste) doit également être prise en compte, et nous avons commencé à étudier ce problème en collaboration avec l'équipe de Hans Burkhardt de l'Université Albert-Ludwig à Freiburg-im-Breisgau en Allemagne, spécialisée dans les invariants géométriques en traitement d'image. Disposer d'une transformée en tout-ou-rien robuste et invariante pour la reconnaissance de motifs permettrait alors d'apporter des réponses morphologiques à la détection des points ou zones d'intérêt.

La morphologie mathématique offre également des solutions pour la description locale des points ou zones d'intérêt (ou de leurs voisinages). En effet, l'ensemble des descripteurs morphologiques basés sur des espaces-échelle (décrits en section 2.1.5 pour les travaux existants, et en section 3.1.3 pour nos contributions), même s'ils considèrent une image dans son ensemble (excepté les profils morphologiques qui sont utilisés dans le contexte de la télédétection), peuvent être aisément calculés de façon locale, et limités au voisinage de points ou zones d'intérêt. Nul doute que les capacités de description de ces approches morphologiques, observées depuis de nombreuses années à l'échelle globale des images, persistent à l'échelle locale. Ces solutions morphologiques offriraient ainsi une alternative intéressante aux descripteurs de type SIFT, présentant plusieurs avantages dont le caractère non-linéaire, un coût calculatoire moindre, et une meilleure compacité.

3.2 Télédétection

Nous présentons dans cette section les travaux que nous avons menés dans le domaine de la télédétection, où nous nous sommes intéressés à l'analyse d'images optiques à très haute résolution spatiale. Ces travaux ont été réalisés en étroite collaboration avec des équipes de géographes de Strasbourg (notamment Anne Puissant, Maître de Conférences au LIVE¹) ou de Caen, mais aussi avec l'équipe FDBT du LSIIT spécialisée en fouille de données, et qui possède une longue expérience quant à l'analyse (principalement en classification) d'images satellite. Ces collaborations se sont notamment traduites par plusieurs co-encadrements : la thèse de doctorat de Sébastien Derivaux a traité de la segmentation et de la classification d'images de télédétection guidées par des connaissances du domaine [Der09], tandis que les stages de master 1 et master 2 recherche de Jonathan Weber ont respectivement porté sur l'extraction d'objets urbains (bâti) et du littoral (trait de côte).

Nous débuterons cette section par une introduction présentant le contexte dans lequel se sont placés nos travaux, en rappelant les usages actuels de la morphologie mathématique en télédétection et en discutant les

¹Laboratoire Image, Ville, Environnement — UMR 7011 CNRS-ULP, <http://imaville.u-strasbg.fr>

particularités des images à très haute résolution spatiale. Nous nous intéresserons ensuite aux images représentant des environnements urbains que nous avons principalement étudiées dans le cadre du projet FoDoMuST. Après avoir décrit les grandes lignes de ce projet, nous présenterons les résultats obtenus en segmentation et en détection de structures bâties qui ont fait l'objet de plusieurs publications en revue nationale [C4] ou en conférences internationales [D14, D17, D20] ou nationales [E9, E10]. Nous commenterons également notre participation à une compétition internationale pour la détection de bâtiments [D27]. Outre les environnements urbains, nous avons aussi étudié les environnements côtiers au travers du projet ECOSGIL. Dans ce projet, nos efforts ont porté sur l'analyse d'images du littoral pour y extraire des entités géographiques telles que le trait de côte. Nous présenterons les résultats obtenus par deux méthodes alternatives (basées respectivement sur une segmentation ou une reconnaissance de formes) qui se sont traduits par plusieurs publications en revue nationale [C5] ou conférences internationales [D24, D25]. Nous conclurons cette section en dégagant les principales conclusions pouvant être tirées de ces expériences et en discutant les perspectives offertes par ces travaux.

3.2.1 Introduction

La télédétection permet l'observation de la terre à partir d'images aériennes ou satellites obtenues par différentes modalités d'images. Nous nous sommes principalement intéressés aux images optiques, acquises dans le spectre du visible ou de l'infra-rouge, et qui permettent une interprétation relativement simple de la part de l'utilisateur. Ces images peuvent être limitées à une seule bande ou image à niveaux de gris ou consister en plusieurs composantes. Dans le premier cas, on utilise le terme d'images panchromatiques, tandis que dans le second cas il peut s'agir d'images couleur (orthophotos acquises depuis un avion), d'images multispectrales (avec un nombre de bandes spectrales restant faible, de l'ordre de 4 à 6 canaux) ou hyperspectrales. Ainsi, pour ne donner qu'un exemple, Spot 5, le dernier né de la famille des satellites SPOT lancé en 2002, offre une information accessible sur 4 canaux (Bleu, Vert, Rouge, Proche Infra-Rouge) à 10 mètres de précision spatiale, et une information panchromatique à 5 mètres de résolution spatiale.

Deux critères essentiels jouent fort logiquement sur la qualité des traitements d'image pouvant être appliqués en télédétection, la résolution spectrale et la résolution spatiale : dans les images panchromatiques (sans information spectrale), la détection, la description et la reconnaissance des objets présents dans les images ne pourront se faire que sur des critères géométriques ou topologiques ; au contraire, l'analyse des images mutispectrales à faible résolution spatiale (la résolution est de 80 mètres par pixel pour les premiers satellites de la gamme LANDSAT) porte essentiellement (voire exclusivement) sur des critères spectraux, sans utilisation de l'information spatiale. Néanmoins, dans un futur proche, il sera possible de disposer d'images multispectrales d'une précision spatiale très intéressante : les satellites PLEIADES dont les lancements sont prévus en 2010 et 2011 proposeront des images panchromatiques à 70 centimètres par pixel et des images multispectrales à 4 canaux et 2,8 mètres par pixel. Le besoin de techniques de traitement d'image tenant compte à la fois de la nature spatiale et spectrale des données devrait donc s'accroître dans les années à venir. Dans ce contexte, la télédétection nous a semblé un domaine d'application opportun pour les travaux fondamentaux que nous avons menés en morphologie mathématique.

En effet, comme dans de nombreux autres domaines, les approches morphologiques restent limitées en télédétection, et loin en pratique des possibilités offertes par les avancées théoriques ou méthodologiques présentées dans la communauté de la morphologie mathématique². Ce constat a été fait par Soille et Pesaresi en 2002 qui, dans leur article de synthèse [SP02], présentaient les principaux usages de la morphologie mathématique et introduisaient quelques applications qui pourraient bénéficier des avancées récentes en morphologie mathématique. Ce constat restant d'actualité, nous rappellerons ici quelques éléments donnés dans [SP02] : les opérateurs morphologiques principalement utilisés sont les filtres morphologiques (ouverture/fermeture ou filtres alternés séquentiels) directionnels ou isotropiques, les chapeaux haut-de-forme, les granulométries ou profils morphologiques, les gradients morphologiques, les opérateurs géodésiques et la ligne de partage des eaux ; les utilisations qui sont faites de ces opérateurs sont la réduction de bruit, l'analyse de texture, la détection d'objets, la seg-

²Notons le cas particulier de Pierre Soille, à l'initiative de nombreuses avancées en morphologie mathématique qui trouvent des applications en télédétection.

mentation, la classification et l'interpolation.

Nous verrons dans les sections suivantes que, pour notre part, nous nous sommes essentiellement intéressés à la segmentation et à la détection d'objets avec des outils tels que la ligne de partage des eaux et la transformée en tout-ou-rien. Nous avons fait le choix de considérer des images à haute et si possible très haute résolution spatiale, de préférence multispectrales ou, à défaut, panchromatiques. Les dénominations utilisées pour définir le niveau de résolution sont celles proposées dans [Pui03] : basse résolution (1000 mètres), moyenne résolution (80 mètres), haute résolution – HR (10 à 30 mètres), très haute résolution – THR (moins de 5 mètres). Ces définitions sont bien sûr valables pour le domaine civil uniquement (dans le domaine militaire, une image moyenne résolution a une précision de 3 mètres par pixel, tandis qu'une image à très haute résolution a une précision de 25 à 60 centimètres par pixel !).

L'avènement des capteurs HR et THR renforce l'intérêt que peut avoir la morphologie mathématique dans le domaine de la télédétection, en rendant incontournable le contexte spatial dans toute analyse d'image. Ce constat est illustré par la figure 3.8 qui présente une seule et même scène à différentes résolutions spatiales. On peut y observer qu'à faible résolution l'information spectrale est suffisante (et nécessaire), tandis que l'importance de l'information spatiale augmente avec le niveau de résolution. A très haute résolution (mais parfois aussi à haute résolution), il faut adopter une approche "orientée-objets" telle que proposée par Anne Puissant dans sa thèse [Pui03].

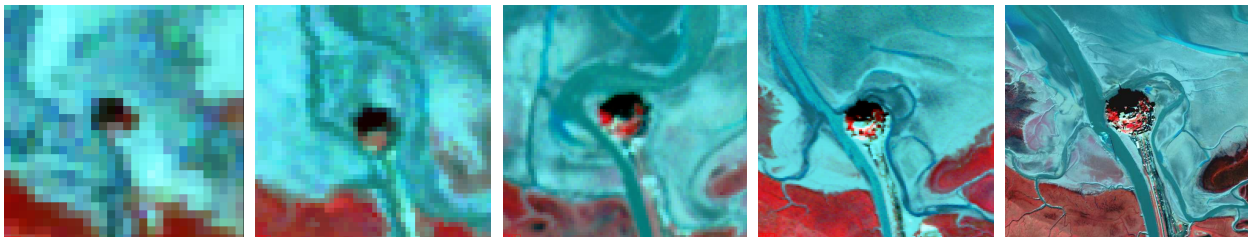


FIG. 3.8 – Images satellites à différentes résolutions représentant une même scène [Pui08], le Mont-Saint-Michel (de gauche à droite) : Landsat 60m, Landsat 30m, Aster 15m, Spot5 10m, Quickbird 2.4m.

L'élément central du processus d'analyse d'image n'est alors plus le pixel et sa réponse spectrale, mais plutôt la région qui peut être décrite par des attributs spectraux, texturaux, et structurels. Dans ce contexte, la ligne de partage des eaux et les profils morphologiques sont sans doute les outils morphologiques parmi ceux bénéficiant aujourd'hui de la plus grande attention, et on pourra trouver dans la littérature récente de nombreux travaux exploitant la première pour segmenter des images THR (cf. l'étude comparative de Carleer et al. [CDW05]) et les seconds pour détecter, classer et reconnaître des objets selon leur contenu structurel (cf. les travaux de Pesaresi et al. sur l'analyse des structures bâties [PP07] ou ceux de Plaza et al. sur la classification d'images hyperspectrales [PMPP04]). La figure 3.9 montre la pertinence d'une classification "orientée-objets" sur une classification directe des pixels, soulignant ainsi l'intérêt de l'étape de segmentation nécessaire pour obtenir les régions à classer.

L'importance de l'étape d'extraction des objets dans le paradigme actuel d'analyse et d'interprétation d'image en télédétection a motivé nos choix quant aux problèmes que nous avons abordé dans ce domaine. Nous nous sommes ainsi intéressés aux étapes cruciales que sont la segmentation et la détection d'objets, dans différents environnements (urbain et littoral).

3.2.2 Télédétection en environnement urbain

Le premier domaine d'étude auquel nous nous sommes intéressés en télédétection a été les zones urbaines, principalement au travers du projet FoDoMuST. Après avoir décrit brièvement ce projet, nous présenterons quelques résultats obtenus en segmentation guidée ou partiellement guidée qui ont été obtenus en utilisant les techniques discutées dans la section 2.3. Nous décrirons également l'utilisation de la transformée en tout-ou-rien pour extraire les structures bâties.

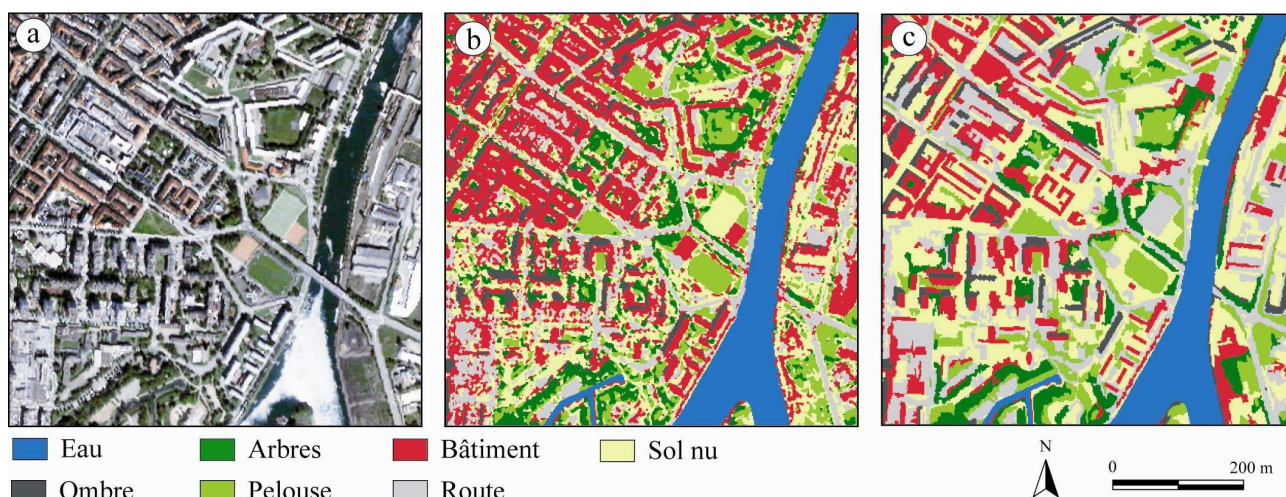


FIG. 3.9 – Classification d’une image à très haute résolution spatiale [Pui08] : image originale (a), classification spectrale des pixels (b), classification des régions à base de règles (c).

FoDoMuST

Dans le cadre de l’ACI Masse de Données, le projet FoDoMuST a été mené conjointement par 3 laboratoires : le LSIIT pour ses compétences en fouille de données et analyse d’image, le LIVE pour sa connaissance du domaine étudié, et l’équipe ERIC³ pour ses compétences en structuration et fouille de données et de connaissances. L’objectif de ce projet était la fouille de données multistratégies pour extraire et qualifier la végétation urbaine.

Ce projet, dont un synopsis général est donné en figure 3.10, était structuré en 4 parties :

1. conceptualisation des connaissances multi-échelles (ontologie du domaine) et aide à l’interprétation,
2. structuration des données et représentation des connaissances,
3. construction des “objets” ou segmentation,
4. classification par une approche multi-stratégie.

Nous nous sommes intéressés et avons coordonné les travaux relatifs à la construction des “objets”. L’objectif consistait à élaborer des méthodes de construction d’objets suffisamment robustes pour pouvoir fournir des objets d’intérêt pertinents à l’étape ultérieure de classification, alors même que les images à traiter étaient de nature très variée. Les connaissances des experts, formalisées lors d’une étape précédente, devaient nous permettre de mener à bien cette tâche.

Deux stratégies différentes ont été suivies pour permettre la construction des objets en se basant sur les connaissances des experts : la segmentation et la détection. Tandis que la segmentation fournit une partition complète de l’image analysée en régions sans se focaliser sur un objet bien particulier, la détection au contraire cherche à extraire de l’image toutes les occurrences d’une classe donnée : pour obtenir la liste complète des régions d’intérêt, il faut répéter cette étape de détection pour chacun des concepts recherchés. Ces stratégies ont été notamment étudiées dans le cadre la thèse de Sébastien Derivaux [Der09]. Ainsi, l’intégration de connaissances dans le processus de segmentation peut se traduire par un changement d’espace (en pré-traitement) ou une évaluation des paramètres de réduction de sursegmentation (en post-traitement), comme nous l’avons vu lors de la section 2.3. Nous illustrerons ici ces différentes approches par quelques résultats obtenus dans le cadre de ce projet. La détection des objets développée dans [Der09], pour sa part, n’a fait qu’un appel limité aux approches morphologiques : nous ne la décrirons donc pas ici, mais illustrerons ce problème par d’autres travaux réalisés en morphologie mathématique dans le cadre du projet FoDoMuST.

Outre ces travaux, le projet FoDoMuST nous a également permis de nous intéresser aux problématiques de la segmentation multi-échelle et de la fusion de segmentations. Néanmoins, faute de temps, nous n’avons pas

³Equipe de Recherche en Ingénierie des Connaissances – EA 3083 Université Lyon II, <http://recherche.univ-lyon2.fr/eric>

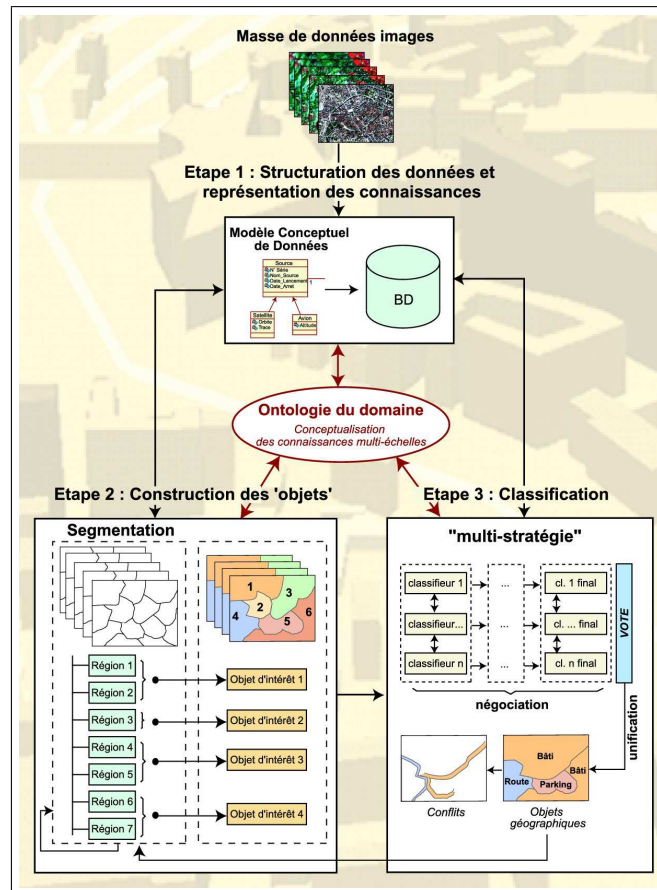


FIG. 3.10 – Architecture proposée dans le projet FoDoMuST

encore pu proposer des solutions pleinement satisfaisantes à ces problèmes, que nous chercherons à résoudre dans des projets à venir, notamment en télédétection (où les données disponibles proviennent de différentes sources et sont associées à différentes échelles).

Segmentation guidée sans marqueur

Nous avons vu dans la section 2.3 qu'il était possible d'intégrer des connaissances au sein du processus de segmentation, en s'appuyant par exemple sur des techniques de classification supervisée. Les exemples fournis à l'algorithme sont alors exploités uniquement au travers de leur contenu spectral (et non spatial). Ils permettent d'une part d'effectuer un changement d'espace sur lequel l'algorithme de ligne de partage des eaux sera appliqué (les régions seront alors des zones homogènes en termes de degrés d'appartenance à des classes prédéfinies, et non plus en termes de réponses spectrales) [D14], et d'autre part d'optimiser les paramètres de réduction de sursegmentation (puisque le changement d'espace ne suffit pas à éviter le problème de sursegmentation) [D20]. Cette méthode de segmentation supervisée (par changement d'espace, optimisation de paramètres, ou les deux) a été appliquée en télédétection dans le cadre de la thèse de Sébastien Derivaux [Der09]. Dans la figure 3.11, nous présentons un exemple de résultat obtenu sur une image de Strasbourg à très haute résolution en considérant 3 classes d'intérêt : la végétation, les routes, et les bâtiments. On peut y constater la complémentarité des approches de changement d'espace et d'optimisation de paramètres, et ainsi la pertinence d'un couplage de ces deux approches. A titre de comparaison, les résultats obtenus à l'aide des logiciels ENVI⁴ et eCognition⁵, fréquemment utilisés dans la communauté de télédétection, sont également indiqués.

De plus, nous avons aussi tenté d'utiliser d'autres descripteurs ou d'autres paradigmes de classification. Ainsi, différents changements d'espaces (modifiant le relief à inonder dans l'algorithme de ligne de partage

⁴ENVI FX est édité par ITT Visual Information Solutions : <http://www.ittvis.com>

⁵eCognition est édité par Definiens : <http://earth.definiens.com>



FIG. 3.11 – Intérêt d'une segmentation supervisée pour l'interprétation / classification d'images satellite à très haute résolution spatiale : (a) image initiale, (b) zones d'apprentissage, (c) classification orientée-objets produite à partir d'une segmentation supervisée. Comparaison des résultats obtenus en segmentation d'images urbaines sans marqueur, par différentes méthodes : (d) segmentation LPE classique, (e) segmentation supervisée par changement d'espace, (f) segmentation supervisée par optimisation de paramètres, (g) segmentation supervisée par changement d'espace et optimisation de paramètres, (h) utilisation du logiciel ENVI, (i) utilisation du logiciel eCognition.

des eaux) ont été effectués, en utilisant des indices de texture [E13] ou une technique de classification non-supervisée [E10]. Les résultats obtenus étant relativement proches de ceux présentés en figure 3.11, nous renvoyons le lecteur aux articles originaux ou à la thèse de Sébastien Derivaux pour plus de détails. Notons néanmoins qu'ils illustrent parfaitement les possibilités d'extension de notre proposition, à d'autres attributs ou méthodes de classification.

Segmentation guidée et partiellement guidée avec marqueurs

Outre les expérimentations menées en segmentation guidée sans faire appel à des marqueurs, nous avons également cherché à appliquer la méthode de segmentation supervisée avec marqueurs [D19] (dont une description est donnée dans la section 2.3) aux images de télédétection représentant des scènes urbaines. Ainsi, nous avons participé au challenge organisé dans le cadre du workshop PRRS 2008 (événement satellite à la conférence ICPR 2008), dont l'objectif était la reconnaissance des bâtiments. Les résultats obtenus [D27] ont été très encourageants : en effet, alors même que la méthode utilisée (celle décrite dans la section 2.3) n'était en aucun point dédiée au problème considéré (la détection des bâtiments dans la ville de Legaspi, aux Philippines), elle a pu fournir des résultats d'un niveau équivalent aux approches concurrentes, ces dernières étant au contraire élaborées de façon *ad hoc*.

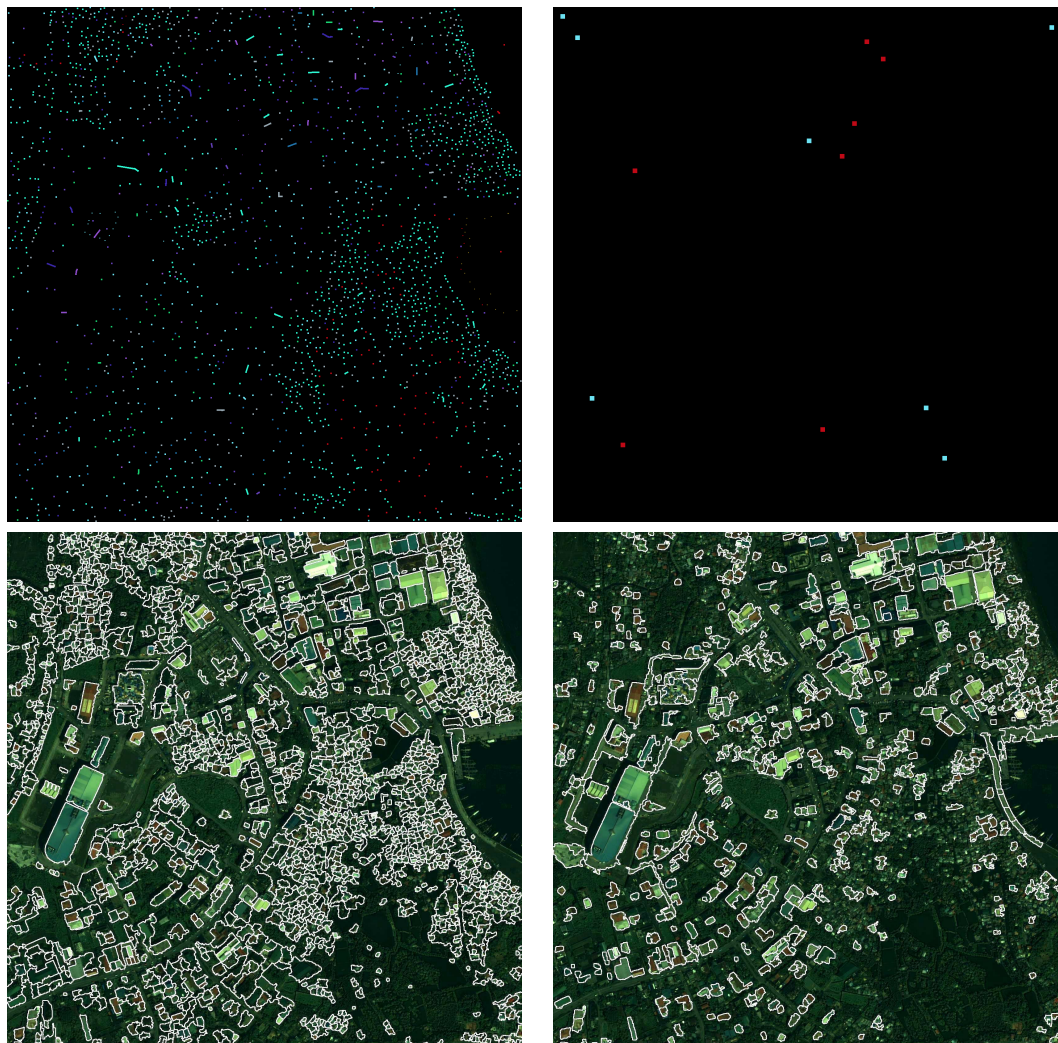


FIG. 3.12 – Marqueurs utilisés (en haut) et résultats obtenus (en bas) lors de la segmentation d'images urbaines (PRRS'08) par une approche supervisée (gauche) et semi-supervisée (droite).

Néanmoins, l'utilisation directe de l'approche de segmentation supervisée avec marqueurs présente un inconvénient majeur, puisqu'il est nécessaire de positionner un marqueur pour chacun des objets à extraire. Sur

l'image à traiter, d'une taille de 1668×1668 pixels et contenant environ 3000 bâtiments, cette initialisation a nécessité l'intervention (fastidieuse : 90 minutes !) d'un utilisateur : celui-ci a saisi 2460 objets (sous la forme de carrés de 5×5 pixels) répartis en 10 classes (6 pour les différents types de surface bâtie, 4 pour les autres classes d'occupation ou de couverture du sol). Cette expérience nous a donc permis de constater les limites d'une telle approche et sa non-pertinence dans le cas où l'image à segmenter contient de très nombreux objets.

Afin de limiter l'intervention de l'utilisateur, nous avons élaboré, très rapidement (dans un laps de temps compatible avec l'échéancier associé à la compétition PRRS 2008) et de façon relativement empirique, un prétraitement permettant une saisie partielle des marqueurs : ainsi, en limitant la saisie à quelques marqueurs, il est possible d'effectuer une classification supervisée de l'image puis d'appliquer un filtrage spatial basé sur des critères morphologiques : les composantes connexes retenues sont alors ajoutées à la liste des marqueurs, et l'algorithme de segmentation supervisé peut s'appliquer. Nous avons obtenu des résultats intéressants, et ce en limitant grandement l'intervention de l'utilisateur : quelques secondes seulement ont été nécessaires pour saisir les 14 marqueurs utilisés (7 pour la classe bâti, 7 pour la classe non-bâti), et la carte de segmentation a ensuite été obtenue en une à deux minutes. Ce temps de calcul, très raisonnable, permet d'envisager l'utilisation de cette méthode dans un cadre interactif, où l'utilisateur pourrait corriger les marqueurs en fonction de la pertinence des résultats obtenus.

Nous illustrons dans la figure 3.12 les résultats obtenus avec l'approche supervisée (à gauche) et l'approche semi-supervisée (à droite). Notons que cette dernière approche aurait certainement fourni de meilleurs résultats si l'étape empirique de prétraitement avait été étudiée de façon plus approfondie. Il s'agit là d'une voie principale d'amélioration de la méthode de segmentation supervisée.

Détection des structures bâties

De façon complémentaire à nos travaux en segmentation, nous avons aussi cherché à extraire les zones d'intérêt en mettant en œuvre des approches de détection d'objet. En nous focalisant sur les structures bâties, nous avons ici adopté une approche moins supervisée, où l'intégration de connaissances portait principalement sur la forme (rectangulaire) des objets recherchés, leur taille (dans un intervalle relativement large), et le nombre de classes spectrales représentées dans l'image.

L'approche utilisée procède par décomposition d'une image (panchromatique ou multispectrale) en un ensemble d'images binaires, celles-ci étant obtenues par un processus de classification non-supervisée (cf. sections 2.2 et 2.3). Chaque image est alors traitée au moyen d'opérateurs morphologiques, en étant préalablement

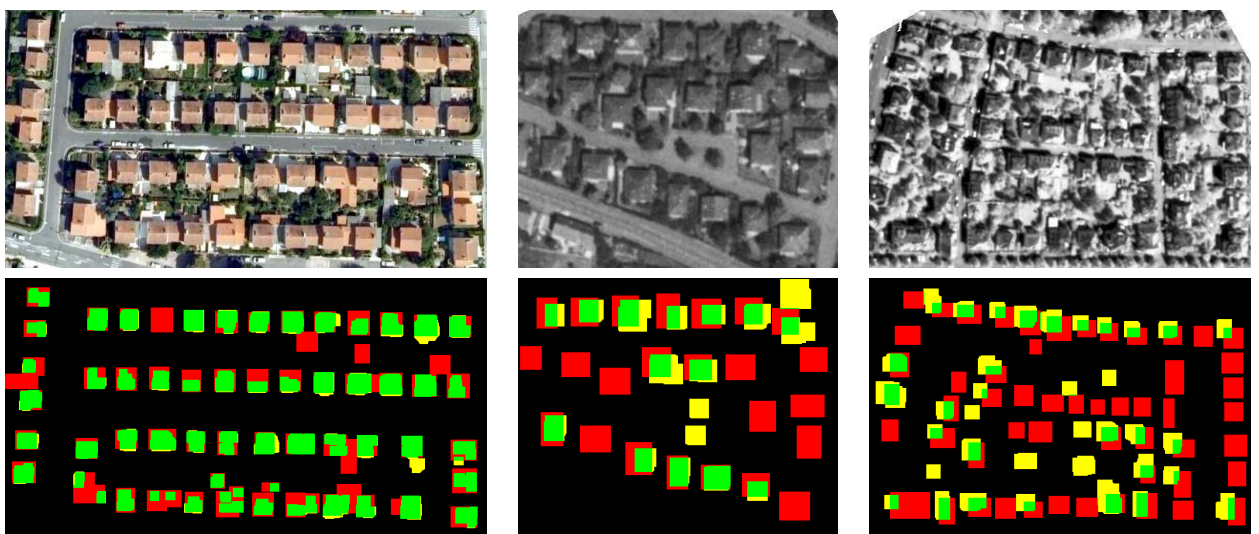


FIG. 3.13 – Images analysées (en haut) et résultats obtenus (en bas) lors d'une détection des structures bâties par combinaison d'une classification non-supervisée et d'une transformée en tout-ou-rien. Les détections correctes sont affichées en vert, les détections manquées ou faux négatifs en rouge, les détections incorrectes ou faux positifs en jaune.

filtrée (les paramètres du filtre morphologique étant obtenus par une granulométrie bidimensionnelle) avant d'être analysée à l'aide d'une transformée en tout-ou-rien avec différentes tailles et orientations pour identifier les occurrences des bâtiments dont la forme exacte est finalement obtenue par un processus de reconstruction géodésique [C4, D17, E9].

La figure 3.13 illustre les résultats obtenus. La méthode n'éprouve pas de difficulté dans le cas des images multispectrales, où l'information spectrale permet de bien séparer les bâtiments des autres classes (à gauche). Dans le cas des images panchromatiques, la détection peut s'avérer plus délicate du fait du peu d'informations utilisées. Néanmoins, sur un ensemble d'images représentatif (principalement tirées des données analysées dans le cadre du projet FoDoMuST), la précision globale de la méthode était de l'ordre de 90 %.

3.2.3 Télédétection en environnement du littoral

En complément des travaux portant sur l'analyse de scènes urbaines, nous avons également traité certains problèmes posés en télédétection lorsque les zones d'étude concernent le littoral. Ces travaux ont été réalisés dans le cadre du projet Ecosgil que nous décrirons brièvement. Nous nous focaliserons ensuite sur un objet d'étude particulier, le trait de côte, pour lequel nous avons expérimenté deux approches alternatives. En effet, de façon similaire au cas des bâtiments dans les scènes urbaines, nous avons procédé à l'extraction du trait de côte, soit à partir d'une segmentation guidée, soit à partir d'une détection d'objet.

ECOSGIL

Le projet Ecosgil⁶ est un projet ANR Jeunes Chercheurs pluridisciplinaire rassemblant différentes équipes de recherche de Strasbourg (LSIT, LIVE) et de Caen (IDEES-Geosyscom, GREYC-Dodola, LETG-Geophen, et M2C). Il avait pour objectif le développement d'une plate-forme interactive d'archivage, d'identification, d'analyse et de représentation de plusieurs objets d'études environnementaux présents en Normandie, à partir de données d'information géographiques, en particulier des données images à très haute résolution. Il était structuré en quatre axes de recherche :

1. évaluation des besoins et apport des images Haute Résolution et Très Haute Résolution,
2. développement de méthodes automatiques d'extraction,
3. définition d'un processus optimal de choix des données et de traitement par le développement d'une plate-forme interactive de gestion des données et des connaissances du domaine,
4. exploitation et valorisation de la plate-forme interactive

Notre participation s'est inscrite principalement dans le second axe et a consisté en la recherche de méthodes automatiques ou semi-automatiques pour l'extraction d'objets du littoral, en s'appuyant sur les connaissances des experts. Un des principaux obstacles que nous avons rencontrés concerne la difficulté que peuvent éprouver les experts à formuler leurs connaissances : or, sans formalisation des connaissances ou informations a priori, il est délicat d'élaborer des approches à bases de connaissances ! Les méthodes auxquelles nous nous sommes intéressées (segmentation ou détection), bien qu'elles soient par essence génériques, ont donc été expérimentées en se focalisant sur un type d'objet particulier : le trait de côte. Celui-ci se présente sous différentes formes selon qu'il s'agisse de marais, de plages, ou de falaises (cf. figure 3.14).

Segmentation guidée avec marqueurs

Comme nous l'avons indiqué précédemment, nous avons constaté lors du projet Ecosgil à quel point il était particulièrement délicat de formaliser les connaissances des experts. Néanmoins, il est possible d'éviter une telle formalisation en introduisant l'expert directement dans le processus de traitement. Ainsi, nous avons fourni aux experts géologues et géomorphologues (notamment Romain Desguée et Franck Levoy du laboratoire M2C⁷) un outil semi-automatique leur permettant d'extraire des objets d'intérêt. Cet outil a consisté en

⁶<http://ecosgil.u-strasbg.fr>

⁷Morphodynamique Continentale et Côtière — UMR 6143 CNRS-UCBN, <http://www.geos.unicaen.fr>

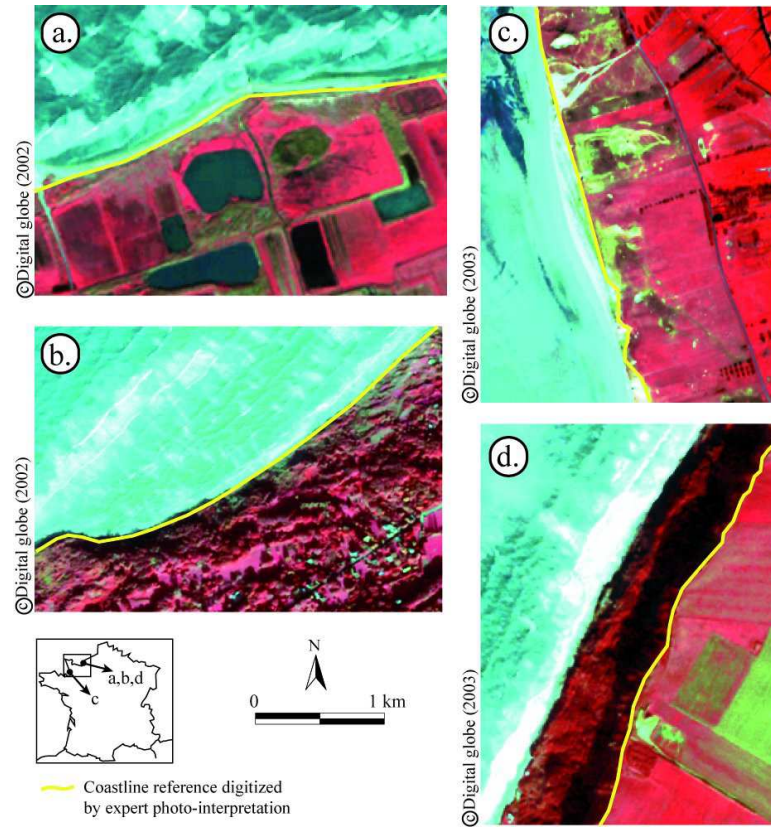


FIG. 3.14 – Illustration de l’objet “trait de côte” : (a) zones marécageuses, (b) côte de roche meuble (soumise à des glissements de terrain), (c) plages de sable avec dunes, (d) côte à falaise (soumise à chute de blocs).

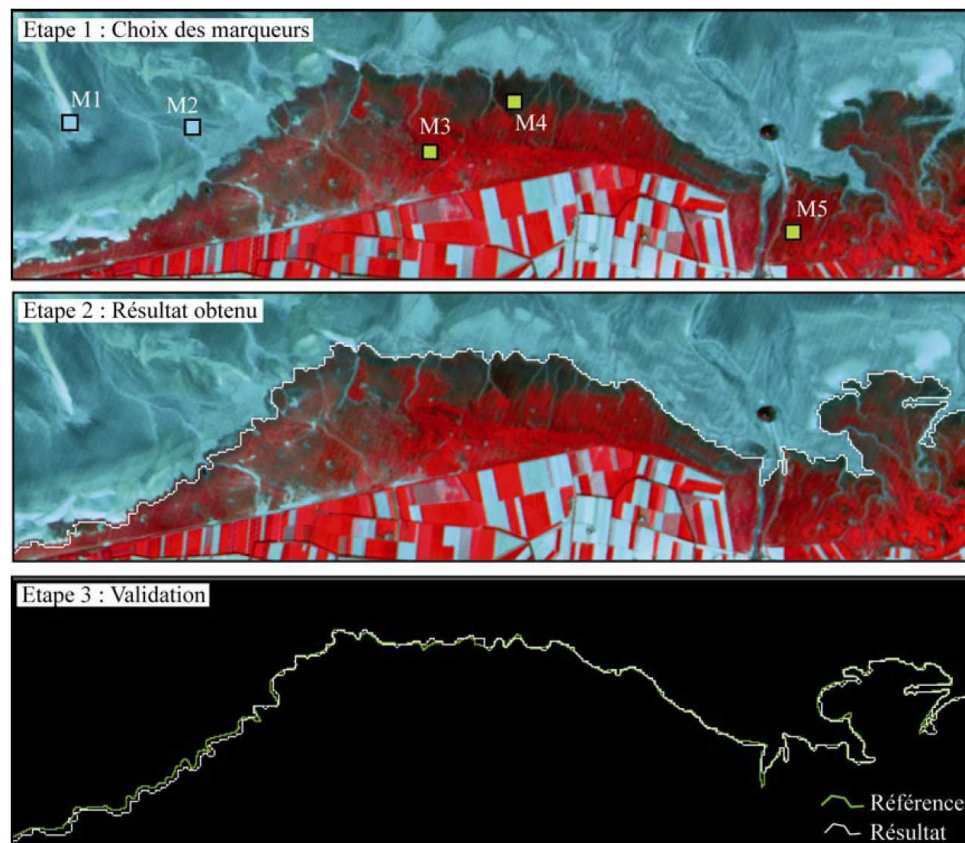


FIG. 3.15 – Segmentation guidée par quelques marqueurs pour l’extraction du trait de côte.

une adaptation de la méthode de segmentation guidée par les marqueurs [D19] aux images multispectrales (la méthode ayant été initialement développée pour des images en couleurs et de taille plus réduite).

Le principal avantage de cette méthode de segmentation est sans doute son extrême simplicité d'utilisation : le trait de côte peut très souvent être extrait à partir de deux clics de souris uniquement (pour saisir un minimum de deux marqueurs : un pour la terre et un pour la mer). Même dans le cas d'une scène plus complexe, l'intervention reste très limitée et le faible temps de calcul de l'algorithme offre à l'utilisateur une véritable interaction lui permettant de modifier ses marqueurs à loisir. Notons que cette méthode a permis aux partenaires du projet Ecosgil de cartographier et de suivre l'évolution du trait de côte (et plus précisément des schorres) dans la baie du Mont-Saint-Michel, de façon (semi-)automatisée, sur les 20 dernières années (un exemple d'utilisation de la méthode est donné en figure 3.15). Cette étude a fait l'objet d'une publication conjointe dans une revue nationale spécialisée [C5].

Détection du trait de côte

Plutôt que de faire appel à l'expert directement lors du processus d'analyse d'image, il est possible de faire l'hypothèse qu'un ensemble de connaissances formalisées soit disponible et puisse servir à la paramétrisation ou la spécialisation d'un détecteur d'objet générique [Web07]. Dans le cas d'un objet de type "frontière" tel que le trait de côte, nous avons élaboré un détecteur de contours spécifiques au moyen de la transformée en tout-ou-rien multispectrale : celle-ci est appliquée dans toutes les orientations possibles pour rechercher les zones de séparation entre terre et mer, en se basant sur des connaissances a priori sur ces deux zones. Les résultats obtenus étaient très prometteurs et ce quelle que soit la résolution considérée, comme l'illustre la figure 3.16. Néanmoins la caractérisation des objets par l'expert s'est avérée bien plus complexe que prévue, diminuant ainsi l'intérêt pratique d'une telle approche (nous avons observé qu'un utilisateur préfère cliquer à la souris sur des pixels de terre et des pixels de mer plutôt que de fournir une description fiable de ces zones en termes spatial et spectral).

Utilisée telle quelle (sans post-traitement), l'application de la transformée en tout-ou-rien dans toutes les

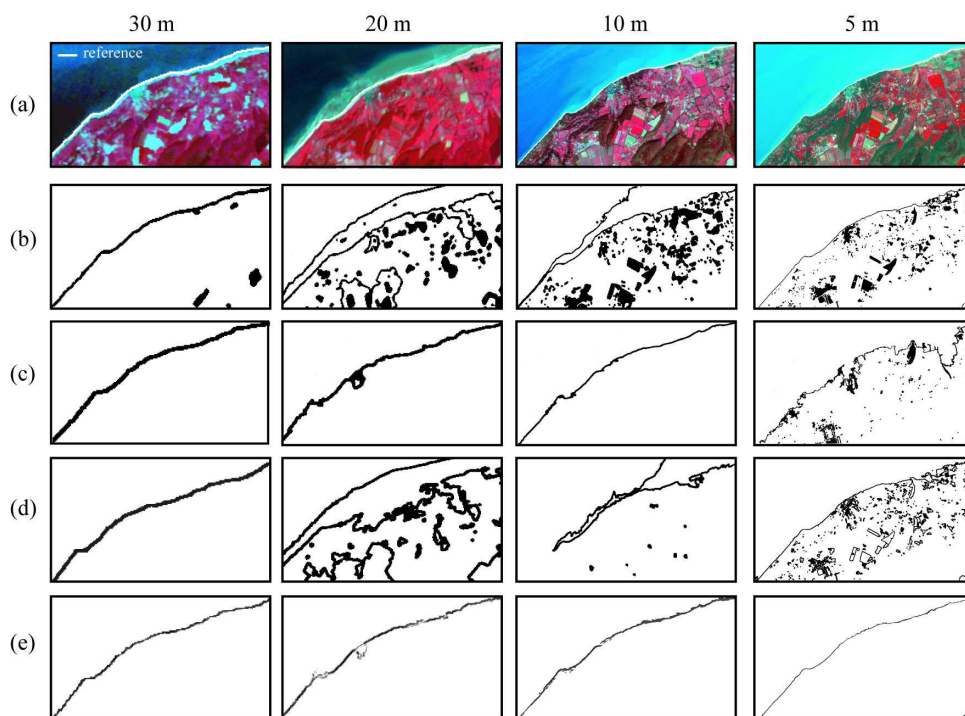


FIG. 3.16 – Comparaison de différentes méthodes de détection du trait de côte : (a) image originale et référence, (b) méthode de Heene et Gautama [HG00], (c) méthode de Jishuang et Chao [JC02], (d) méthode de Bagli et Soille [BS03], et (e) méthode proposée par transformée en tout-ou-rien multispectrale.

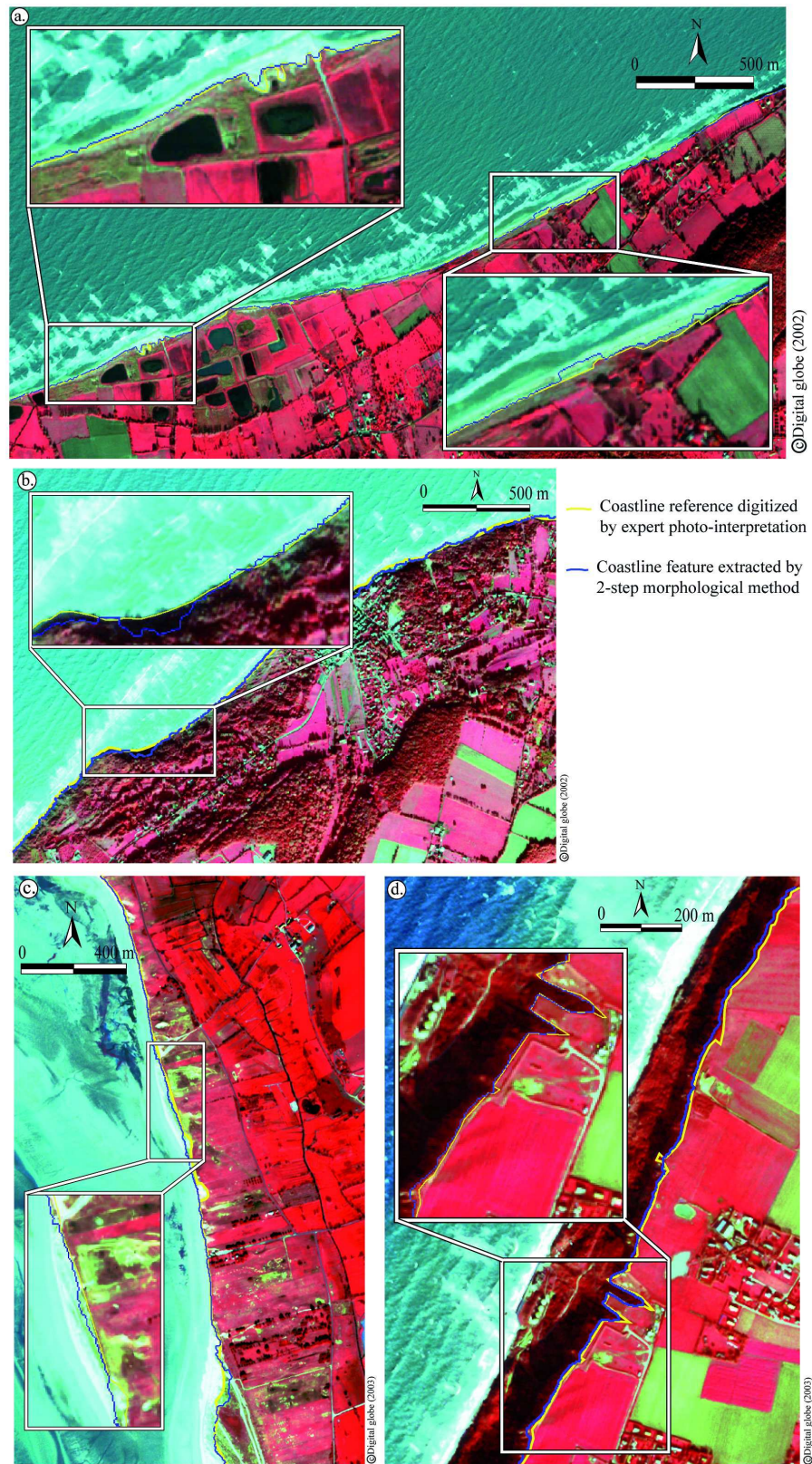


FIG. 3.17 – Détection du trait de côte dans les différents environnements du littoral de la figure 3.14.

orientations n'offre pas la garantie de disposer d'un résultat sous la forme d'un trait continu. Nous avons donc par la suite intégré cet opérateur dans un processus plus complexe, impliquant également un double seuillage, une reconstruction géodésique, et une segmentation par ligne de partage des eaux. Les résultats obtenus sont très intéressants (figure 3.17), bien qu'ils nécessitent de formaliser les connaissances à utiliser [D24,D25], étape qui peut s'avérer délicate.

3.2.4 Conclusion

Dans cette section, nous avons présenté les travaux que nous avons menés dans le domaine de la télédétection, et qui ont consisté en la mise en œuvre (voire l'adaptation aux données satellites) de certaines des approches multivaluées et/ou supervisées présentées dans les sections 2.2 et 2.3. Plus précisément, ces travaux ont essentiellement porté sur la segmentation automatique ou semi-automatique des images en régions et la détection d'objets particuliers : bâtiments dans des scènes urbaines, trait de côte dans des scènes du littoral. Ces tâches sont essentielles dans l'approche "orientée-objets", paradigme actuel d'analyse et d'interprétation d'image en télédétection, rendue possible et nécessaire avec l'avènement des capteurs à très haute résolution spatiale.

Les techniques de segmentation ne se focalisent pas sur un type d'objet particulier et permettent d'obtenir une partition complète de l'image. La complexité des images de télédétection limite l'intérêt des approches complètement non-supervisées. Nous avons donc choisi des approches basées sur des connaissances des experts, ces dernières étant fournies aux algorithmes sous la forme de zones exemples (cf. section 2.3). L'utilisation d'informations spatiales, au travers de marqueurs, produit des résultats particulièrement intéressants : ainsi, dans le cadre du projet Ecosgil, elle a permis la détection du trait de côte à partir d'images du littoral [C5], puisque la partition de l'image en deux zones (terre et mer) se prête bien à l'utilisation de marqueurs. Cependant, fournir une information spatiale exhaustive (c'est-à-dire ici une liste complète de marqueurs) n'est pas toujours possible : dans le cas de scènes contenant un grand nombre d'objets à extraire (comme des scènes urbaines), il est nécessaire d'automatiser au moins partiellement le traitement et de limiter l'intervention de l'utilisateur dans l'étape de saisie des marqueurs, comme nous l'avons fait dans le cadre d'une compétition relative à la détection de bâtiments [D27]. Il est également possible de ne pas tenir compte de l'information spatiale portée par les zones exemples, et de n'exploiter que leur contenu spectral ou textural. Ainsi, dans le contexte du projet FoDoMuST, nous nous sommes intéressés à la segmentation d'images urbaines en procédant à des changements d'espace [D14,E13] et en optimisant les paramètres de réduction de sursegmentation [D20] dans le processus de segmentation par ligne de partage des eaux.

L'approche alternative en interprétation d'images satellite à très haute résolution spatiale consiste à remplacer le processus de segmentation-classification (notons que nous nous sommes essentiellement intéressés à la première de ces deux étapes) par une seule étape de détection d'objet. Combiner des détecteurs d'objets permet ainsi d'obtenir un résultat similaire à celui fourni par une classification [Der09]. Dans ce contexte, nous nous sommes intéressés à la détection d'objets spécifiques, tels que les structures bâties dans les scènes urbaines [C4,D17,E9] lors du projet FoDoMuST ou le trait de côte dans les images du littoral [D24,D25] lors du projet Ecosgil. Pour cela nous avons mis en œuvre des solutions basées sur la transformée en tout-ou-rien, que nous avons appliquée de façon omnidirectionnelle sur une décomposition de l'image en différentes composantes (cf. section 2.2.4), celles-ci pouvant être les bandes spectrales de l'image ou des cartes produites par une procédure de classification non-supervisée. Les connaissances, ici une description spatiale et/ou spectrale des objets d'intérêt, ont permis la définition des éléments structurants utilisés lors de l'application de la transformée en tout-ou-rien.

Les conclusions que nous pouvons tirer de nos travaux en télédétection sont les suivantes :

- l'avènement de nouveaux satellites a profondément modifié les usages du traitement d'image (et notamment de la morphologie mathématique) en télédétection : avec l'arrivée de la haute et de la très haute résolution, le paradigme de l'orienté-objets est devenu incontournable [Pui03] et des méthodes d'extraction des objets sont nécessaires ; ces objets peuvent être obtenus suivant deux stratégies alternatives, soit en effectuant une segmentation de l'image en régions (ce qui suppose que les régions soient construites de façon à correspondre aux objets recherchés), soit en utilisant une ou plusieurs méthodes de détection d'objets particuliers [Der09] ;

- pour éviter que l’analyse de chaque nouvelle image ne nécessite un paramétrage spécifique, voire l’élaboration d’une nouvelle méthode *ad hoc*, l’intégration de connaissances sur les objets à identifier est nécessaire ; il est irréaliste de croire que des approches totalement non-supervisées pourront fournir des résultats pertinents quelle que soit la scène étudiée ;
- les connaissances pouvant être introduites par un expert dans un processus d’analyse d’image en télédétection sont de deux types : naturellement il peut s’agir d’une description (géométrique, spectrale, texturale, etc.) relativement exhaustive des objets ou concepts du domaine, et des relations entre ces objets ; cependant, comme nous l’avons observé dans le cadre de la thèse de Sébastien Derivaux [Der09], face à la difficulté qu’ont les experts à formaliser ces connaissances du domaine, il est plus simple de considérer les connaissances disponibles sous la forme d’exemples (pixels ou zones) étiquetés par l’expert ;
- la stratégie la plus utilisée dans le paradigme de l’orienté-objets consiste à segmenter l’image avant d’effectuer une classification des régions obtenues ; si les régions produites par la segmentation sont incorrectes, une méthode de classification, même parfaite, ne pourra donner un résultat acceptable ; il est donc judicieux d’intégrer des connaissances dès l’étape de segmentation, et nous avons observé l’intérêt des méthodes s’appuyant sur le contenu spectral et/ou spatial de zones exemples fournies par l’expert ;
- l’approche alternative à la segmentation est la détection d’objets particuliers, aussi divers que les bâtiments ou le trait de côte par exemple ; ces objets peuvent être extraits à l’aide de la transformée en tout-ou-rien ; dans ce contexte, nous n’avons pas exploité de zones exemples mais plutôt des connaissances du domaine telles qu’une description spatiale (taille, forme) et spectrale des objets à détecter, mais aussi la connaissance du nombre de classes d’intérêt présentes dans l’image ou les relations spatiales entre les objets d’intérêt ;

3.2.5 Perspectives

Profils morphologiques

Comme nous l’avons indiqué dans cette section, l’arrivée de nouveaux capteurs offrant des images à très haute résolution spatiale a modifié en profondeur les usages du traitement d’image en télédétection. L’interprétation d’image peut bénéficier aujourd’hui d’informations de nature très diverses, et non plus se limiter aux réponses spectrales individuelles de chaque pixel. Les approches tenant compte de l’information spatiale (en plus de l’information spectrale) présentent un intérêt évident, ce qui s’est traduit par un regain d’attention porté aux outils morphologiques. Les profils morphologiques, introduits par Pesaresi et Benediktsson [PB01, BPA03], en sont certainement le meilleur exemple, et ils sont très fréquemment utilisés pour décrire le contenu (structurel) des images à très haute résolution spatiale, seuls ou en combinaison avec les informations spectrales [PP08].

Initialement formulés pour des images panchromatiques, ils ont fait l’objet de plusieurs études dans le cas des images à plusieurs composantes. Ainsi, nous nous y sommes intéressés en réalisant une première étude comparative sur les ordres vectoriels [B3], tandis que Plaza et al. [PMPP04] et Benediktsson et al. [BPA05] ont pour leur part utilisé ces profils dans un contexte hyperspectral. Néanmoins, il nous semble que l’adaptation des profils morphologiques aux images multispectrales ou hyperspectrales mérite d’être étudiée de façon plus approfondie. Nous pouvons ainsi nous interroger sur la stratégie à employer (marginale ou vectorielle) et, le cas échéant, sur l’ordre vectoriel à utiliser (la préférence allant aux ordres ayant une signification physique adaptée aux images de télédétection). La stratégie basée sur une décomposition des images (cf. section 2.2) pourrait également être envisagée. Suite à un séminaire donné au laboratoire IPSC⁸, nous avons initié une collaboration avec l’équipe de Martino Pesaresi (qui a été à l’origine des profils morphologiques), notamment Elodie Pagot et Annett Wania, pour étudier la formulation de ces profils dans un contexte d’imagerie multispectrale et hyperspectrale.

Une autre piste à explorer pour combiner l’information spectrale et spatiale au sein des profils morphologiques serait de procéder à une classification spectrale non pas après avoir calculé les profils mais plutôt avant : nous disposerions alors de profils morphologiques supervisés, c’est-à-dire de descripteurs guidés par l’utilisateur. L’intervention de l’expert, initialement située lors de l’étape de classification, et que nous avons introduite

⁸Institute for the Protection and Security of the Citizen — European Commission Joint Research Center (JRC), Ispra (Italie)

lors du processus préalable de segmentation ou de détection, serait alors effectuée tout en amont du processus d'analyse d'image.

Approches multi-sources

Au cours des expérimentations menées dans le cadre des projets FoDoMuST et Ecosgil et décrites dans cette section, nous n'avons pas considéré plusieurs images simultanément : en effet, de par leur conception, les différents traitements (segmentation ou détection) ne pouvaient tenir compte que d'une unique image fournie en entrée. La complexité des scènes étudiées en télédétection, et la grande variété des images offertes par les différents capteurs (satellites ou aéroportés), nécessite d'adopter une approche multi-sources pour tirer profit au mieux des données disponibles.

Contrairement au cas de la classification pour laquelle plusieurs approches adaptées à des données multi-sources ont été proposées dans la littérature [GW07], l'exploitation conjointe de plusieurs sources dans les processus de segmentation ou de détection reste peu étudiée. Elle semble néanmoins pertinente dans le paradigme de l'approche orientée-objets qui est suivi lors de l'analyse d'images à très haute résolution.

La complexité de mise en œuvre des approches multi-sources en segmentation ou en détection d'objets dépend de la nature des différentes sources. Dans le cas le plus simple, il peut s'agir d'images satellites de résolution spatiale identique, obtenues à des dates relativement proches, où la complémentarité des informations porte essentiellement sur la dimension spectrale. Il faut alors procéder à un recalage des images, ou orthorectification, pour pouvoir comparer les données portées par les différentes images, qui peuvent être représentées comme autant de bandes d'une unique image multi-composantes (ainsi, les approches discutées dans cette section sont directement utilisables).

Si les images ont été acquises à des dates différentes, la représentation des objets peut varier d'une image à l'autre (notamment pour les objets naturels, la végétation par exemple, ou suite à des événements de type inondation, tremblement de terre, etc.). Le recalage géométrique est nécessaire mais pas suffisant, il faut également savoir appréhender cette variation temporelle. En télédétection, l'étude de l'évolution temporelle des objets est appelée détection de changements, et elle a fait l'objet de nombreux travaux basés sur des classifications de pixels (donc inadaptés au cas de la très haute résolution), mais de très peu de travaux basés sur une approche objet (notons par exemple l'utilisation des profils morphologiques pour l'analyse des bâtiments dans des situations de conflit [PP07]). En effet, outre les changements spectraux observables en chaque pixel (ou en chaque objet), la forme des objets peut également évoluer au cours du temps. De façon alternative aux solutions proposées dans le contexte de la détection de changement, il nous semble donc pertinent d'envisager ici l'application de techniques de suivi d'objet, plus fréquemment utilisées en analyse de séquences vidéo.

Les images peuvent également représenter une même scène mais à des échelles différentes. Il faut alors savoir appréhender ces différentes représentations et les exploiter conjointement. Ce dernier cas de figure est sans doute le plus délicat à gérer : avec l'amélioration de la résolution, des objets peuvent apparaître dans l'image alors qu'ils n'étaient pas visible à plus faible résolution. Au contraire, des structures composites peuvent être visibles à faible résolution du fait de l'absence de détails, mais difficilement observables à très haute résolution. En outre, la diminution de la résolution provoque un effet connu sous le nom de pixels mixtes : du fait de la trop faible résolution spatiale, chaque pixel représente différents objets au sol, et ses valeurs dans les différentes bandes spectrales sont donc obtenues par composition des signatures spectrales des différents objets qu'il représente. Plus les différences entre échelles sont importantes, plus il sera délicat de traiter conjointement les différentes images (la figure 3.8 en donne une illustration en présentant une même scène à des résolutions différentes).

Quelle que soit la forme que peuvent prendre les données multi-sources (multi-capteurs, multi-échelles, multi-dates), nous pouvons nous interroger sur l'adaptation des processus de segmentation ou de détection au cas de plusieurs images. Au vu de nos expérimentations en télédétection, cette adaptation est certainement plus aisée dans le cas des approches par détection. En effet, nous avons déjà considéré pour l'extraction du trait de côte un ensemble d'éléments structurants de différentes formes et/ou tailles, chacun d'entre eux étant associé à une bande spectrale donnée. L'approche proposée dans [D24, D25] devrait donc être aisément transposable à des données multi-sources.

Le processus de segmentation, pour sa part, doit certainement être revu pour tenir compte de données multi-sources. Dans le cas où les images utilisées sont de résolution ou d'échelle identique, nous avons identifié lors du projet FoDoMuST trois stratégies pouvant être mises en œuvre : fusionner les images pour n'obtenir qu'une seule image sur laquelle appliquer la segmentation, appliquer une segmentation sur chacune des images et fusionner les résultats (cartes de segmentation) obtenus, et enfin tenir compte des différentes images directement au sein du processus de segmentation. Cette dernière approche offre l'avantage d'éviter d'introduire une étape de fusion qui se traduira nécessairement par une perte d'information, mais elle est la plus délicate à réaliser. Néanmoins la méthode de segmentation avec marqueurs spatiaux et spectraux que nous avons élaborée [D19, E16] exploite d'ores et déjà cette idée, il faudrait alors s'affranchir de l'utilisation des marqueurs pour obtenir un processus générique de segmentation multi-image.

Si les échelles associées aux différentes images à traiter sont différentes, d'autres stratégies doivent être envisagées. Ainsi, il est possible de segmenter tout d'abord l'image à l'échelle la moins précise, et de guider les segmentations aux résolutions plus fines par le résultat obtenu (en imposant par exemple des contraintes sur les frontières des objets). Notons que cette approche est utilisée dans certains logiciels commerciaux comme eCognition où il est possible de contraindre une segmentation par une carte thématique (cette dernière pouvant être le résultat d'une autre segmentation). L'ordre des images à segmenter peut être inversé, l'imposition des contraintes étant alors remplacé par un processus de fusion de régions. Cette stratégie de segmentations successives présente cependant l'inconvénient d'une trop forte importance donnée aux premières images analysées dans la segmentation de l'ensemble des images. Une approche par segmentation conjointe, tenant compte simultanément des différentes échelles dans le processus de segmentation semble plus pertinente. Adaptée au cas de la ligne de partage des eaux, elle nécessite d'inonder les reliefs associées aux différentes échelles de façon parallèle, comme nous l'avons déjà fait pour la segmentation supervisée avec marqueurs. Toutefois, il faut ici tenir compte du fait que les reliefs à inonder ne correspondent pas à la même résolution.

3.3 Imagerie astronomique

Cette section traite des applications de la morphologie mathématique au domaine de l'imagerie astronomique. Ces travaux se sont déroulés en étroite collaboration avec Christophe Collet, responsable du groupe PASEO (Probabilistic Analysis for Space and Earth Observation) avec qui j'ai notamment encadré deux stages de master recherche (Erchan Aptoula en 2005, Benjamin Perret en 2007). Ils se poursuivent actuellement avec la thèse de Benjamin Perret, sous la responsabilité de Christophe Collet et d'Eric Slezak (Observatoire de Nice – Côte d'Azur), et dont j'assure une partie de l'encadrement avec Vincent Mazet (Maître de Conférences dans l'équipe MIV, et spécialiste des problèmes inverses en traitement du signal et des images et de leur résolution par inférence bayésienne). Tous ces travaux ont également fait l'objet de partenariats avec des astronomes, principalement de l'Observatoire Astronomique de Strasbourg et de l'Observatoire de Nice – Côte d'Azur.

Après une rapide introduction permettant de rappeler nos motivations, nous présenterons les travaux que nous avons menés pour résoudre les problèmes de la segmentation, la détection, et la classification de galaxies qui ont fait l'objet de plusieurs publications en revue [B9] ou en conférences internationales [D15, D26]. Nous indiquerons enfin les conclusions que nous pouvons tirer de ces travaux et les suites que nous souhaiterions y donner.

3.3.1 Introduction

Les approches morphologiques restent aujourd'hui très peu utilisées en imagerie astronomique. En outre, il s'agit le plus souvent de méthodes morphologiques déjà anciennes, et non adaptées aux images multispectrales. Parmi les applications morphologiques en imagerie astronomique, nous pouvons citer les travaux en compression (notamment au travers de méthodes de squelettisation [BBH96, HB91]) ou en réduction de bruit (avec des filtres tels qu'ouverture et fermeture [LK89, LLH94, Mac96, MTD92]). Plus proches de nos travaux en segmentation, détection et classification de galaxies, notons l'utilisation du spectre morphologique pour aider à la classification étoiles / galaxies [CBNF97] ou encore la segmentation d'images astronomiques contenant plusieurs objets, au moyen d'une technique de ligne de partage des eaux suivie par une fusion de région [FL04].

Pourquoi les approches morphologiques sont-elles si peu exploitées dans le domaine de l'imagerie astronomique ? La principale raison est certainement relative à la nature même de ces images et des objets qui peuvent y être observés, comme l'illustre les images de la figure 3.18. Nous pouvons constater sur ces images que les objets astronomiques sont diffus. De plus la taille des observations est importante et le niveau de bruit souvent élevé (figure 3.19). C'est pourquoi les approches statistiques (telles que celles développées dans le groupe PA-SEO) sont les plus fréquemment utilisées en imagerie astronomique, au détriment par exemple des approches morphologiques.



FIG. 3.18 – Images astronomiques : (de gauche à droite) galaxie *whirlpool* M51 / NGC 5194, galaxie d'Andromède M31 / NGC 224, et deux représentations de la nébuleuse d'Orion M42 / NGC 1976.

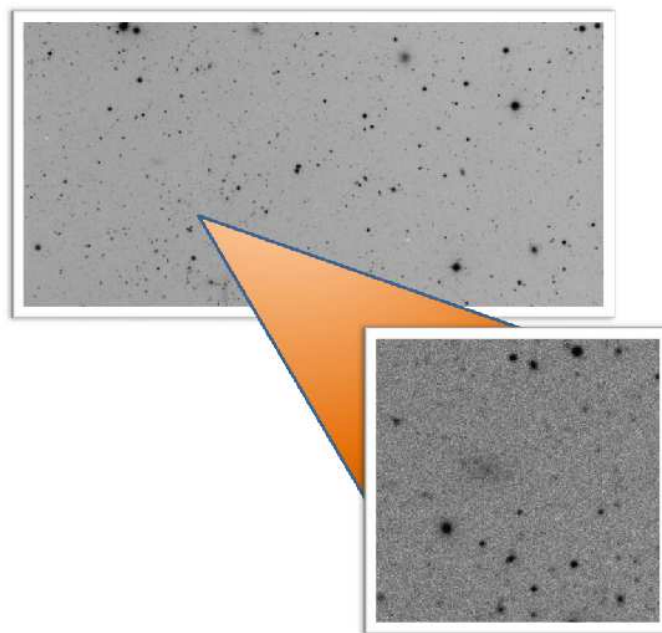


FIG. 3.19 – Exemple d'une planche de 4096×2048 pixels représentant l'amas de la Vierge (observation composée de 80 planches).

Pourtant les avancées récentes en morphologie mathématique permettraient de résoudre certains problèmes en imagerie astronomique. Pour expliquer cette situation, nous pouvons avancer une raison supplémentaire que nous avons déjà soulignée précédemment : la difficulté rencontrée par la communauté de morphologie mathématique à disséminer correctement les fruits de ses travaux dans les différents domaines utilisateur du traitement d'image (cf. par exemple la couverture du sujet dans un des ouvrages de référence, le *Gonzalez et Woods* [GW08]). Ce constat motive nos travaux en imagerie astronomique, dans le but de démontrer le potentiel

des approches morphologiques et de renforcer leur utilisation.

Pour cela, nous nous sommes intéressés à des problèmes déjà traités dans d'autres domaines (cf. sections 3.1 et 3.2) et également présents en imagerie astronomique. Nous avons ainsi étudié la segmentation, la détection et la classification de galaxies dans le but de fournir des outils morphologiques d'aide à l'interprétation d'images astronomiques.

3.3.2 Segmentation de galaxies

La segmentation de galaxies n'est généralement pas un but en soi. Elle a plutôt pour objectif de cerner correctement la zone d'analyse correspondant à l'objet (la galaxie) en cours d'étude. Ainsi la description de la galaxie par différents critères quantitatifs sera plus pertinente, et sa classification plus fiable. Nous avons donc cherché à résoudre le problème de la segmentation d'images astronomiques multi-bandes, dont un exemple est donné en figure 3.20. Nous pouvons observer sur ces imagerie (de taille 128×128 pixels) la faible résolution spatiale des galaxies ainsi que l'hétérogénéité des informations apportées par les différentes bandes spectrales, ce qui rend particulièrement délicat le processus de segmentation. En effet, le niveau de bruit est variable d'une bande à l'autre, et il est possible que l'objet à segmenter ne soit pas observable à certaines longueurs d'onde (il est alors absent des images correspondantes).

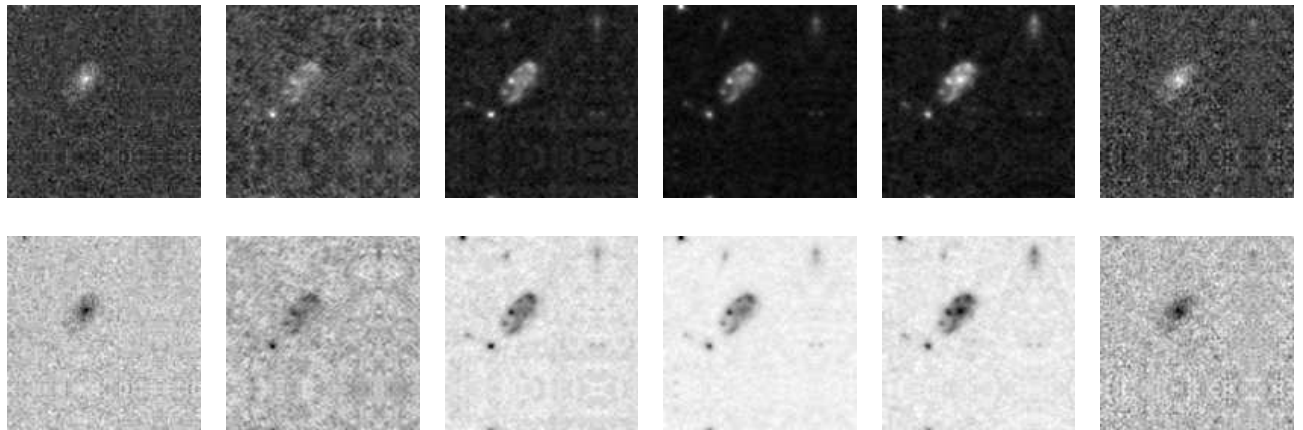


FIG. 3.20 – Représentation multispectrale de la Galaxie HDF4-378 (de gauche à droite : 160, 300, 450, 606, 814 et 1100 nm) en utilisant l'échelle originale des niveaux de gris (haut) et son inverse permettant une meilleure visualisation des faibles intensités (bas).

La méthode que nous avons élaborée est constituée de deux étapes [Apt05] : un prétraitement est tout d'abord effectué pour réduire le bruit présent dans les images, avant de procéder à la segmentation proprement dite. Nous adoptons ici une stratégie marginale (cf. section 2.2.2), évitant ainsi le choix délicat d'un ordre vectoriel, au prix toutefois des deux inconvénients majeurs des approches marginales. D'une part, l'utilisation d'un ordre partiel tel qu'induit par une stratégie marginale ne garantit pas la préservation des vecteurs d'origine. Cependant, puisque notre but ici est la segmentation de l'image (et non le filtrage par exemple), c'est-à-dire la partition de l'image en différentes régions, il n'est pas nécessaire de préserver les valeurs d'origine. D'autre part, les approches marginales traitent chaque bande indépendamment les unes des autres. Elles ignorent donc les informations corrélées entre les différentes bandes. Ce problème est plus critique ici, et nous avons fait le choix d'effectuer une fusion des images à l'issue de l'étape de réduction de bruit. Ces choix s'expliquent notamment par la meilleure capacité des approches marginales à réduire le bruit et la difficulté à appliquer la méthode de segmentation par ligne de partage des eaux sur un relief multibandes (notons que, par la suite, nous nous sommes intéressés à de telles méthodes, cf. en particulier [D19]). Puisque seule une galaxie est présente dans chaque image, l'objectif est d'identifier les bords de celles-ci ou autrement dit les frontières entre la galaxie et le fond. Nous avons donc utilisé une approche de segmentation avec marqueurs, en associant un marqueur interne à la galaxie et un marqueur externe au fond, ce qui permet en outre d'éviter le problème de sursegmentation. La définition de ces deux marqueurs est faite à partir d'un traitement itératif sur chaque bande spectrale en s'appuyant sur des hypothèses quant à l'intensité moyenne de la bande et les propriétés (taille,

position) de l'objet détecté après binarisation. Afin d'obtenir un seul couple de marqueurs (interne / externe), nous fusionnons les marqueurs obtenus dans les différentes bandes spectrales par des opérations de minimum et maximum. Le gradient multispectral est quant à lui calculé comme un maximum des gradients des différentes bandes (on limite ainsi la perte d'information, au prix d'un niveau de bruit plus important mais non critique).

Ces travaux, parmi les premiers que nous avons menés en segmentation d'images multispectrales, ont fourni des résultats intéressants (figure 3.21) et fait l'objet d'une communication à EUSIPCO [D15]. Ils ont également permis de convaincre nos partenaires astronomes de l'intérêt des approches morphologiques, et ainsi motivé des projets complémentaires en détection et classification de galaxies.

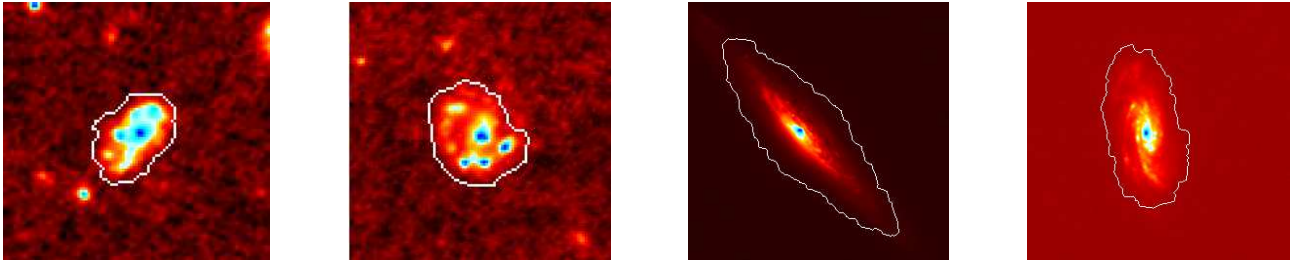


FIG. 3.21 – Quelques résultats en segmentation d'images astronomiques.

3.3.3 Détection de galaxies

La détection de galaxies consiste à localiser les zones de l'image contenant des objets d'intérêt astronomiques. Elle peut donc être obtenue au travers d'une segmentation telle que décrite dans la section précédente. Cependant nous considérons ici le cas d'objets difficilement visibles appelés galaxies à faible brillance de surface (ou galaxies LSB – *low surface brightness*). Pour ce type d'objet, la méthode de segmentation décrite précédemment n'est pas adaptée (puisque prévue pour les galaxies HSB – *high surface brightness*). La figure 3.22 présente deux exemples de galaxies HSB et LSB. On peut y constater le faible rapport signal sur bruit des galaxies LSB (inférieur à 0 dB⁹), ce qui rend l'utilisation de techniques morphologiques particulièrement délicate. Nous avons néanmoins proposé une telle approche pour la détection de ces objets dans des images mono-bandes [Per07].

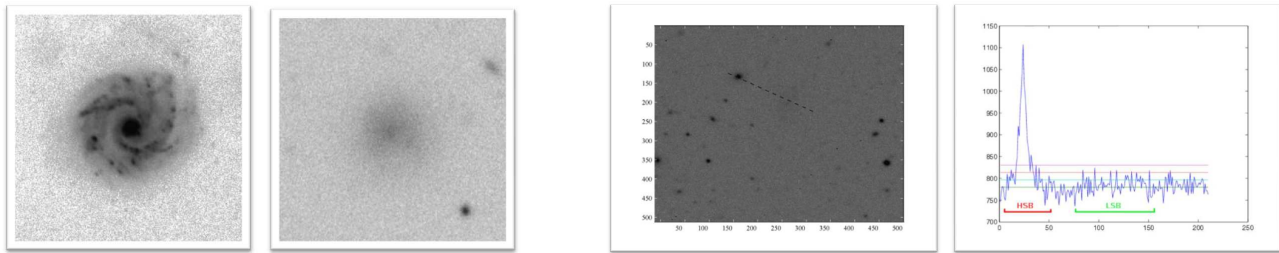


FIG. 3.22 – De gauche à droite, galaxies à forte brillance de surface (HSB) et à faible brillance de surface (LSB), image contenant ces deux types de galaxies et profil correspondant (*les images astronomiques sont affichées en couleurs inversées*).

Les galaxies LSB, même noyées dans le bruit, ont une forme et une intensité caractéristique : il s'agit d'objets elliptiques au profil de luminosité proche d'une décroissance exponentielle (le pic d'intensité étant situé au centre de la galaxie). De plus, le niveau de bruit de l'image peut être estimé, de préférence localement pour tenir compte de ses variations dans l'image et limiter le temps de calcul. Nous exploitons ces connaissances au sein d'un opérateur morphologique de mise en correspondance, la transformée en tout-ou-rien. Cependant la formulation de cette dernière (équation (2.1.28)) n'est pas adaptée aux images bruitées puisqu'il sera impossible

⁹le rapport signal sur bruit ou *signal to noise ratio* est défini par $SNR = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_n^2} \right)$ dB, avec σ_s et σ_n les écarts-types du signal et du bruit, respectivement.

d'obtenir une correspondance parfaite entre un motif théorique elliptique au profil d'intensité exponentielle et la réalité observée dans un environnement particulièrement bruité. Nous avons donc élaboré un nouvel opérateur de mise en correspondance inspiré des travaux de Maragos pour les images binaires [Mar88]. Dans la transformée en tout-ou-rien robuste que nous avons proposée, et contrairement à la formulation initiale, l'étape d'ajustage (cf. section 2.1.4) ne fournit pas un résultat binaire (c'est-à-dire avec une réponse de type : *il y a* ou *il n'y a pas* correspondance) mais plutôt un ratio mesurant la qualité de la correspondance obtenue :

$$\eta_{b_1, b_2}(f)(\mathbf{p}) = \max_{h \in \mathcal{T}} \frac{\text{card}\{\mathbf{q} \in \mathcal{D}_{b_1, b_2} \mid b_1(\mathbf{q}) + h \leq f(\mathbf{p} + \mathbf{q}) \leq b_2(\mathbf{q}) + h\}}{\text{card}(\mathcal{D}_{b_1, b_2})} \quad (3.3.1)$$

avec $\mathcal{D}_{b_1, b_2} = \mathcal{D}_{b_1} \cup \mathcal{D}_{b_2}$ l'union des supports des deux fonctions structurantes et la variable h représentant les différents niveaux de gris. Il faut donc construire correctement les motifs à rechercher, en considérant différentes tailles, élongations, orientations d'ellipses. Différentes optimisations permettent d'assurer une détection en 2 minutes pour une image de taille 512×512 pixels, dont un résultat est donné en figure 3.23.

Les résultats obtenus sont intéressants et cette méthode est actuellement utilisée par des astronomes pour identifier de nouvelles galaxies LSB potentielles qu'il leur faut ensuite étudier plus finement. En outre, nous avons publié ces travaux dans la conférence EUSIPCO [D26] et la revue *Pattern Recognition* [B9].

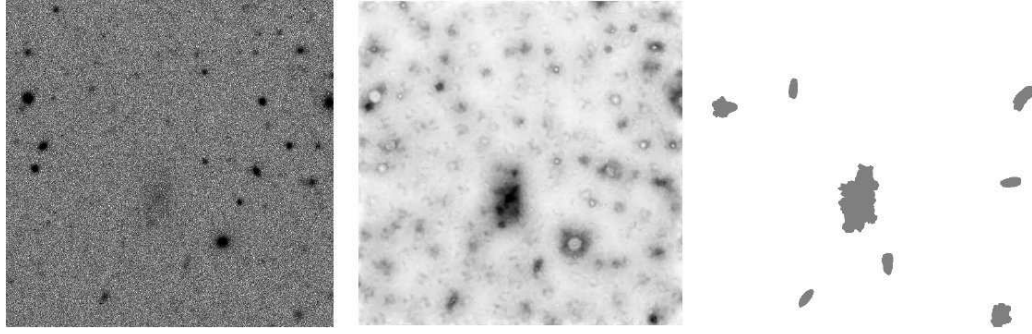


FIG. 3.23 – De gauche à droite, image initiale, résultat fourni par la transformée en tout-ou-rien robuste, et seuillage final pour identifier les galaxies LSB potentielles.

3.3.4 Classification de galaxies

Le dernier problème que nous avons étudié en imagerie astronomique concerne la classification de galaxies. Ici la morphologie mathématique n'est pas utilisée directement pour effectuer la classification, mais plutôt pour fournir une description pertinente des objets à classer. L'objectif est d'aboutir à des solutions de classification automatique respectant la proposition de Hubble (figure 3.24)



FIG. 3.24 – De gauche à droite : classification de Hubble et exemples de galaxies elliptique et spirale.

Suite à la segmentation de galaxies (cf. section 3.3.2), nos premières tentatives pour la classification de galaxies ont porté sur la distinction entre galaxies elliptiques et spirales illustrées en figure 3.24. En considérant que les jeunes étoiles (les plus brillantes) sont fréquentes dans les galaxies spirales qui contiennent de nombreux sites de formation stellaire (tandis que les étoiles vieilles et froides sont dominantes dans les galaxies

elliptiques), nous avons procédé à une application marginale de l'opérateur du chapeau haut-de-forme afin d'obtenir une mesure du nombre de sites de formation stellaire qui soit invariante du point de vue. L'utilisation de cet unique attribut fournit des taux de classification de 80 % environ [D15]. L'utilisation de descripteurs morphologiques tels que la covariance permet d'améliorer la classification obtenue.

Pour obtenir une classification plus précise (par exemple distinguer les différentes classes de galaxie elliptique ou spirale), les critères morphologiques sont sans doute insuffisants. Aussi, dans le cadre de la thèse de Benjamin Perret, ils sont couplés avec des approches statistiques (par exemple des algorithmes de Monte Carlo par chaînes de Markov – MCMC [PMCS09]). La classification s'appuie sur une modélisation paramétrique très fine des galaxies et de leurs principaux composants (bulbe, disque, bras spiraux, etc). Nous utilisons ici la morphologie mathématique au travers des arbres des coupes (cf. section 2.1.3). Ils nous permettent de résoudre différents problèmes [Per09] tels que la séparation de sources (pour isoler la galaxies des objets parasites), l'initialisation déterministe du modèle paramétrique par extraction de la composante principale de la galaxie, la localisation des zones de formation stellaire et des bras spiraux, etc. De plus, la nature bruitée des images astronomiques nous amène à travailler sur les notions plus générales d'hyperconnexité [Ser98] et d'arbres en composantes hyperconnexes. Enfin, nous étudions l'adaptation de ces outils aux images multispectrales, notamment à l'aide d'une variante de l'ordre lexicographique tenant compte de l'énergie présente dans les différentes bandes.

3.3.5 Conclusion

Dans cette section, nous avons décrit nos travaux en imagerie astronomique. Les problèmes classiques de traitement d'image (segmentation, détection, description et classification par exemple), déjà abordés dans d'autres sections, sont également présents dans ce domaine. Cependant les images astronomiques et les objets qu'elles contiennent se caractérisent par certaines spécificités qu'il nous a fallu étudier : un niveau de bruit important, l'absence de frontières nettes des objets et leur caractère diffus, la possible répartition des informations au sein de plusieurs bandes spectrales.

La principale difficulté est certainement liée au niveau de bruit présent dans les images. Lorsque cela est encore possible, l'utilisation systématique de techniques de filtrage permet d'obtenir des images de qualité suffisante pour les traitements à venir (segmentation ou classification par exemple). Dans ce contexte, nous avons observé qu'il était préférable de suivre une stratégie marginale plutôt que vectorielle dans le cas d'images multispectrales (même si nous nous sommes également intéressés à des ordres vectoriels spécifiques aux images astronomiques, tenant compte par exemple du rapport signal / bruit ou de l'énergie contenue dans chaque bande). Comme indiqué dans la section 2.2, cette stratégie marginale se traduit néanmoins par une non préservation des valeurs spectrales originales et l'absence de prise en compte des informations corrélées entre les bandes. Tandis qu'il n'est pas ici nécessaire de préserver les valeurs originales (notre but final étant la détection ou la classification, mais pas le filtrage ou la restauration par exemple), nous avons choisi de procéder à des fusions de bandes pour réduire les images multispectrales (initiales ou obtenues après traitement marginal) à des images à niveaux de gris, et ainsi combiner les informations présentes dans les différentes bandes spectrales. Notons que l'étape de fusion se caractérisant par une perte d'information, sa position dans le processus global d'analyse ou de traitement d'image influe sur la qualité du résultat final (par exemple, en segmentation d'image, elle a directement précédé l'application de l'opérateur de ligne de partage des eaux).

Dans certains cas le niveau de bruit est trop important pour pouvoir être appréhendé au travers de méthodes de filtrage ou de restauration sans dégradation irréversible du résultat final. Il est alors préférable d'intégrer directement la gestion du bruit au sein de l'opérateur morphologique. Ainsi, nous avons élaboré une approche de mise en correspondance par transformée en tout-ou-rien robuste que nous avons utilisée pour la détection de galaxies à faible brillance de surface. Contrairement aux formulations originales, elle s'intéresse à la qualité d'une possible mise en correspondance sans chercher à déterminer s'il y a réellement ou non une correspondance. La décision finale revient alors à l'utilisateur qui peut ajuster par seuillage le résultat obtenu. En outre, il est possible de tenir compte explicitement du niveau de bruit lors de la formulation de cet opérateur. Pour résoudre un plus large panel de problèmes, nous nous intéressons aussi à la formulation d'opérateurs robustes issus d'une représentation des images sous forme d'arbres des coupes. Pour traiter correctement le bruit présent dans les images, nous étudions la notion de composantes hyperconnexes, évitant ainsi la contrainte de zones

plates (ou niveaux de gris constants) imposée dans la définition des composantes connexes. Cette représentation en arbres de composantes hyperconnexes va également nous permettre de mieux appréhender la nature des objets astronomiques, et notamment leur caractère diffus [Per09].

En résumé, les conclusions que nous pouvons tirer de nos travaux en imagerie astronomique sont les suivantes :

- pour tenir compte du niveau de bruit important dans les images astronomiques, il est possible d’employer des techniques de filtrage ou d’intégrer la gestion du bruit directement au sein des opérateurs morphologiques ;
- lorsque la réduction de bruit est obtenue par filtrage d’images multispectrales, les approches marginales sont à privilégier [D15] ; en l’absence de prise en compte explicite de la corrélation entre les bandes, il est alors nécessaire de fusionner les informations contenues dans les différentes bandes, et l’étape de fusion doit être judicieusement positionnée dans le processus global d’analyse d’image ;
- lorsque le niveau de bruit est trop important, l’élaboration de techniques morphologiques robustes au bruit est nécessaire ; dans ce contexte, nous avons proposé une transformée en tout-ou-rien robuste pour la détection des galaxies à faible brillance de surface [B9,D26], et travaillons actuellement sur le concept d’arbre en composantes hyperconnexes pour identifier les zones de formation stellaire, les bras spiraux, ou encore d’autres structures internes des galaxies en vue de leur classification.

3.3.6 Perspectives

Images astronomiques multispectrales et hyperspectrales

Les travaux décrits dans cette section se poursuivent actuellement au travers de la thèse de Benjamin Perret débutée en 2007. Nous nous intéressons en particulier à la notion d’arbres de composantes hyperconnexes et à leur application à des images astronomiques multispectrales. Au vu des utilisations existantes des arbres de composantes connexes, nous pensons qu’une telle représentation, tenant compte à la fois de l’aspect multispectral et du caractère diffus et bruité des observations, est particulièrement pertinente pour l’analyse d’images astronomiques [Per09].

Outre ces travaux en cours, l’application des approches morphologiques aux images astronomiques est source de questions qu’il nous semble important de traiter. Ainsi, la formulation de nouveaux opérateurs (transformée en tout-ou-rien, arbres des coupes), voire également l’utilisation de techniques existantes (ligne de partage des eaux) dans le domaine de l’imagerie astronomique pose la question de la compatibilité de ces techniques avec les images multispectrales (cf. section 2.2). Nous avons ici choisi essentiellement d’appliquer une stratégie marginale suivie d’une étape de fusion. Cette stratégie a ses limites, et d’autres alternatives doivent être envisagées. Le choix d’une stratégie vectorielle nécessite la définition d’un ordre pertinent dans le contexte astronomique (nous nous sommes ainsi intéressés aux ordres lexicographiques où la priorité des bandes est définie selon leur rapport signal / bruit ou l’énergie qu’elles contiennent). En outre, des concepts tels que la représentation en arbre de composantes (hyper)connexes ou la technique de segmentation par ligne de partage des eaux doivent être étudiés dans le contexte d’images multispectrales : qu’est-ce qu’une composante connexe multispectrale ? comment définir la ligne de partage des eaux d’un relief multispectral ?

Ces questions sont encore plus importantes si l’on s’intéresse aux images hyperspectrales également présentes en imagerie astronomique. Dans de telles images, les bandes sont bien plus nombreuses (d’une dizaine à plusieurs centaines) et couvrent le spectre de façon plus fine. De plus certaines bandes sont souvent inutilisables, rendant l’utilisation d’une stratégie marginale couplée à une étape de fusion complètement irréaliste. Il est alors nécessaire de procéder directement au traitement morphologique des images hyperspectrales, c’est-à-dire à la détection, la segmentation, ou la description d’objets représentés par des dizaines ou des centaines de bandes spectrales. L’application de la morphologie mathématique à de telles images est extrêmement rare, un des travaux représentatifs dans ce domaine et très proche de nos préoccupations (segmentation par LPE, composantes connexes, etc) est la thèse de Noyel [Noy08] mais celle-ci ne porte cependant pas sur l’imagerie astronomique. Suivant le schéma classique en imagerie hyperspectrale (filtrage, réduction de dimension, puis classification ou segmentation – ici segmentation basée sur une classification), elle propose des développements intéressants pour combiner la dimension spatiale induite par les opérateurs morphologiques avec la

nature spectrale des données analysées. Une autre approche consisterait à appréhender l'image hyperspectrale astronomique comme un cube de données, et ainsi à utiliser les opérateurs morphologiques définis pour des volumes (et utilisés par exemple en imagerie biologique ou médicale 3-D) et traiter la dimension spectrale de façon spatiale. Cet axe de recherche permettrait une mise en application des questions posées quant à l'analyse d'images hypervaluées (cf. section 2.2.6).

Une autre voie à explorer concerne l'utilisation en imagerie astronomique d'approches morphologiques développées pour d'autres types d'images multispectrales. En particulier, il serait intéressant d'appliquer les différents traitements élaborés en télédétection (section 3.2) au cas des images astronomiques. L'utilisation d'approches de segmentation ou de détection morphologique intégrant une classification spectrale (supervisée ou non) est un choix très pertinent en imagerie astronomique, mais les méthodes introduites en télédétection supposent des frontières entre objets relativement nettes, ce qui est rarement le cas pour les objets astronomiques. Nous envisageons également d'étudier l'intérêt des profils morphologiques pour la description multi-échelle et la caractérisation du contenu des images astronomiques. En outre, ces profils pourraient être calculés efficacement à partir des représentations en arbres de composantes (hyper)connexes.

Morphologie statistique

La collaboration avec le groupe PASEO initiée par ces projets en imagerie astronomique ouvre également des perspectives quant au lien entre morphologie mathématique et cadre probabiliste. Comme nous l'avons observé précédemment, la définition d'opérateurs morphologiques robustes nécessite de faire appel à des outils statistiques (même si les usages que nous en avons fait restent encore limités). Les approches morphologiques et statistiques étant complémentaires, il semble intéressant de les combiner pour tenir compte au mieux des informations spatiales et spectrales, modéliser plus finement les a priori disponibles, assurer une robustesse au bruit et disposer d'algorithmes efficaces. Cette combinaison peut s'effectuer de différentes façons :

- appliquer des opérateurs statistiques sur le résultat de traitement morphologiques (Patras et al. [PHL01] effectuent une présegmentation par ligne de partage des eaux qu'ils affinent ensuite à l'aide d'une modélisation par champs de Markov) ;
- effectuer une modélisation statistique des traitements morphologiques (par exemple les formulations du *watersnake* de Nguyen et al. [NWvdB03] ou du *stochastic watershed* d'Angulo et Jeulin [AJ07]) ;
- baser les traitements morphologiques sur des résultats obtenus de façon probabiliste (Lezoray et Carot [LC02] utilisent une classification probabiliste en tant que prétraitement d'une segmentation par ligne de partage des eaux pour obtenir les marqueurs à utiliser, et des mesures statistiques permettant de guider la segmentation) ;
- considérer les traitements morphologiques comme des étapes d'un processus statistique global (Forbes et Raftery [FR99] proposent de remplacer l'algorithme ICM par des opérations morphologiques pour réduire le temps de calcul de la restauration d'image dans un cadre Bayésien).

Dans le contexte astronomique, nous nous intéressons à cette dernière approche dans le cadre de la classification de galaxie au travers d'une modélisation explicite de ses composants [PMCS09], en définissant certains éléments du modèle de galaxie comme des filtres morphologiques (associés à une représentation en arbres des coupes). Les paramètres de ces filtres font donc partie des variables à optimiser.

3.4 PELICAN : une plate-forme pour l'analyse et le traitement des images

En complément des sections précédentes dédiées à nos travaux fondamentaux et applicatifs en morphologie mathématique, nous présentons ici une plate-forme pour l'analyse et le traitement des images numériques. Celle-ci est utilisée par différents chercheurs au sein du laboratoire pour concevoir, expérimenter, ou utiliser les méthodes de traitement d'image auxquelles nous nous intéressons. Elle ne se limite pas au cadre morphologique, et certains de ses principes nous semblent suffisamment pertinents pour justifier une présentation dans ce mémoire. L'implantation de cette plate-forme a nécessité l'implication de plusieurs personnes, notamment Florent Sollier qui a réalisé différentes interfaces du système et une procédure de conformité dans le cadre de son master en apprentissage, et plusieurs étudiants inscrits en master ou en doctorat et/ou proches collabora-

teurs (pour n'en citer que quelques-uns : Erchan Aptoula, Jonathan Weber, Benjamin Perret, Régis Witz, et il y a plus longtemps Loïc Fortchantre et Abdelhak Mokthari).

Nous débuterons cette section par une introduction permettant de positionner notre proposition par rapport aux systèmes existants, et précisant nos motivations et nos contributions. Nous présenterons ensuite l'architecture et les principales fonctionnalités de notre système intitulé PELICAN. Nous dresserons ensuite une liste des usages possibles de PELICAN mis en œuvre au travers de différentes interfaces. Cette section se terminera, comme les précédentes, par des conclusions et perspectives.

3.4.1 Introduction

La conception d'une solution d'analyse et de traitement d'image peut s'avérer fastidieuse. Il faut bien sûr définir les algorithmes à utiliser, qu'il s'agisse de contributions existantes ou de travaux élaborés pour l'occasion, puis assembler les éléments adéquats pour former un système complexe. La réflexion menée en amont sur les méthodes de traitement d'image à employer pour résoudre un problème donné, bien que primordiale, n'est pas suffisante. En effet, une fois ces méthodes choisies, leur intégration dans un système d'analyse et de traitement d'image nécessite de disposer d'un environnement logiciel adéquat pour produire l'outil final qui sera fourni à l'utilisateur.

Il existe aujourd'hui différents logiciels permettant de concevoir des solutions d'analyse et de traitements d'images, adaptés à différents besoins et usages. Les logiciels en ligne de commande (comme *Imagemagick*¹⁰) sont utiles lorsqu'aucune interaction avec l'utilisateur n'est nécessaire, permettent d'éviter un affichage des images à l'écran, et sont adaptés au traitement des images par lots. Les environnements de développement (comme *Matlab*¹¹) permettent d'élaborer rapidement des traitements complexes et de réaliser des prototypes démontrant l'intérêt de la solution adoptée : ce sont des scripts écrits dans un langage propriétaire interprété (offrant des instructions et structures de données de niveau intermédiaire). Les bibliothèques d'opérateurs (comme *JAI*¹² pour le langage Java, *ITK*¹³ et *OpenCV*¹⁴ pour les langages C/C++) s'adressent aux développeurs informatiques et offrent beaucoup plus de possibilités, au prix néanmoins d'un investissement de temps plus important pour réaliser la solution : celle-ci nécessite l'écriture d'un programme dans un langage compilé standard, n'offrant initialement que des instructions de bas-niveau. Les logiciels de manipulation d'image tels que *Photoshop*¹⁵ ou *The Gimp*¹⁶ offrent une interface plus intuitive, accessible au grand public : adaptés au traitement rapide d'une seule image et à la conception empirique d'une solution de traitement d'image, leur usage est plus délicat en cas d'automatisation (plusieurs images à traiter, ou plusieurs traitements à effectuer). Notons que certains logiciels de cette catégorie offrent des capacités poussées d'analyse d'image (par exemple *ImageJ*¹⁷ fortement utilisé en imagerie biologique). Enfin, les logiciels de programmation visuelle (par exemple *Neat-Vision*¹⁸) permettent de concevoir et d'exécuter des traitements complexes sans imposer la connaissance d'un langage de programmation : l'utilisateur exploite des éléments graphiques (opérateurs de traitement d'image, connexions entre ces opérateurs) pour créer la solution à son problème d'analyse et de traitement d'image.

La grande majorité de ces systèmes sont destinés à un type d'utilisateur particulier (le traiteur ou l'analyste d'image, le détenteurs de fonds d'images, le développeur d'algorithmes, etc.). De plus, l'ajout de fonctionnalités peut dans certains cas s'avérer délicate, notamment du fait du lien trop fort entre l'interface utilisateur et les algorithmes de traitement d'image proposés par le système. Enfin, un nombre important de logiciels ne sont dédiés qu'à des images représentées comme un tableau 2-D ou 3-D, voire 4-D (les dimensions pouvant être spatiale, temporelle, spectrale) de données, au type prédéfini (entier 8-bits par exemple). Ils peuvent être spécialisés dans un domaine particulier (imagerie biologique par exemple) avec des traitements dédiés à ce

¹⁰<http://www.imagemagick.org>

¹¹<http://www.mathworks.fr>

¹²<http://java.sun.com/products/java-media/jai>

¹³<http://www.itk.org>

¹⁴<http://opencv.willowgarage.com>

¹⁵<http://www.adobe.com/fr/products/photoshop>

¹⁶<http://www.gimp-fr.org>

¹⁷<http://rsb.info.nih.gov/ij>

¹⁸<http://neatvision.eeng.dcu.ie>

domaine. Rares sont les systèmes commerciaux qui ne présentent pas au moins une des limites mentionnées ci-dessus.

Or les différentes collaborations que nous avons entreprises, les différents projets dans lesquels nous nous sommes impliqués nous ont amené à appréhender des images de nature différente, dans des domaines d'application variés, et à proposer des méthodes d'analyse et de traitement d'images destinées à des utilisateurs aux profils bien différents. Il est bien évidemment possible d'utiliser un environnement spécifique pour chaque application résultant d'un nouveau projet ou d'une nouvelle collaboration, mais nous n'avons pas opté pour cette stratégie qui va à l'encontre du principe de réutilisation et empêche de bénéficier au mieux de l'expérience acquise dans le domaine de l'analyse et le traitement des images. Au contraire, nous avons observé à quel point il pouvait être pertinent de disposer d'une plate-forme permettant l'élaboration, l'utilisation et la réutilisation d'algorithmes et méthodes génériques en traitement d'image, à partir desquels peuvent être conçues des solutions aux problèmes posés par nos partenaires dans le cadre de différents projets.

Ce constat a inspiré la conception et la réalisation d'une nouvelle plate-forme intitulée PELICAN (Polyvalent Extensible Library for Image Computing and ANalysis). Celle-ci permet de traiter de façon similaire ou non des données variées : images 2-D à niveaux de gris, couleur, multi- ou hyperspectrales, images 3-D volumiques, séquences vidéo, etc. Elle est destinée à un large panel d'utilisateurs, et peut être très facilement enrichie de nouveaux algorithmes ou de nouvelles interfaces utilisateurs. Ces deux propriétés, la polyvalence (de profils d'utilisateurs, de données traitées, de domaines considérés) et l'extensibilité sont les principes fondateurs de notre projet et les éléments qui le différencie le plus de l'existant. L'élaboration de cette plate-forme nécessite des efforts importants en développement et ingénierie logicielle, mais nous pensons qu'il s'agit là d'une contribution aux thèmes de recherche du groupe de travail SCATI¹⁹, qui s'intéresse à la notion de systèmes pour le traitement et l'analyse d'image, en faisant appel aux outils et méthodes relevant du traitement d'image, de la fusion de données, de l'intelligence artificielle, de l'ingénierie de la connaissance, de la cognition et de la perception visuelle.

Certains des systèmes académiques présentés dans le cadre de ce groupe de travail, ainsi que d'autres systèmes élaborés par des équipes de recherche en France ou à l'étranger, s'affranchissent de la plupart des limites des systèmes commerciaux, notamment en garantissant un niveau plus élevé de généricité et d'extensibilité. Ainsi, Pandore²⁰ regroupe des opérateurs traitant d'images 1-D, 2-D et 3-D, en niveaux de gris, en couleurs et multispectrales sous forme de programmes exécutables et de fonctions C++ (qui peuvent être intégrées dans une solution globale par un module extérieur de programmation visuelle intitulé ARIANA). Gmic²¹ propose un environnement basé sur un langage de script, pouvant être intégré à d'autres logiciels (tels que The Gimp). Milena, dédiée à l'analyse d'images par morphologie mathématique et faisant partie de la plate-forme Olena²², offre un niveau très élevé de généricité (par exemple relative aux dimensions des images, au type des pixels, à la nature du voisinage, etc.). ImaLab²³ propose un environnement interactif couplé à un ensemble de fonctions C++ où il est aisé d'ajouter de nouvelles fonctionnalités. Dans ces différents systèmes, la caractéristique de généricité est plutôt relative aux types de données qu'aux usages des algorithmes de traitement d'image. Concevoir un environnement apte à traiter tout type d'image et destiné à tout type d'utilisateur reste encore un problème ouvert auquel nous nous sommes intéressés au travers du projet PELICAN.

3.4.2 Architecture et fonctionnalités

La motivation de concevoir une nouvelle plate-forme d'analyse et de traitement d'image plus adaptée à nos besoins remonte à 2004. C'est à cette période qu'ont été implantés ses premiers composants (réécrits depuis !) et posés ses principes fondateurs. Cette plate-forme est régulièrement utilisée comme un dépôt commun d'algorithmes de traitement d'image par des chercheurs et étudiants, dans un but de recherche ou d'enseigne-

¹⁹Systèmes Complexes pour l'Analyse et le Traitement d'Images, groupe des GDR ISIS et I3 animé par Rémy Mullot (L3i, La Rochelle), Régis Clouard (GREYC, Caen) et Didier Coquin (LISTIC, Annecy), <http://www.greyc.ensicaen.fr/scati>

²⁰Pandore : Une bibliothèque d'opérateurs de traitement d'images (Greyc), <http://www.greyc.ensicaen.fr/EquipeImage/Pandore>

²¹GREYC's magic image converter, <http://gmic.sourceforge.net>

²²Olena, <http://olena.lrde.epita.fr>

²³<http://www-prima.inrialpes.fr/Imalab>

ment. Elle est composée d'environ un millier de classes (la plupart d'entre elles représentant des algorithmes d'analyse ou traitement d'image), mais ce nombre varie au gré des départs et arrivées dans l'équipe des développeurs / utilisateurs, et des cures d'amincissement entreprises régulièrement (ainsi, PELICAN a connu un pic de 185 000 lignes de code mais la version actuelle n'en comporte qu'environ la moitié). Son architecture actuelle, qui fait l'objet de constantes évolutions et améliorations, date de 2007.

Comme indiqué précédemment, les deux principales raisons qui ont motivé la conception d'une nouvelle plate-forme furent la polyvalence et l'extensibilité. Nous associons au concept de polyvalence la capacité du système à s'adapter à différents besoins : ces besoins peuvent être liés à des usages, des données, des domaines d'application variés. Ainsi, le logiciel doit être accessible à des utilisateurs aux profils variés : néophyte souhaitant une interface "grand public" lui permettant empiriquement et interactivement de traiter ses images, traiteur d'images cherchant à rapidement prototyper ses solutions réutilisables par la suite, développeur informatique élaborant ses algorithmes élémentaires ou les combinant pour résoudre un problème, etc. L'extensibilité signifie que l'ajout de nouvelles fonctionnalités (nouveaux algorithmes, nouvelles données, nouveaux usages, etc.) est possible au sein du système. Cet ajout doit être simple à effectuer, et n'impliquer qu'un minimum de modifications dans le code source existant. Dans PELICAN, l'ajout de nouveaux algorithmes et de nouvelles interfaces ne nécessite aucune modification de ce type.

Pour garantir ces propriétés de polyvalence et d'extensibilité, la réalisation de PELICAN est basée sur le paradigme orienté-objet et la technologie Java, et son architecture est modulaire et composée de trois couches (figure 3.25) : le noyau, les algorithmes, et les interfaces. Le noyau est la partie centrale de la plate-forme, il contient l'ensemble des structures de données et des fonctions pouvant être utilisées par les éléments des autres couches. Médiateur situé entre les algorithmes et les interfaces, il évite une communication directe entre ces deux parties : il assure ainsi qu'aucune modification effectuée sur un algorithme ou une interface n'aura une répercussion sur la partie opposée, et permet de rendre les interfaces complètement indépendantes des algo-

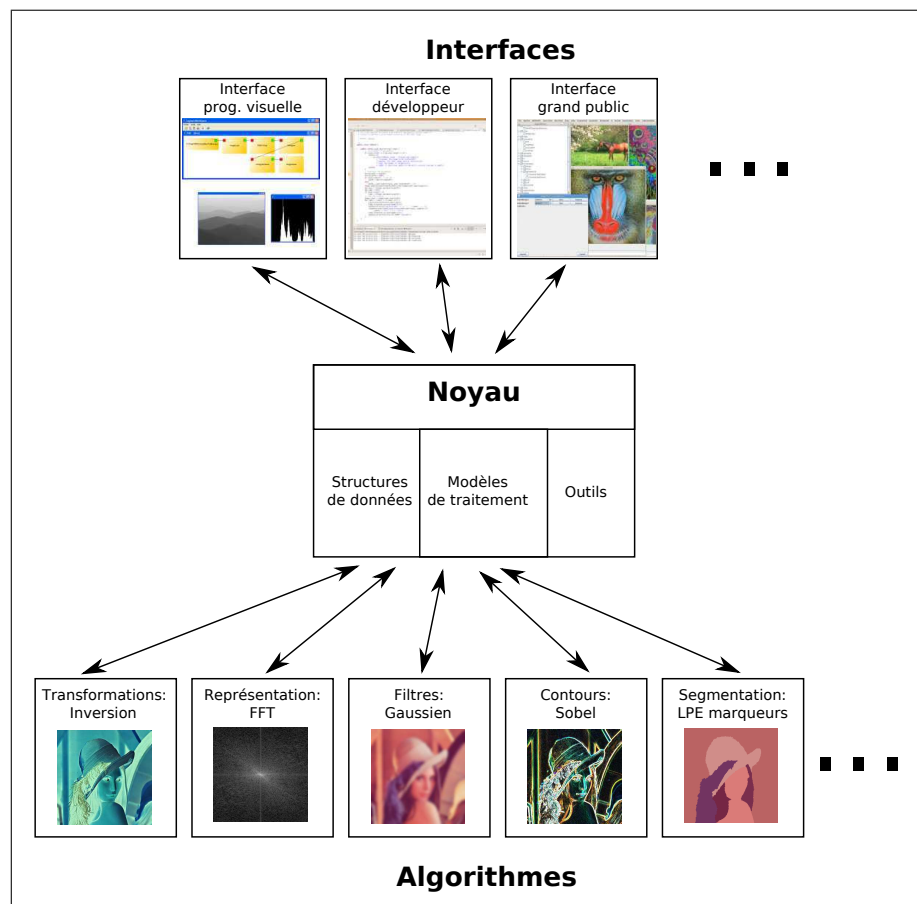


FIG. 3.25 – Architecture de la plate-forme PELICAN

algorithmes. Les algorithmes sont les techniques d'analyse et de traitement d'images présentes dans la plate-forme sous forme de classes : il peut s'agir de contributions originales de chercheurs de l'équipe ou d'implémentations d'outils publiés dans la littérature. L'écriture d'un algorithme est facilitée par les fonctionnalités offertes par le noyau, mais doit suivre un ensemble de règles prédéfinies, ce qui permet d'assurer sa compatibilité avec les autres éléments de la plate-forme. Les interfaces, dont une liste non-exhaustive sera donnée dans la prochaine section, permettent l'accès à la plate-forme et à ses algorithmes. Elles répondent à des cas d'utilisation spécifiques et sont adaptés aux différents profils d'utilisateurs.

Spécifier a priori les caractéristiques des algorithmes d'analyse ou de traitement d'image permet d'alléger la charge du développeur lors de l'écriture d'un algorithme donné, en exploitant différentes techniques de génie logiciel (généricité, abstraction, introspection, etc.). En effet, comme illustré par la figure 3.26 qui présente un exemple de code source associé à un algorithme dans PELICAN (ici une extraction des maxima régionaux

```

1 package fr.unistra.pelican.algorithms.morphology.gray.geodesic;
2 import fr.unistra.pelican.*;
3 import fr.unistra.pelican.algorithms.arithmetic.*;
4
5 /**
6  * Regional maxima extraction using morphological reconstruction
7  * algorithm.
8  * @author Aptoula, Lefevre
9  */
10 public class GrayRegionalMaxima extends Algorithm {
11
12     /**
13      * The input from which maxima should be extracted
14      */
15     public Image input;
16
17     /**
18      * The output image containing non-null values for regional maxima
19      * pixels
20      */
21     public Image output;
22
23     public GrayRegionalMaxima() {
24         super.inputs = "input";
25         super.outputs = "output";
26     }
27
28     public void launch() {
29         Image inputPlusOne = AdditionConstantChecked.exec(input, 1.0 /
30             255);
31         Image reconstruction = FastGrayReconstruction.exec(input,
32             inputPlusOne);
33         output = Difference.exec(inputPlusOne, reconstruction);
34     }
35
36     public static Image exec(Image input) {
37         return (Image) new GrayRegionalMaxima().process(input);
38     }
39 }

```

FIG. 3.26 – Exemple d'un code source définissant un algorithme dans PELICAN (ici une extraction des maxima régionaux par reconstruction morphologique).

par reconstruction morphologique), il n'est demandé au développeur que de lister les entrées (obligatoires ou optionnelles) et sorties de son traitement dans le constructeur (lignes 21 à 24), et de fournir l'implantation de son algorithme en renseignant la méthode principale de la classe (lignes 26 à 30). Une procédure de conformité exploitant les codes sources et compilés a été élaborée, pour s'assurer que les contributions (algorithmes) des différents participants au projet PELICAN respectent les règles d'écriture prédéfinies, et pour fournir toutes les informations nécessaires au bon fonctionnement des interfaces utilisateur. Elle tient compte notamment de la documentation écrite par le développeur (ici lignes 6–7, 12, et 17).

3.4.3 Interfaces et usages

Comme indiqué dans la section précédente, la conception modulaire du système PELICAN permet de séparer complètement la couche utilisateur de la couche développeur. Ainsi, quel que soit l'usage cible visé initialement lors de l'élaboration d'une méthode d'analyse ou de traitement d'image, il sera possible d'accéder à cette méthode de différentes façons, à partir de différentes interfaces. Pour illustrer l'intérêt d'un tel principe, nous donnons ici une liste d'interfaces offertes (ou qui le seront puisque, faute de ressources humaines, elles n'ont pas encore pu être toutes implantées) aux utilisateurs pour accéder à PELICAN. Certaines de ces interfaces sont fort logiquement inspirées des usages et logiciels existants, tandis que d'autres sont plus originales (et à notre connaissance encore peu ou pas présentes dans les solutions commerciales ou académiques). Notons que cette liste n'est en aucun cas exhaustive, et qu'il est très aisé, sans toucher au cœur de la plate-forme PELICAN ni à ses algorithmes, d'ajouter de nouvelles interfaces pour faire face à de nouveaux usages.

L'accès en mode console est adapté au cas d'un utilisateur connaissant parfaitement le traitement qu'il doit effectuer, sur une ou plusieurs images (par exemple convertir des images multispectrales à 4 ou 6 bandes en images en pseudo-couleurs). Il évite ainsi une perte de temps liée à la recherche dans le logiciel de la solution à son problème. Cet accès offre néanmoins des possibilités d'entrée-sortie, sous forme textuelle ou image par exemple.

L'interface graphique "grand public" s'adresse au plus grand nombre. A l'instar de Photoshop ou The Gimp, elle offre un accès aisé à l'ensemble des traitements disponibles dans PELICAN, et est donc destinée à l'utilisateur qui cherche de façon interactive et empirique une solution à son problème de traitement d'image. Cependant, à la différence des logiciels existants, aucune intervention de la part du concepteur d'un algorithme ou d'un traitement n'est ici requise pour la réalisation des composants graphiques permettant l'accès au traitement dans la plate-forme. En effet, cette interface est construite entièrement automatiquement, en s'appuyant notamment sur les mécanismes d'introspection offerts par Java.

A l'opposé d'un accès graphique et interactif, il est possible d'utiliser PELICAN comme une bibliothèque d'opérateurs et de développer en langage Java sa propre solution en s'appuyant sur les éléments déjà présents dans PELICAN. Ce mode d'utilisation permet une grande souplesse mais nécessite des compétences en programmation. Il s'adresse donc en priorité aux ingénieurs et chercheurs qui possèdent les notions algorithmiques nécessaires à la réalisation d'une solution informatique de traitement d'image. En pratique, cette interface est la plus employée puisque PELICAN reste encore principalement utilisé au LSIIT (dans les équipes MIV et FDBT), en attendant sa diffusion prochaine.

La connaissance d'un langage de programmation tel que Java n'est pas nécessaire pour concevoir une solution complexe de traitement d'image. En effet, il est possible d'utiliser une interface de programmation visuelle où chaque élément du système est représenté par un bloc, et où l'assemblage de ces blocs (par la création de relations d'entrée-sortie entre ceux-ci) permet la conception de la solution attendue. Autrement dit, le système d'analyse d'image est représenté comme un graphe, où les nœuds sont les algorithmes, et les arêtes les relations d'antériorité entre les algorithmes. Là encore, la réalisation de la représentation graphique des différents algorithmes est automatique et ne nécessite aucun effort de la part des concepteurs d'algorithmes.

Dans certains cas, l'utilisateur intéressé par certains traitements présents dans PELICAN dispose déjà d'un environnement de développement dans lequel il conçoit ses solutions de traitement d'image (par exemple le logiciel Matlab, la bibliothèque C++ ITK, etc.). L'utilisation de multiples logiciels (celui de l'utilisateur et PELICAN) pour produire le résultat souhaité n'est pas un choix très pertinent et s'avère particulièrement fastidieux. Aussi il est plus intéressant d'offrir un accès aux fonctionnalités de PELICAN directement au sein de produits tiers (voire également d'accéder aux fonctionnalités de ces produits au sein de PELICAN). Une

expérimentation a été menée dans ce sens avec la bibliothèque ITK et nous observé qu'il était relativement aisé d'accéder aux fonctionnalités de PELICAN depuis ITK, et à celles d'ITK depuis PELICAN. Notre principale cible est maintenant d'assurer une intéropérabilité avec Matlab, solution fréquemment utilisée dans la communauté du traitement d'image.

Les différents usages discutés précédemment supposent l'exécution du logiciel de traitement d'image sur le poste client, ce qui n'est pas toujours souhaitable. Aussi il est possible de proposer un accès à distance et nous nous intéressons à différentes alternatives. La première consiste à offrir une interface de type client / serveur Java : pour le client, l'utilisation semble identique à l'interface graphique grand public, mais les traitements ne sont pas effectués localement sur son poste de travail ; ils font l'objet de requêtes au serveur (avec envoi des données telles que les images si elles ne sont pas déjà disponibles sur le serveur) qui produira les résultats et les retournera à la partie cliente. Pour éviter d'imposer la présence d'un environnement d'exécution Java sur le poste client, le client Java peut être avantageusement remplacé par un client léger accessible au travers d'un navigateur web. Ce dernier est réalisé avec le langage HTML, mais il pourrait sans doute l'être également avec la technologie Flash. Outre l'accès à distance à PELICAN selon un mode client / serveur, il est possible d'offrir les traitements présents dans PELICAN sous la forme de services web. Cette interface présente l'avantage de garantir une très grande souplesse dans l'appel aux fonctionnalités de PELICAN puisqu'aucune contrainte n'est imposée du côté du client en dehors de la capacité à émettre des requêtes et à communiquer avec ces services web.

Enfin, lorsque le problème posé en analyse d'image amène différentes solutions, et que nous souhaitons mettre à dispositions différentes alternatives de traitement à l'utilisateur, il faut concevoir un logiciel spécifique (par exemple propre à la détection d'objets du littoral en télédétection). Pour limiter le coût nécessaire à une telle réalisation, il est possible de faire appel à un générateur d'exécutable qui sera chargé de la réalisation du logiciel : sélection des classes Java, assemblage au sein d'une interface utilisateur graphique ou non (pouvant nécessiter la création de classes dédiées). Spécifier les fonctionnalités offertes par le logiciel, tâche incombant bien sûr toujours au concepteur, peut être effectué à l'aide de fichiers de configurations structurés selon une syntaxe prédéfinie (par exemple des fichiers XML). Notons que cette démarche peut également s'appliquer pour la réalisation d'un logiciel plus simple, offrant par exemple l'accès en mode console à une unique fonctionnalité (qui peut néanmoins nécessiter l'appel successif à plusieurs algorithmes). Dans ce cas, la présence d'un générateur d'exécutable permet au concepteur de produire et livrer sans effort un logiciel ou exécutable Java ne contenant que le strict nécessaire au besoin des utilisateurs finaux.

3.4.4 Conclusion

Dans cette section, nous avons brièvement décrit la plate-forme PELICAN que nous utilisons d'une part pour résoudre les différents problèmes d'analyse et de traitement d'image qui nous sont posés par nos partenaires dans le cadre de collaborations variées, et d'autre part pour concevoir les algorithmes par lesquels nous contribuons aux travaux de la communauté scientifique en analyse d'image par morphologie mathématique.

La décision de concevoir une nouvelle plate-forme, alors même que de nombreuses solutions académiques ou commerciales existent, a été prise en 2004 face au constat de l'absence de généralité de la plupart des logiciels existants. En effet, nous souhaitons disposer d'un environnement nous permettant de concevoir des algorithmes aptes à traiter des types de données variées (images 2-D ou 3-D, à niveaux de gris, couleur, ou multispectrales, séquences d'images, etc.) indépendamment du domaine d'application, et destinés à des utilisateurs aux profils différents. De plus, nous avons à cœur de favoriser la réutilisation de nos méthodes, à l'instar des objectifs du groupe de travail SCATI des GDR ISIS et I3.

Notre contribution, un système polyvalent et extensible pour l'analyse et le traitement des images, est conçue selon une architecture générique à trois couches, qui permet de séparer complètement la manière dont un algorithme est écrit de la manière dont il sera utilisé. Il combine les fonctionnalités proposées par différents logiciels existants, tant au niveau des usages que des données traitées.

3.4.5 Perspectives

Concevoir un environnement permettant le développement, l'expérimentation, et l'utilisation de méthodes en analyse et traitement des images n'est pas trivial. Lorsqu'une certaine généricité (se traduisant ici par les propriétés de polyvalence et d'extensibilité) est requise, cette tâche s'avère particulièrement délicate et nécessite des ressources humaines certaines. Le projet PELICAN n'a malheureusement pas encore pu bénéficier d'un tel soutien, et (outre mon investissement personnel) son implantation est principalement le fruit du travail de trois stagiaires, les doctorants impliqués dans le projet apportant essentiellement leur contribution à l'enrichissement des opérateurs de traitement d'image. Ce manque de ressources a eu une répercussion directe sur l'avancement de ce projet, laissant pour l'instant plusieurs des interfaces présentées précédemment à l'état de spécifications, sans implantation réelle. La liste des futures interfaces s'étoffe pour sa part au fil des avancées de la communauté scientifique : citons par exemple la plate-forme AdessoWiki [LMKR09] qui propose un wiki pour l'apprentissage du traitement d'image et permet l'exécution des traitements directement sur le site, la bibliothèque Milena [LGN09] qui garantit un haut-niveau de généricité, ou l'environnement ImaLab qui propose un langage de script extensible [Lux04].

Cette situation a également retardé la distribution officielle du projet, qui sera accessible sous licence libre lorsqu'il aura atteint un niveau suffisant de maturité. Ainsi nous pourrons également présenter le projet et ses caractéristiques à la communauté du GT SCATI, bénéficier d'interactions riches avec les différents membres de ce groupe de travail et nous y impliquer plus fortement.

Chapitre 4

Programme de recherche

Dans les chapitres précédents, nos différentes contributions à la morphologie mathématique, qu’elles soient de nature fondamentale ou appliquée, ont été présentées. En amont, nous avons cherché à adapter la morphologie mathématique aux images multivaluées, en proposant des solutions basées d’une part sur des ordres vectoriels totaux équilibrés, et d’autre part sur des décompositions en ensembles d’images. Pour assurer l’intégration de connaissance dans les traitements morphologiques, nous avons étudié le rôle pouvant être joué par les éléments structurants, mais aussi comment il était possible de s’appuyer sur des techniques de classification supervisée ou non-supervisée, notamment dans un processus de segmentation d’image.

Ces différents travaux ont été appliqués principalement en analyse d’image couleur (pour l’annotation et la recherche par le contenu), en télédétection (à très haute résolution spatiale), et en imagerie astronomique. Dans ces domaines, les problèmes étudiés ont concerné la détection d’objet, la segmentation en régions, et la description et classification du contenu des images. En outre, nous avons présenté la plate-forme PELICAN que nous avons élaborée pour faciliter la conception de solutions d’analyse et de traitement d’image.

Les différents thèmes de recherche auxquels nous nous sommes intéressés en morphologie mathématique sont tous porteurs de différentes perspectives que nous avons systématiquement détaillées. Ainsi, nous souhaitons étudier l’adaptation de la morphologie mathématique à d’autres types d’images : images hypervaluées, séquences d’images, images d’étiquette. L’intégration de connaissances pourrait être améliorée par une modélisation avancée des connaissances ou par l’utilisation de nouveaux paradigmes en apprentissage et fouille de données. Elle serait aussi facilitée par la conception itérative d’une solution morphologique, pouvant aboutir au principe de morphologie mathématique par l’exemple. Dans le domaine de la recherche d’images couleur par le contenu, nos travaux pourraient être poursuivis dans l’objectif de réaliser l’indexation et la recherche dans de grandes bases d’images, en étudiant notamment les points d’intérêt et descripteurs locaux d’un point de vue morphologique. En télédétection, ce sont plutôt les profils morphologiques qui motivent de futurs travaux, ainsi que l’élaboration d’approches multi-sources. Enfin, l’analyse d’images astronomiques multispectrales et hyperspectrales mériterait d’être étudiée. Ce contexte applicatif est particulièrement propice pour mener une réflexion sur la morphologie statistique.

Parmi ces nombreuses perspectives, lesquelles étudier en priorité ? Il est bien difficile de répondre à cette question, et nous avons plutôt choisi de présenter par la suite trois thèmes de recherche que nous souhaiterions développer dans le contexte de différentes collaborations. Le premier porte sur l’imagerie du vivant et s’inscrit exclusivement dans le contexte géographique du LSIIT, au travers de partenariats avec d’autres laboratoires des universités de Strasbourg et de Freiburg. Le second est relatif à l’analyse de séquences vidéo (ou le retour aux problématiques chères à mon travail doctoral). Même s’il a pour origine des collaborations initiées sur Strasbourg avec le laboratoire IPGS (projet BQR UDS 2008–2009) d’une part, et la société RBS (thèse CIFRE de Jonathan Weber débutée en 2007 en partenariat avec l’équipe FDBT) d’autre part, mon accueil en délégation en 2009–2010 dans l’équipe TEXMEX de l’IRISA à Rennes me semble particulièrement propice à son étude. Enfin, le dernier thème relatif à l’utilisation de la morphologie mathématique hors du domaine du traitement d’image, s’inscrit exclusivement dans le contexte de mon accueil dans l’équipe TEXMEX, qui s’intéresse à l’exploitation de documents multimédia, donc ne contenant pas uniquement des images (mais également des données textuelles, sonores, etc.).

4.1 Imagerie du vivant

La morphologie mathématique fournit depuis longtemps des solutions pertinentes à différents problèmes posés en imagerie biologique ou médicale [GB00]. Dans ces deux domaines, les applications peuvent avoir en commun la nature tridimensionnelle des images et le caractère imprécis des structures à localiser, identifier et/ou caractériser. Afin de pouvoir élaborer des méthodes robustes et efficaces de détection, segmentation, ou description d'objet en se basant sur la morphologie mathématique, il est nécessaire de s'intéresser aux problèmes d'invariance des opérateurs et d'imprécision des objets considérés.

4.1.1 Invariance et imprécision

Dans des domaines aussi variés que l'imagerie couleur, la télédétection, et l'imagerie astronomique, nous avons apporté au problème de la détection d'objets différentes réponses basées sur la transformée en tout-ou-rien, nos contributions portant notamment sur la gestion de l'information spectrale et du bruit. Pour faire face à la variabilité des formes et/ou des orientations, nous avons suivi l'approche communément adoptée consistant à considérer plusieurs éléments structurants et à fusionner les résultats fournis par la transformée en tout-ou-rien. Dans le cas d'images 3-D telles que celles rencontrées en imagerie du vivant, le nombre de motifs nécessaires augmente de façon drastique et rend impraticable cette stratégie. Il faut donc adopter une autre approche, en tentant d'assurer l'invariance des opérateurs morphologiques (tels que la transformée en tout-ou-rien) à différentes transformations géométriques (autre que la translation et le contraste qui sont intrinsèquement pris en compte en morphologie mathématique) : rotation, changement d'échelle, etc.

En collaboration avec l'équipe LMB (cf. section suivante sur l'imagerie biologique) dont un des domaines d'expertise concerne les invariants géométriques [BS01], je m'intéresse donc à la question théorique de l'invariance des approches morphologiques aux principales transformations géométriques. Dans le cas général, différentes stratégies (notamment la normalisation, la dérivation, et l'intégration) permettent d'assurer l'invariance des descripteurs ou des opérateurs, et nous étudions l'application de ces stratégies au cadre morphologique. Cette question reste peu étudiée (notons néanmoins les travaux de Roerdink [Roe00]) alors qu'elle nous semble fondamentale. Les débouchés potentiels de ces travaux sont nombreux et non limités à l'imagerie biologique : ainsi, l'élaboration d'une méthode morphologique de détection d'objet (par transformée en tout-ou-rien) qui soit robuste aux transformations géométriques pourrait par exemple être utilisée en recherche par le contenu dans les images couleur ou les séquences vidéo.

En complément à cette question d'invariance aux transformations géométriques, nous pensons que les approches morphologiques pourraient également bénéficier du cadre flou pour accroître leur robustesse et mieux prendre en compte des structures de nature imprécise. Nous avons pour l'instant limité notre usage de la logique floue à l'exploitation de différentes techniques de classification floue (supervisée ou non) dans le but de permettre l'intégration de connaissances (lors de processus de détection ou de segmentation notamment). Il nous semble intéressant d'étudier de façon plus approfondie les travaux en morphologie mathématique floue (notamment ceux d'Isabelle Bloch qui œuvre depuis longtemps dans ce domaine [BM95, Blo08]) pour tirer pleinement bénéfice du cadre flou dans nos travaux futurs.

4.1.2 Applications en imagerie biologique

L'imagerie biologique est au cœur des préoccupations de l'équipe LMB¹ de l'Université Albert-Ludwig de Freiburg-im-Breisgau (responsable : Pr. Hans Burkhardt), au travers de son implication dans le projet BIOSS (Centre for Biological Signalling Studies) reconnu comme projet d'excellence en Allemagne. Dans le cadre d'un séjour de recherche au printemps 2008 dans cette équipe (puis de visites régulières en 2008–2009), j'ai eu l'opportunité de m'initier à ce domaine et d'étudier les contributions que la morphologie mathématique pouvait y apporter.

Plus précisément, je me suis intéressé aux problèmes de segmentation et localisation d'objets (ici des cellules) que nous avons réussi à résoudre avec succès dans d'autres domaines (imagerie couleur, télédétection,

¹Lehrstuhl für Mustererkennung und Bildverarbeitung – Reconnaissance de Formes et Analyse d'Image, <http://lmb.informatik.uni-freiburg.de>

imagerie astronomique). J'ai été confronté à deux principales difficultés : d'une part, le volume et la nature tridimensionnelle des données à traiter qui nécessitent d'adapter ou de repenser les algorithmes (de segmentation par exemple) qui étaient pertinents et efficaces dans le cas d'images 2-D (les domaines auxquels nous nous étions intéressés étant exclusivement sources d'images 2-D) ; d'autre part, le manque de robustesse des approches morphologiques pour la détection d'objet, que nous avons observé lorsque les objets et les images sont de nature tridimensionnelles, et qui est à l'origine des travaux théoriques sur l'invariance en morphologie mathématique.

Nul doute que munie de solutions plus robustes et plus efficaces, la morphologie mathématique pourrait apporter des réponses intéressantes aux problèmes posés en imagerie biologique, et je souhaite donc poursuivre la collaboration initiée avec l'équipe LMB dans ce domaine.

4.1.3 Applications en imagerie médicale

L'imagerie médicale est le principal domaine d'application des travaux en analyse d'image (que ce soit par des approches morphologiques ou statistiques) de l'équipe MIV, au travers de sa participation au programme de recherche IRMC (initialement une EPML du département STIC du CNRS, puis un PPF soutenu par l'UDS et le MESR, et maintenant un des deux axes du pôle de compétitivité alsacien à vocation mondiale "Innovations Thérapeutiques"). J'ai donc naturellement participé à différents projets en imagerie médicale, avec néanmoins un investissement bien moindre (faute de temps et de lien évident avec mes travaux fondamentaux) que celui accordé aux autres domaines d'application présentés dans le chapitre 3. Cet investissement s'est notamment traduit par l'encadrement de plusieurs étudiants en projet de master 1ère ou 2ème année.

Ainsi, nous nous sommes intéressés avec Vincent Agnus de l'IRCAD à l'utilisation de la morphologie mathématique floue pour résoudre différents problèmes de segmentation de volumes 3-D représentant les organes de l'appareil digestif. Dans ce contexte, les approches floues semblent adaptées d'une part à l'identification du réseau hépatique dans le foie (le cadre flou permettant de gérer la variabilité des propriétés des vaisseaux et ainsi d'extraire correctement les vaisseaux secondaires), et d'autre part à l'identification de tumeurs dans le foie (le cadre flou permettant ici de tenir compte du faible contraste entre les tumeurs et le reste du foie, ainsi que de l'absence de frontières nettes entre ces zones).

De plus, je me suis intéressé à l'analyse des structures cérébrales à partir de volumes 3-D acquis par IRM (Imagerie par Résonance Magnétique) en partenariat avec le LINC. Alors que j'avais déjà mené des travaux sur de telles images lors de mon stage de DEA, il s'agissait ici d'étudier l'intérêt des approches morphologiques pour résoudre différents problèmes : identification, quantification, et suivi des lésions cérébrales au cours du temps ; segmentation maxillo-faciale s'appuyant sur des connaissances disponibles sous forme d'atlas.

Même si les résultats obtenus jusqu'à aujourd'hui sont limités et loin d'être satisfaisants, des avancées sont attendues dans ce domaine, en s'appuyant notamment sur nos développements théoriques et méthodologiques en intégration de connaissances dans les traitements morphologiques (que nous envisageons de transposer au domaine de l'imagerie médicale).

4.2 Analyse de séquences vidéo

Alors que mes travaux doctoraux ont porté essentiellement sur l'analyse de séquences vidéo au moyen de différents outils, les développements que j'ai effectués par la suite en morphologie mathématique n'ont concerné que des images fixes. Récemment, différents projets ont motivé l'utilisation des approches morphologiques dans le contexte des séquences vidéo, notamment dans un but de segmentation spatio-temporelle. Mon accueil en délégation dans l'équipe TEXMEX de l'IRISA devrait également me permettre de m'intéresser à ce type de données, dans le contexte de l'indexation multimédia.

4.2.1 Segmentation spatio-temporelle

La question de la segmentation d'image a motivé une partie importante des travaux que j'ai menés en morphologie mathématique, que ce soit de par les aspects fondamentaux mis en œuvre (comment segmenter des images multivaluées ? comment guider le processus de segmentation par les connaissances de l'expert ?), ou

les applications (en imagerie couleur, télédétection, imagerie astronomique). C'est donc tout naturellement que je souhaite continuer à m'intéresser à cette question dans le contexte de séquences d'image (où les approches morphologiques ont montré [SBCP96], et continuent de montrer [RA08, FL09], leur intérêt), au travers de différents projets.

Ainsi, dans le contexte d'un projet sur les écoulements gravitaires menés avec des géophysiciens de l'IPGS (et soutenu par le conseil scientifique de l'UDS en 2008 – 2009), je souhaite étudier la capacité des méthodes de segmentation (automatique ou non) à appréhender correctement des séquences d'images haute résolution (spatiale et/ou fréquentielle) pour effectuer le suivi des écoulements gravitaires. Deux questions principales méritent une attention particulière : comment étendre les méthodes de segmentation (principalement celle guidée par des marqueurs spatiaux et spectraux [D19, E16]) aux séquences d'images (par une approche $2D + t$ ou $3D$) ? Comment prendre en compte conjointement les informations de couleur, de texture et de mouvement dans le processus de segmentation ? Il s'agit là d'un nouvel exemple de domaine d'application où la morphologie mathématique peut potentiellement offrir des solutions originales et pertinentes.

À une autre échelle (des séquences de moins bonne résolution mais en bien plus grand nombre), la segmentation est considérée comme une étape cruciale d'un processus de fouille de séquences vidéo dans la thèse de doctorat CIFRE de Jonathan Weber débutée en 2007, en collaboration avec la société RBS et Pierre Gançarski de l'équipe FDBT. Nous avons choisi ici de considérer les objets (spatio-temporels) présents dans les séquences vidéo comme élément central du processus de fouille de données, ce qui nécessite de les extraire correctement avant de pouvoir les décrire et les classer. Autrement dit, il faut disposer d'un processus de segmentation qui soit capable de produire un ensemble de régions spatio-temporelles correspondant aux objets (sémantiques) d'intérêt pour l'utilisateur. Il est bien sûr complètement irréaliste d'élaborer une telle méthode de segmentation dans un contexte non-supervisé, et nous nous intéressons à la question de l'apprentissage de segmentations (ici spatio-temporelles) : en s'appuyant sur quelques segmentations obtenues avec l'aide d'un utilisateur, on souhaite guider automatiquement le processus de segmentation morphologique sur le reste de la base de séquences vidéo. Les concepts d'apprentissage actif et d'apprentissage semi-supervisé semblent particulièrement adaptés à la problématique considérée ici, et nous étudions comment ils peuvent être mis en œuvre dans le contexte morphologique.

En outre, face au volume des données à traiter dans le contexte de la fouille de séquences vidéo, nous cherchons à élaborer une méthode de segmentation morphologique particulièrement efficace, procédant d'abord par simplification non-supervisée des données (les régions initiales pouvant par exemple être construites par les zones plates ou quasi-plates, spatiales ou spatio-temporelles) puis par manipulation supervisée (fusion, modification, suppression) de régions au travers d'une représentation de celles-ci sous une forme adaptée (graphe, arbre, etc.).

4.2.2 Indexation multimédia

L'application de la morphologie mathématique au domaine du multimédia reste faible, et cela même si les quelques résultats obtenus sont prometteurs. Mis à part le problème de la description morphologique des images dans un but d'indexation et de recherche par le contenu auquel nous nous sommes directement intéressés [Apt08], les principaux travaux se sont déroulés essentiellement au sein du Centre de Morphologie Mathématique (CMM) de l'Ecole des Mines – ParisTech : structuration de documents vidéo [Dem00] (nous avons également effectué une première tentative pour aborder ce problème lors du stage de DEA de Fateh Beder), et plus récemment détection du texte incrusté ou enfoui [Ret07] (ce travail ayant obtenu la première position à la compétition nationale ImagEval 2006) ou encore restauration de la piste son optique des films cinématographique [BHBD07].

Dans le cadre de mon accueil en délégation en 2009–2010 dans l'équipe TEXMEX de l'IRISA spécialisée dans l'exploitation de documents multimédia, je souhaite m'intéresser notamment aux solutions que la morphologie mathématique peut offrir aux différents problèmes d'analyse et de traitement d'image posés en indexation multimédia. En effet, les questions de filtrage, détection, segmentation, description, restent toujours pertinentes dans le cas de séquences vidéo, où les problèmes à résoudre sont par exemple l'extraction de l'information de mouvement (global ou local), la segmentation spatio-temporelle ou le suivi des objets (discuté dans la section précédente), le découpage en plans ou en scènes et la construction d'images-clé (dans un but de structuration

de la vidéo), l'extraction de texte incrusté, etc.

Parmi ces pistes de recherche, celle qui me semble devoir mériter le plus d'attention est la construction de descripteurs morphologiques locaux, qu'ils soient de nature purement spatiale ou spatio-temporelle. En effet, les descripteurs locaux connaissent ces dernières années un vif succès (non limité par ailleurs au domaine de l'indexation multimédia mais également rencontré en reconnaissance de formes, vision robotique 3-D, etc). Les approches locales procèdent en deux étapes : la détection de points caractéristiques dans l'image, appelés points d'intérêt, suivie de la description du voisinage de ces points, le plus souvent en assurant une invariance à certaines transformations géométriques (translation, rotation, changement d'échelle, etc). L'approche SIFT (pour Scale-Invariant Feature Transform) introduite par David Lowe [Low99,Low04] est certainement la plus connue de ces approches, et celle ayant ouvert la voie à cette orientation récente en traitement des images [MS05]. Afin d'assurer une invariance à l'échelle, l'image est le plus souvent analysée au travers d'un espace-échelle produit par des filtres linéaires (par exemple un filtre Gaussien). Le cadre morphologique fournit lui aussi des représentations en espaces-échelle (tout en remplaçant l'effet de flou obtenu avec les filtres Gaussiens par une suppression localisée des détails, conservant ainsi les contours principaux), qui permettent la construction de descripteurs globaux ou locaux tels que les profils morphologiques (cf. section 2.1.5). Ces derniers me semblent présenter plusieurs intérêts vis-à-vis des descripteurs locaux traditionnellement utilisés en indexation multimédia : caractère non-linéaire, préservation du contenu (contours) dans l'espace-échelle, capacité à produire des descriptions de taille réduite (puisque issus de représentations creuses), temps de calcul limité (du fait d'algorithmes morphologiques efficaces), etc. En outre, le problème de la production des points d'intérêt (voire des régions d'intérêt ou patches) à partir des espaces-échelle morphologiques mériterait également d'être étudié.

4.3 Morphologie mathématique : plus loin que l'analyse d'image

Les axes de recherche discutés dans les sections précédentes concernaient différents développements théoriques et/ou applicatifs de la morphologie mathématique, dans des domaines variés (imagerie du vivant, multimédia) mais toujours liés au traitement et à l'analyse des images. Une autre orientation que je souhaite donner à mes recherches porte sur l'analyse et le traitement morphologique de données non-images : en effet, cette théorie n'est par définition aucunement limitée à des données de type image, même si l'essentiel des applications de la morphologie mathématique se limitent à des données de cette nature. La possibilité d'utiliser les outils morphologiques hors de leur champ d'application historique reste encore très largement inexplorée aujourd'hui, et me semble une perspective très prometteuse. Mon accueil en délégation au sein de l'équipe TEXMEX, confrontée à l'analyse de différents média (image, vidéo, mais aussi son, texte, etc.), et à des problématiques telles que l'indexation multidimensionnelle, me paraît un contexte particulièrement favorable pour mener à bien cette étude.

4.3.1 Traitements des différents média

Par définition, un document multimédia est composé de différents média. Bien sûr, l'image reste une composante incontournable en indexation multimédia (surtout lorsque l'intéressé est un traiteur d'image !), mais d'autres sources d'information sont également importantes. Citons par exemple les données sonores ou textuelles, respectivement de nature temporelle et symbolique.

Ainsi, il est possible d'appliquer les outils morphologiques sur des séquences temporelles, dans leur espace initial ou après transformation (fréquentielle par exemple). Ces données temporelles peuvent être monodimensionnelles, le voisinage défini par l'élément structurant représentant alors une fenêtre temporelle d'analyse avec ou sans pondération, ou multidimensionnelles, lorsqu'on associe un ensemble de caractéristiques à chaque instant d'une piste audio ou à chaque image d'une séquence vidéo. Le traitement de ces données temporelles, qu'elles soient scalaires ou vectorielles, requiert donc d'adapter le cadre morphologique à de telles données : il faut redéfinir les opérateurs morphologiques de base, mais aussi les outils plus évolués tels que la ligne de partage des eaux, etc). Il en résulte des usages originaux de la morphologie mathématique où les différents traitements existants prennent un sens nouveau lorsqu'ils sont appliqués à des données temporelles. Ainsi, dans le cas de données audio, le filtrage pourrait supprimer les *outliers*, la segmentation fournirait une partition tem-

porelle de la séquence, l'appariement permettrait la reconnaissance de motifs prédéfinis (copies par exemple, mais aussi recherche par le contenu [Kar08]), et la description s'appliquerait soit à des données brutes soit aux caractéristiques numériques qui en ont été extraites.

D'autre part, nous avons indiqué précédemment (section 2.2.6) que même si la morphologie mathématique requiert qu'une structure de treillis soit imposée sur les données sur lesquelles elle est appliquée [Ron90], il n'est pas nécessaire que ce treillis soit produit à partir de données numériques. Ainsi, une définition du cadre morphologique a été donnée récemment [RA05, RA08] dans le contexte des images d'étiquettes telles que les images produites par une segmentation, où à chaque pixel est associé un identifiant sans qu'on puisse ordonner a priori les différents identifiants. Ces résultats ne sont pas limités à de telles images de labels mais peuvent être utilisés sur des données symboliques en général. Dans ce contexte, les opérateurs morphologiques pour le filtrage, la description, la reconnaissance, ou la segmentation peuvent être appliqués sur des données non-images et sont autant d'outils disponibles pour analyser et traiter de telles données de façon originale. Dans le cas du média textuel, il pourrait s'agir des classes, des thèmes, des mots, etc.

4.3.2 Classification et indexation de données multidimensionnelles

Plus généralement, la morphologie mathématique pourrait être appréhendée comme une nouvelle boîte à outils pour résoudre différents problèmes en analyse de données (il s'agit là d'un cas particulier des généralisations possibles des approches introduites dans le domaine du traitement d'image). En effet, une vision simplifiée du fonctionnement des opérateurs morphologiques consiste à les assimiler à des analyses locales non-linéaires dans des espaces mono ou multidimensionnels. L'analyse et le traitement des images n'en sont que des cas particuliers, où la morphologie mathématique a démontré avec succès son intérêt au cours des quarante dernières années. Ainsi, en complément des perspectives relatives à l'application du cadre morphologique à des données issues d'autres média tels que le son ou le texte, et suite à l'étude que j'ai menée récemment sur le lien entre segmentation morphologique et classification non-supervisée [E15], il me semble que les approches morphologiques pourraient fournir des solutions intéressantes aux problématiques inhérentes à l'analyse de données, telles que la classification ou l'indexation multidimensionnelle.

Même si les premiers travaux en classification morphologique remontent à une quinzaine d'années (cités par exemple [PZLB93, Soi96, Bre98]), l'utilisation de la morphologie mathématique pour résoudre des problèmes d'analyse de données reste rare lorsque ces dernières ne sont pas des images. L'exemple le plus représentatif en classification non-supervisée est certainement le *morphological clustering* qui consiste en l'application d'une ligne de partage des eaux sur un relief construit à partir de la densité des points dans l'espace (les bassins de rétention correspondent aux zones de forte densité, et les lignes de partage des eaux sont les zones de faible densité séparant ces bassins). Dans le contexte de l'analyse d'image, l'espace peut représenter les niveaux de gris (la ligne de partage des eaux est alors appliquée sur le complément de l'histogramme des niveaux de gris) ou des attributs multidimensionnels (par exemple les couleurs des pixels) [Soi96, LC09]. L'utilisation en classification supervisée reste pour sa part exceptionnelle [BP02, PFM⁺07]. Les travaux récents [PFM⁺07] exploitent une représentation des données sous forme de graphe (l'affectation d'un élément à une classe dépendant alors du plus court chemin aux prototypes des différentes classes) et semblent une alternative intéressante aux SVM.

Outre les problématiques relatives à la classification, la morphologie mathématique pourrait également apporter des solutions à d'autres problèmes rencontrés en analyse de données. Dans le cadre de mon accueil en délégation dans l'équipe TEXMEX, il semble pertinent d'étudier l'apport des approches morphologiques à l'indexation de grands volumes de données dans des espaces hautement multidimensionnels, puisqu'une partie des travaux dans le domaine de la morphologie mathématique porte sur l'élaboration d'algorithmes particulièrement efficaces, et que l'utilisation d'approches non-linéaires doit permettre de s'affranchir de la distribution des données (problème rencontré par exemple lorsque les descripteurs utilisés en indexation d'image fournissent des points très proches dans l'espace multidimensionnel). Les travaux que nous avons menés en morphologie mathématique multivaluée [B3], en particulier pour savoir ordonner des vecteurs, pourraient être exploités pour construire des techniques de projection sur des espaces k -D (avec $k \ll n$ et n le nombre de dimensions de l'espace) à comparer avec les projections 1-D étudiées dans l'équipe TEXMEX [LAJA09]. Les représentations en espace-échelle pourraient également être exploitées pour produire des techniques performantes d'indexation.

Pour mener à bien cette généralisation des applications de la morphologie mathématique, il serait pertinent d'étudier d'autres cadres théoriques que ceux que nous avons étudiés jusqu'à présent. En particulier, je souhaite m'intéresser à l'approche proposée par Vinh-Thong Ta et Olivier Lezoray [TEL08], qui permet l'analyse et le traitement de tout type de données pouvant être représentées par un graphe, et où les opérateurs morphologiques sont définis au moyen d'équations aux dérivées partielles. Ces travaux sont prometteurs dans le contexte des récents développements scientifiques en apprentissage topologique.

4.4 Conclusion générale

Dans ce mémoire, j'ai présenté les différents travaux de recherche menés à Strasbourg depuis mon recrutement comme Maître de Conférences. Après une thèse CIFRE à forte connotation applicative traitant de l'analyse et de l'indexation de séquences vidéo, j'ai effectué une reconversion thématique en m'intéressant à la morphologie mathématique et à deux problèmes fondamentaux qui me semblaient présenter un intérêt particulier. J'ai ainsi abordé la question de l'adaptation des approches morphologiques aux données multivaluées : nous avons proposé des solutions basées sur des ordres vectoriels dans le cadre de la thèse de doctorat d'Erchan Aptoula, mais aussi suggéré d'autres approches consistant à décomposer l'image multivaluée en ensemble d'images binaires ou à niveaux de gris (cette stratégie ayant notamment fait l'objet des stages de master de Jonathan Weber). J'ai également étudié le problème de l'intégration de connaissances en morphologie mathématique, qui peut être assurée par exemple à l'aide d'éléments structurants adaptés, ou d'introduction de procédures de classification supervisée ou non. Outre des travaux personnels, les résultats obtenus sont également le fruit de la thèse de doctorat de Sébastien Derivaux (et dans une moindre mesure des stages de master de Benjamin Perret et Jonathan Weber).

L'étude de ces deux problèmes fondamentaux n'est pas restée sans retombée applicative. Ainsi, nous nous sommes intéressés avec Erchan Aptoula à l'analyse d'image couleur au moyen de méthodes morphologiques. Les techniques élaborées pour la détection, la segmentation et la description des objets d'une image nous ont permis de construire une solution d'annotation et de recherche par le contenu, dans le cadre d'un projet soutenu par Oséo – ANVAR. De plus, j'ai participé à plusieurs projets nationaux (ANR) en télédétection avec des collègues géographes et l'équipe FDBT du LSIIT. Ma contribution à ces projets a porté sur l'élaboration de méthodes de détection et de segmentation d'objets géographiques, exploitant les connaissances des experts du domaine. Dans le cadre notamment des stages de master d'Erchan Aptoula et de Benjamin Perret, ainsi que de la thèse de doctorat entamée par ce dernier, nous nous sommes intéressés aux problématiques de détection, segmentation, et classification de galaxie en imagerie astronomique. Enfin, l'ensemble des développements effectués, qu'ils soient de nature méthodologique ou répondant à un besoin applicatif particulier, s'est appuyé sur une plate-forme d'imagerie intitulée PELICAN. Ce projet, que j'ai initié localement depuis quelques années, s'appuie sur une structure générique pour faciliter l'ajout de fonctionnalités, qu'il s'agisse d'algorithmes de traitement d'image écrits par les chercheurs de l'équipe, ou d'interfaces accessibles par les utilisateurs de nos outils.

Les travaux que j'ai menés en morphologie mathématique, seul ou dans le cadre de collaborations ou d'encadrements de jeunes chercheurs, ont apporté quelques solutions à différents problèmes de nature fondamentale ou appliquée. Loin de résoudre complètement les problèmes encore ouverts en morphologie mathématique, ils ont bien plus aidé à identifier de nouvelles pistes de recherche que je souhaite maintenant explorer. Ces perspectives ont été discutées dans les différents chapitres de ce mémoire, et les principales décrites plus en détail dans les pages précédentes. Parmi ces perspectives, si je ne devais en retenir qu'une à explorer (faute de temps ?), il s'agirait certainement de m'éloigner de mon domaine de recherche historique (l'image), et d'étudier les contributions qui pourraient être apportées par la morphologie mathématique (à laquelle je m'intéresse depuis quelques années) à des domaines variés de l'informatique et du traitement des données. En effet, rien ne restreint par définition la théorie de la morphologie mathématique au domaine de l'analyse et du traitement d'image, et pourtant rares sont les tentatives pour l'utiliser hors de ce champ d'application privilégié. Dans le cadre d'une délégation INRIA dans l'équipe TEXMEX de l'IRISA à Rennes, je souhaite prendre cette direction et, qui sait, cette mobilité géographique se traduira peut-être par une mobilité thématique, afin de m'ouvrir à de nouveaux domaines en m'appuyant sur ma culture en analyse d'image...

Bibliographie

Cette bibliographie ne contient pas les travaux propres à l’auteur, dont une liste est donnée en pages 28 à 33.

- [AJ07] J. ANGULO et D. JEULIN : Stochastic watershed segmentation. *Dans International Symposium on Mathematical Morphology (ISMM)*, pages 265–276, Rio de Janeiro, Brésil, Octobre 2007.
- [Ang07] J. ANGULO : Morphological colour operators in totally ordered lattices based on distances : Application to image filtering, enhancement and analysis. *Computer Vision and Image Understanding*, 107(1-2):56–73, Juillet 2007.
- [AO03] H.M. AL-OTUM : Morphological operators for color image processing based on Mahalanobis distance measure. *Optical Engineering*, 42(9):2595–2606, Septembre 2003.
- [Apt05] E. APTOULA : *Segmentation et classification des images astronomiques multibandes au moyen de la morphologie mathématique*. Mémoire de DEA, Université Louis Pasteur, Strasbourg, 2005.
- [Apt08] E. APTOULA : *Analyse d’Images Couleur par Morphologie Mathématique. Application à la Description, l’Annotation, et la Recherche d’Images — Colour Image Analysis with Mathematical Morphology. Application to Image Description, Annotation, and Retrieval*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, 2008.
- [AS03] J. ANGULO et J. SERRA : Morphological coding of color images by vector connected filters. *Dans IEEE International Symposium on Signal Processing and Its Applications (ISSPA)*, pages 69–72, Paris, Juillet 2003.
- [AS07] J. ANGULO et J. SERRA : Modelling and segmentation of colour images in polar representations. *Image and Vision Computing*, 25(4):475–495, Avril 2007.
- [Bar76] V. BARNETT : The ordering of multivariate data. *Journal of the Statistical Society A*, 139(3):318–355, 1976.
- [BBH96] A. BIJAOU, Y. BOBICHON et L. HUANG : Digital image compression in astronomy : morphology or wavelets. *Vistas in astronomy*, 40(4):587–594, 1996.
- [Bea06] R. BEARE : A locally constrained watershed transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7):1063–1074, Juillet 2006.
- [Ber04] S.A. BERRANI : *Recherche approximative de plus proches voisins avec contrôle probabiliste de la précision ; application à la recherche d’images par le contenu*. Thèse de doctorat, Université de Rennes 1, Février 2004.
- [BGK06] A. BLANSCHÉ, P. GANÇARSKI et J.J. KORCZAK : MACLAW : A modular approach for clustering with local attribute weighting. *Pattern Recognition Letters*, 27(11):1299–1306, 2006.
- [BHBD07] E. BRUN, A. HASSAINE, B. BESSERER et E. DECENCIÈRE : Restoration of variable area soundtracks. *Dans IEEE International Conference on Image Processing (ICIP)*, volume 4, pages 13–16, San Antonio, TX, Etats-Unis, 2007.
- [Blo08] I. BLOCH : Morphologie mathématique floue, applications en raisonnement spatial et en logique. *Dans Rencontres francophones sur la Logique Floue et ses Applications*, pages 2–9, Lens, 2008.

- [BM95] I. BLOCH et H. MAÎTRE : Fuzzy mathematical morphologies : A comparative study. *Pattern Recognition*, 28(9):1341–1387, 1995.
- [BP02] T. BARATA et P. PINA : Construction of decision region borders by geometric modelling of training sets. *Dans International Symposium on Mathematical Morphology (ISMM)*, pages 147–156, Sydney, Australie, 2002.
- [BPA03] J.A. BENEDIKTSSON, M. PESARESI et K. ARNASON : Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1940–1949, Septembre 2003.
- [BPA05] J.A. BENEDIKTSSON, J.A. PALMASON et K. ARNASON : Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):480–491, Mars 2005.
- [Bre98] E. BREEN : Morphological segmentation of non-image data – the closing function. *Dans International Symposium on Mathematical Morphology (ISMM)*, pages 223–230, Amsterdam, Pays-Bas, Juin 1998.
- [BS01] H. BURKHARDT et S. SIGGELKOW : Invariant features in pattern recognition – fundamentals and applications. *Dans C. KOTROPOULOS et I. PITAS, éditeurs : Nonlinear Model-Based Image/Video Processing and Analysis*, pages 269–307. John Wiley & Sons, 2001.
- [BS03] S. BAGLI et P. SOILLE : Morphological automatic extraction of pan-european coastline from landsat ETM+ images. *Dans International symposium on GIS and computer cartography for coastal zone management (COASTGIS)*, Octobre 2003.
- [BV00] D. BLOOMBERG et L. VINCENT : Pattern matching using the blur hit-miss transform. *Journal of Electronic Imaging*, 9(2):140–150, 2000.
- [BZQ95] P. BHATTACHARYA, W. ZHU et K. QIAN : Shape recognition method using morphological hit-or-miss transform. *Optical Engineering*, 34(6):1718–1725, 1995.
- [CBNF97] A. J. CANDEAS, U. BRAGA-NETO et E. Carvalho FILHO : A mathematical morphology approach to the star/galaxy characterization. *Journal of the Brazilian Computer Society*, 3(3):14–29, Avril 1997.
- [CD99] M. COMER et E. DELP : Morphological operations for color image processing. *Journal of Electronic Imaging*, 8(3):279–289, Juillet 1999.
- [CDW05] A.P. CARLEER, O. DEBEIR et E. WOLFF : Assessment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering and Remote Sensing*, 71(11):1285–1294, 2005.
- [Cha98] J. CHANUSSOT : *Approches Vectorielles ou Marginales pour le Traitement d’Images Multi-composantes*. Thèse de doctorat, Université de Savoie, 1998.
- [DB02] E.R. DOUGHERTY et J. BARRERA : Pattern recognition theory in nonlinear signal processing. *Journal of Mathematical Imaging and Vision*, 16:181–197, 2002.
- [Dem00] C.H. DEMARTY : *Segmentation et Structuration d’un Document Vidéo pour la Caractérisation et l’Indexation de son Contenu Sémantique – Application aux Journaux Télévisés*. Thèse de doctorat, Ecole des Mines de Paris, Janvier 2000.
- [Der09] S. DERIVAUX : *Construction et Classification d’Objets à partir d’Images de Télédétection par une Approche Itérative Guidée par des Connaissances du Domaine*. Thèse de doctorat, Université de Strasbourg, Strasbourg, 2009.
- [DHS01] R.O. DUDA, P.E. HART et D.G. STORK : *Pattern Classification*. Wiley, New York, 2ème édition, 2001.
- [DJLW08] R. DATTA, D. JOSHI, J. LI et J.Z. WANG : Image retrieval : Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 40(2):38–97, Avril 2008.

- [Fer05] M. FERECATU : *Image retrieval with active relevance feedback using both visual and keyword-based descriptors*. Thèse de doctorat, Université de Versailles, Juillet 2005.
- [FL04] M. FRUCCI et G. LONGO : Watershed transform and the segmentation of astronomical images. *Dans Proceedings of Astronomical Data Analysis III*, Naples, Italy, April 2004.
- [FL09] F.C. FLORES et R.d.A. LOTUFO : Watershed from propagated markers : An interactive method to morphological object segmentation in image sequences. *Image and Vision Computing*, 2009. A paraître, doi :10.1016/j.imavis.2009.06.015.
- [FR99] F. FORBES et A. RAFTERY : Bayesian morphology : Fast unsupervised bayesian image analysis. *Journal of the American Statistical Association*, 94(446):555–568, Juin 1999.
- [GB00] J. GOUTSIAS et S. BATMAN : Morphological methods for biomedical image analysis. *Dans* M. SONKA et J.M. FITZPATRICK, éditeurs : *Handbook of medical imaging*, volume 2, pages 175–272. SPIE Press, 2000.
- [GBL04] P. GOUTON, P. BONTON et P. LAMBERT, éditeurs. *Traitement du Signal, Numéro Spécial sur l’Imagerie Numérique Couleur*, 21(5-6), 2004.
- [GHS95] J. GOUTSIAS, H. HEIJMANS et K. SIVAKUMAR : Morphological operators for image sequences. *Computer Vision and Image Understanding*, 62(3):326–346, Novembre 1995.
- [GMA⁺04] V. GRAU, A.U.J. MEWES, M. ALCANIZ, R. KIKINIS et S.K. WARFIELD : Improved watershed transform for medical image segmentation using prior information. *IEEE Transactions on Medical Imaging*, 23(4):447–458, Avril 2004.
- [Gol89] D. GOLDBERG : *Genetic Algorithms in Search, Optimization and Machine Learning*. Kluwer Academic Publishers, 1989.
- [Gro07] P. GROS, éditeur. *L’indexation multimédia : description et recherche automatiques*. Traité IC2 (série traitement du signal et de l’image). Hermès – Lavoisier, 2007.
- [GW07] P. GANCARSKI et C. WEMMERT : Collaborative multi-step mono-level multi-strategy classification. *International Journal of Multimedia Tools and Applications*, 35(1):1–27, Octobre 2007.
- [GW08] R.C. GONZALEZ et R.E. WOODS : *Digital Image Processing*. Prentice Hall, 3ème édition, 2008.
- [Han02] A. HANBURY : *Morphologie Mathématique sur le Cercle Unité : avec Applications aux Teintes et aux Textures Orientées*. Thèse de doctorat, Ecole des Mines, Paris, 2002.
- [Han08] A. HANBURY : Constructing cylindrical coordinate colour models. *Pattern Recognition Letters*, 29(4):494–500, Mars 2008.
- [HB91] L. HUANG et A. BIJAOU : Astronomical image data compression by morphological skeleton transformation. *Experimental Astronomy*, 1(5):311–327, 1991.
- [HG00] G. HEENE et S. GAUTAMA : Optimisation of a coastline extraction algorithm for object-oriented matching of multisensor satellite imagery. *Dans IEEE International Geosciences and Remote Sensing Symposium (IGARSS)*, volume 6, pages 2632–2634, 2000.
- [HGN04] E. HADJIDEMETRIOU, M.D. GROSSBERG et S.K. NAYAR : Multiresolution histograms and their use in recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7):831–847, Juillet 2004.
- [HHM03] M.S. HAMID, N.R. HARVEY et S. MARSHALL : Genetic algorithm optimization of multidimensional grayscale soft morphological filters with applications in film archive restoration. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(5):406–416, Mai 2003.
- [HM09] A. HANBURY et B. MARCOTEGUI : Morphological segmentation on learned boundaries. *Image and Vision Computing*, 27(4):480–488, 2009.
- [HR90] H. HEIJMANS et C. RONSE : The algebraic basis of mathematical morphology : I. Dilations and erosions. *Computer Vision, Graphics and Image Processing*, 50:245–295, 1990.
- [HS85] R.M. HARALICK et L.G. SHAPIRO : Image segmentation techniques. *Computer Vision, Graphics and Image Processing*, 29:100–132, 1985.

- [HS01] A. HANBURY et J. SERRA : Morphological operators on the unit circle. *IEEE Transactions on Image Processing*, 10(12):1842–1850, Décembre 2001.
- [Jai09] A.K. JAIN : Data clustering : 50 years beyond k-means. *Pattern Recognition Letters*, 2009. A paraître, doi :10.1016/j.patrec.2009.09.011.
- [JC02] Q. JISHUANG et W. CHAO : A multi-threshold based morphological approach for extracting coastal line feature in remote sensing. *Dans ISPRS Commission I Symposium on Integrated Remote Sensing at the Global, Regional and Local Scale*, volume XXXIV part 1, 2002.
- [JMF99] A.K. JAIN, M.N. MURTY et P.J. FLYNN : Data clustering : A review. *ACM Computing Surveys*, 31(3):264–323, 1999.
- [Jol05] A. JOLY : *Recherche par similarité statistique dans une grande base de signatures locales pour l'identification rapide d'extraits vidéo*. Thèse de doctorat, Université de La Rochelle, Février 2005.
- [Kar08] M. KARVONEN : *Using mathematical morphology for geometric music retrieval*. Mémoire de master, Université de Helsinki, Avril 2008.
- [KGG85] J.M. KELLER, M.R. GRAY et J.A. GIVENS : A fuzzy k-nearest neighbor algorithm. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(4):580–585, Juillet/Août 1985.
- [LAJA09] H. LEJSEK, F.H. ASMUNDSSON, B. JONSSON et L. AMSALEG : NV-Tree : An efficient disk-based index for approximate search in very large high-dimensional collections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(5):869–883, 2009.
- [Lav05] N. LAVEAU : *Mouvement et vidéo : estimation, compression et filtrage morphologique*. Thèse de doctorat, Ecole des Mines de Paris, Décembre 2005.
- [LC02] O. LEZORAY et H. CARDOT : Cooperation of color pixel classification schemes and color watershed : a study for microscopical images. *IEEE Transactions on Image Processing*, 11(7):783–789, Juillet 2002.
- [LC09] O. LEZORAY et C. CHARRIER : Color image segmentation using morphological clustering and fusion with automatic scale selection. *Pattern Recognition Letters*, 30(4):397–406, Mars 2009.
- [LGN09] R. LEVILLAIN, T. GÉRAUD et L. NAJMAN : Milena : Write generic morphological algorithms once, run on many kinds of images. *Dans International Symposium on Mathematical Morphology (ISMM)*, volume 5720 de *Lecture Notes in Computer Science*, pages 295–306, Groningen, Pays-Bas, Août 2009. Springer-Verlag.
- [LK89] S. M. LEA et L. A. KELLAR : An algorithm to smooth and find objects in astronomical images. *The Astronomical Journal*, 97(4):1238–1276, Avril 1989.
- [LLH94] M. LYBANON, S. M. LEA et S. M. HIMES : Segmentation of diverse image types using opening and closing. *Dans IAPR International Conference on Pattern Recognition (ICPR)*, pages 347–351, Jérusalem, Israël, 1994.
- [LMKR09] R. LOTUFO, R. MACHADO, A. KORBES et R.G. RAMOS : AdessoWiki – on-line collaborative scientific programming platform. *Dans International Symposium on Wikis and Open Collaboration (WikiSym)*, Orlando, Etats-Unis, Octobre 2009.
- [Low99] D. LOWE : Object recognition from local scale-invariant features. *Dans IEEE International Conference on Computer Vision (ICCV)*, pages 1150–1157, Corfu, Grèce, Septembre 1999.
- [Low04] D. LOWE : Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [Lux04] A. LUX : The imalab method for vision systems. *Journal of Machine Vision and Applications*, 16(1):21–26, Décembre 2004.
- [LZ07] I. LEVNER et H. ZHANG : Classification driven-watershed segmentation. *IEEE Transactions on Image Processing*, 16(5):1437–1445, 2007.

- [Mac96] M. MACCARONE : Fuzzy mathematical morphology : concepts and applications. *Vistas in astronomy*, 40(4):469–477, 1996.
- [Mai05] N. MAILLOT : *Ontology Based Object Learning and Recognition*. Thèse de doctorat, Université de Nice – Sophia Antipolis, 2005.
- [Mar88] P. MARAGOS : Optimal morphological approaches to image matching and object detection. *Dans IEEE International Conference on Computer Vision (ICCV)*, pages 695–699, 1988.
- [Mar89] P. MARAGOS : Pattern spectrum and multiscale shape representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):701–716, Juillet 1989.
- [Mat75] G. MATHERON : *Random sets and integral geometry*. Wiley, New York, 1975.
- [Mey91] F. MEYER : Un algorithme optimal de ligne de partage des eaux. *Dans Congrès de Reconnaissance des Formes et d'Intelligence Artificielle (RFIA)*, pages 847–857, Lyon, 1991.
- [Mey98] F. MEYER : The levelings. *Dans International Symposium on Mathematical Morphology (ISMM)*, pages 199–206, Amsterdam, Pays-Bas, Juin 1998.
- [Mey04] F. MEYER : Levelings, image simplification filters for segmentation. *Journal of Mathematical Imaging and Vision*, 20:59–72, 2004.
- [MFM04] D. MARTIN, C. FOWLKES et J. MALIK : Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5):530–549, 2004.
- [MFTM01] D. MARTIN, C. FOWLKES, D. TAL et J. MALIK : A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *Dans IEEE International Conference on Computer Vision (ICCV)*, pages 416–423, Juillet 2001.
- [MH05] B. MICUSIK et A. HANBURY : Steerable semi-automatic segmentation of textured images. *Dans Scandinavian Conference on Image Analysis (SCIA)*, volume 3540 de *Lecture Notes in Computer Science*, pages 35–44. Springer-Verlag, 2005.
- [MM00] F. MEYER et P. MARAGOS : Nonlinear scale-space representation with morphological levelings. *Journal of Visual Communication and Image Representation*, 11(2):245–265, 2000.
- [MS02] G. MATHERON et J. SERRA : The birth of mathematical morphology. *Dans International Symposium on Mathematical Morphology (ISMM)*, pages 1–16, Sydney, Australie, 2002.
- [MS05] K. MIKOLAJCZYK et C. SCHMID : A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005.
- [MTD92] M. MACCARONE, M. TRIPICIANO et V. DI GESU : Fuzzy mathematical morphology to analyse astronomical images. *Dans IAPR International Conference on Pattern Recognition (ICPR)*, volume 3, pages 248–251, La Haye, Pays-Bas, Septembre 1992.
- [Noy08] G. NOYEL : *Filtrage, réduction de dimension, classification et segmentation morphologique hyperspectrale*. Thèse de doctorat, Ecole des Mines de Paris, Septembre 2008.
- [NPR07a] B. NAEGEL, N. PASSAT et C. RONSE : Grey-level hit-or-miss transforms - part 1 : Unified theory. *Pattern Recognition*, 40(2):635–647, Février 2007.
- [NPR07b] B. NAEGEL, N. PASSAT et C. RONSE : Grey-level hit-or-miss transforms - part 2 : Application to angiographic image processing. *Pattern Recognition*, 40(2):648–658, Février 2007.
- [NT08] L. NAJMAN et H. TALBOT, éditeurs. *Morphologie mathématique (tome 1 : approches déterministes)*. Hermès – Lavoisier, 2008.
- [NWvdB03] H.T. NGUYEN, M. WORRING et R. van den BOOMGAARD : Watersnakes : energy-driven watershed segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):330–342, Mars 2003.
- [OTAP01] F. ORTIZ, F. TORRES, J. ANGULO et S. PUENTE : Comparative study of vectorial morphological operations in different color spaces. *Dans SPIE Conference on Intelligent Robots and Computer Vision*, volume 4572, pages 259–268, Novembre 2001.

- [PB01] M. PESARESI et J.A. BENEDIKTSSON : A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(2):309–320, Février 2001.
- [Per07] B. PERRET : *Détection et caractérisation de galaxies fantômes en imagerie astronomique bi-bande*. Mémoire de master, Université Louis Pasteur, Strasbourg, 2007.
- [Per09] B. PERRET : Arbres de composantes connexes généralisés et applications en imagerie astronomique. Rapport de seconde année de thèse, Université de Strasbourg, 2009.
- [PF00] S. PHILIPP-FOLIGUET : Segmentation d’images en régions floues. *Dans Rencontres Franco-phones sur la Logique Floue et ses Applications*, pages 186–196, La Rochelle, 2000.
- [PFM⁺07] J.P. PAPA, A.X. FALCAO, P.A.V. MIRANDA, C.T.N. SUZUKI et N.D.A. MASCARENHAS : Design of robust pattern classifiers based on optimum-path forests. *Dans International Symposium on Mathematical Morphology (ISMM)*, pages 337–348, Rio de Janeiro, Brésil, Octobre 2007.
- [PHL01] I. PATRAS, E.A. HENDRIKS et R.L. LAGENDIJK : Video segmentation by map labeling of watershed segments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):326–332, Mars 2001.
- [PMCS09] B. PERRET, V. MAZET, C. COLLET et E. SLEZAK : Décomposition d’images multispectrales de galaxies au moyen d’algorithmes de monte carlo par chaines de markov. *Dans Colloque GRETSI sur le Traitement du Signal et des Images*, Troyes, France, Septembre 2009.
- [PMPP04] A. PLAZA, P. MARTINEZ, R. PEREZ et J. PLAZA : A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles. *Pattern Recognition*, 37(6):1097–1116, Juin 2004.
- [PP07] M. PESARESI et E. PAGOT : Post-conflict reconstruction assessment using image morphological profile and fuzzy multicriteria approach on 1-m-resolution satellite data. *Dans IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Paris, Avril 2007.
- [PP08] E. PAGOT et M. PESARESI : Systematic study of the urban postconflict change classification performance using spectral and structural features in a support vector machine. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(2):120–128, Juin 2008.
- [Pui03] A. PUISSANT : *Information géographique et images à très haute résolution : utilité et applications en milieu urbain*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, France, 2003.
- [Pui08] A. PUISSANT : Les échelles de la télédétection pour l’analyse des territoires. Journées Géomatiques de l’Ouest, Juin 2008.
- [PZLB93] J.G. POSTAIRE, R. ZHANG et C. LECOCQ-BOTTE : Cluster analysis by binary morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(2):170–180, 1993.
- [RA05] C. RONSE et V. AGNUS : Morphology on label images : flat-type operators and connections. *Journal of Mathematical Imaging and Vision*, 22(2/3):283–307, 2005.
- [RA08] C. RONSE et V. AGNUS : Geodesy on label images, and applications to video sequence processing. *Journal of Visual Communication and Image Representation*, 19(6):392–408, Août 2008.
- [RBD92] J.F. RIVEST, S. BEUCHER et J. DELHOMME : Marker-controlled segmentation : an application to electrical borehole imaging. *Journal of Electronic Imaging*, 1(2):136–142, Avril 1992.
- [Ren07] A. RENOUF : *Modélisation de la formulation d’applications de traitement d’images*. Thèse de doctorat, Université de Caen, 2007.
- [Ret07] T. RETORNAZ : *Localisation de textes enfouis dans des bases d’images généralistes : un descripteur sémantique pour l’indexation*. Thèse de doctorat, Ecole des Mines de Paris, Octobre 2007.
- [RH91] C. RONSE et H. HEIJMANS : The algebraic basis of mathematical morphology : II. Openings and closings. *Computer Vision, Graphics and Image Processing : Image Understanding*, 54(1):74–97, 1991.

- [Riv04] J.F. RIVEST : Morphological operators on complex signals. *Signal Processing*, 84(1):133–139, Janvier 2004.
- [RM00] J. ROERDINK et A. MEIJSTER : The watershed transform : Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae*, 41(1/2):187–228, 2000.
- [Roe00] J.B.T.M. ROERDINK : Group morphology. *Pattern Recognition*, 33(6):877–895, 2000.
- [Ron90] C. RONSE : Why mathematical morphology needs complete lattices. *Signal Processing*, 21(2):129–154, Octobre 1990.
- [Ron96] C. RONSE : A lattice-theoretical morphological view on template extraction in images. *Journal of Visual Communication and Image Representation*, 7(3):273–295, 1996.
- [RSB93] J.F. RIVEST, P. SOILLE et S. BEUCHER : Morphological gradients. *Journal of Electronic Imaging*, 2(4):326–336, Décembre 1993.
- [Sal08] P. SALEMBIER : Opérateurs connexes et arbre des coupes. Dans L. NAJMAN et H. TALBOT, éditeurs : *Morphologie mathématique (tome 1 : approches déterministes)*, pages 153–172. Hermès – Lavoisier, 2008.
- [SBCP96] P. SALEMBIER, P. BRIGGER, J.R. CASAS et M. PARDÀS : Morphological operators for image and video compression. *IEEE Transactions on Image Processing*, 5(6):881–898, Juin 1996.
- [Ser82] J. SERRA : *Image Analysis and Mathematical Morphology, Volume 1*. Academic Press, 1982.
- [Ser88] J. SERRA : *Image Analysis and Mathematical Morphology, Volume 2 : Theoretical Advances*. Academic Press, 1988.
- [Ser98] J. SERRA : Connectivity on complete lattices. *Journal of Mathematical Imaging and Vision*, 9(3):231–251, 1998.
- [SOG98] P. SALEMBIER, A. OLIVERAS et L. GARRIDO : Antiextensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, 7(4):555–570, Avril 1998.
- [Soi96] P. SOILLE : Morphological partitioning of multispectral images. *Journal of Electronic Imaging*, 5(3):252–265, Juillet 1996.
- [Soi03] P. SOILLE : *Morphological Image Analysis : Principles and Applications*. Springer-Verlag, Berlin, 2003.
- [SP02] P. SOILLE et M. PESARESI : Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 40(9):2042–2055, Septembre 2002.
- [Ste86] S.R. STERNBERG : Grayscale morphology. *Computer Vision, Graphics and Image Processing*, 35(3):333–355, 1986.
- [TEJ98] H. TALBOT, C. EVANS et R. JONES : Complete ordering and multivariate mathematical morphology. Dans H. HEIJMANS et J. ROERDINK, éditeurs : *Mathematical Morphology and Its Applications to Image and Signal Processing*, pages 27–34, Amsterdam, 1998. Kluwer Academic Press.
- [TEL08] V.T. TA, A. ELMOATAZ et O. LEZORAY : Partial difference equations over graphs : morphological processing of arbitrary discrete data. Dans *European Conference on Computer Vision (ECCV)*, volume 5304 de *Lecture Notes in Computer Science*, pages 668–680. Springer-Verlag, 2008.
- [TFMB04] A. TRÉMEAU, C. FERNANDEZ-MALOIGNE et P. BONTON : *Imagerie Numérique Couleur*. Dunod, 2004.
- [TZR95] L.A. TAMBURINO, M.A. ZMUDA et M.M. RIZKI : Generating pattern-recognition systems using evolutionary learning. *IEEE Expert*, 10(4):63–68, Août 1995.
- [Vin00] L. VINCENT : Granulometries and opening trees. *Fundamenta Informaticae*, 41(1-2):57–90, Janvier 2000.

- [VPS03] I. VANHAMEL, I. PRATIKAKIS et H. SAHLI : Multiscale gradient watersheds of color images. *IEEE Transactions on Image Processing*, 12(6):617–626, Juin 2003.
- [VS91] L. VINCENT et P. SOILLE : Watersheds in digital spaces : An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, Juin 1991.
- [Web07] J. WEBER : *Détection d’objets géographiques du littoral dans des images satellites multispectrales*. Mémoire de master, Université Louis Pasteur, Strasbourg, 2007.
- [Wit83] A.P. WITKIN : Scale-space filtering. Dans *International Joint Conference on Artificial Intelligence*, pages 1019–1022, Karlsruhe, Germany, 1983.

Annexe A

Publications jointes

Pour compléter la présentation de nos travaux en morphologie mathématique, et approfondir les aspects scientifiques et techniques de certains sujets, différentes publications représentatives sont incluses en annexe :

- A comparative study on multivariate mathematical morphology A. 1
(*Pattern Recognition*, 2007 ; avec E. Aptoula)
où nous dressons un état de l’art des ordres vectoriels pour la morphologie mathématique multivaluée (section 2.2)
- On lexicographical ordering in multivariate mathematical morphology A. 17
(*Pattern Recognition Letters*, 2008 ; avec E. Aptoula)
où nous discutons le problème de dissymétrie induit par l’ordre lexicographique et proposons différents solutions basées sur des modèles de quantification (section 2.2)
- Segmentation par ligne de partage des eaux avec marqueurs spatiaux et spectraux A. 27
(actes du *Colloque GRETSI sur le Traitement du Signal et des Images*, 2009)
où nous suggérons de mieux exploiter le contenu des marqueurs pour guider le processus de segmentation par ligne de partage des eaux (sections 2.2 et 2.3)
- A new approach for unsupervised classification in image segmentation A. 31
(chapitre pour un livre à paraître suite à EGC’2009 : *Advances in Knowledge Discovery and Management*)
où nous étudions le lien entre segmentation et classification et proposons une reformulation de l’algorithme k-means pour la segmentation, basée sur des algorithmes morphologiques (section 2.3)
- Morphological description of colour images for content-based image retrieval A. 51
(*IEEE Transactions on Image Processing*, 2009 ; avec E. Aptoula)
où nous introduisons de nouveaux descripteurs morphologiques dans le contexte de la recherche d’image couleur par le contenu (sections 2.2 et 3.1)
- Supervised segmentation using machine learning and evolutionary computation A. 65
(soumis à *Pattern Recognition Letters*, 2009 ; avec S. Derivaux, G. Forestier et C. Wemmert)
où nous présentons comment des zones d’apprentissage peuvent être utilisées pour guider la segmentation d’images de télédétection (sections 2.3 et 3.2)
- A robust hit-or-miss transform for template matching in very noisy astronomical images A. 85
(*Pattern Recognition*, 2009 ; avec B. Perret et C. Collet)
où nous proposons une formulation de la transformée en tout-ou-rien robuste au bruit et son application à la détection d’objets astronomiques dans des images à niveaux de gris (sections 2.3 et 3.3)

Available online at www.sciencedirect.com

Pattern Recognition 40 (2007) 2914–2929

**PATTERN
RECOGNITION**

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/pr

A comparative study on multivariate mathematical morphology

E. Aptoula*, S. Lefèvre

*Image Sciences, Computer Sciences and Remote Sensing Laboratory, LSIIT-UMR-7005-CNRS Louis Pasteur University, Pôle API, Blvd Sébastien Brant,
P.O. Box 10413, 67412 Illkirch Cedex, France*

Received 24 May 2006; received in revised form 7 February 2007; accepted 8 February 2007

Abstract

The successful application of univariate morphological operators on several domains, along with the increasing need for processing the plethora of available multivalued images, have been the main motives behind the efforts concentrated on extending the mathematical morphology framework to multivariate data. The few theoretical requirements of this extension, consisting primarily of a ranking scheme as well as extrema operators for vectorial data, have led to numerous suggestions with diverse properties. However, none of them has yet been widely accepted. Furthermore, the comparison research work in the current literature, evaluating the results obtained from these approaches, is either outdated or limited to a particular application domain. In this paper, a comprehensive review of the proposed multivariate morphological frameworks is provided. In particular, they are examined mainly with respect to their data ordering methodologies. Additionally, the results of a brief series of illustrative application oriented tests of selected vector orderings on colour and multispectral remote sensing data are also discussed.

© 2007 Published by Elsevier Ltd on behalf of Pattern Recognition Society.

Keywords: Multivariate mathematical morphology; Vector ordering; Vectorial processing; Colour images; Multispectral images

1. Introduction

The mathematical morphology (MM) theory, founded by Matheron [1] and Serra [2,3], is a powerful image analysis framework, nowadays fully developed for both binary and greyscale images. Its popularity in the image processing community is mainly due to its rigorous mathematical foundation as well as its inherent ability to exploit the spatial relationships of pixels. The morphological framework provides a rich set of tools able to perform from the simplest to the most demanding tasks: noise reduction, edge detection, segmentation, texture and shape analysis, etc. As a methodology, it has been applied to almost all application areas dealing with digital image processing [4]. Consequently, it was only a matter of time before attempting to extend the same concepts to colour and more generally multivalued images.

Unfortunately, this extension is not straightforward. Specifically, the morphological framework is based on complete lattices, thus in order to accommodate multivalued images, a way

of calculating the extrema of vectorial data is essential. Yet unlike scalars, there is no unambiguous way of ordering vectors. And besides, most of the known vector ordering schemes have already been employed for defining multivariate morphological operators. However, none of them has yet been widely accepted.

A few surveys on the different approaches have appeared occasionally in the literature [5,6]; although of excellent quality, they have become nevertheless relatively outdated and miss the newest developments. In the light of these remarks, the primary contribution of this paper is to provide an up to date overview of the existing approaches used for extending the fundamental operators of MM to multivariate data; and that is why cases limited only to specific operators have been omitted (e.g. colour image segmentation). Due to their rich diversity and overwhelming number, they are examined according to their methodology of ordering their multivalued input.

The secondary contribution of our work, differentiating it from previous surveys on this topic, is in terms of an application oriented comparative study of the major approaches. More precisely, the available comparison research work, is restricted either to a particular application (e.g. noise reduction) or image type (e.g. colour images) or vector space [5,7,8]. In this paper,

* Corresponding author. Tel.: +33 3 90 24 44 99; fax: +33 3 90 24 44 55.

E-mail address: aptoula@lsiit.u-strasbg.fr (E. Aptoula).

however, for the sake of objectivity, we provide the results of three types of tests realised with two kinds of multivalued data. In particular, we used colour images for noise reduction and texture classification, as well as multispectral remote sensing image data for pixel classification. The tests were conducted with the end of both comparing the suitability of different vector ordering schemes as basis for morphological operators, and for asserting their theoretical properties with the help of experimental results.

The organisation of the paper is as follows. The next section introduces the relevant fundamentals of extending morphological operators to multivariate images, whereas Section 3 presents the current implementation variants of multivariate morphological frameworks. A discussion on the results of a series of illustrative application oriented tests of selected vector orderings takes place in Section 4; and finally Section 5 is devoted to concluding remarks.

2. Extension of MM to multivariate data

In this section, we recall briefly the theoretical concepts behind the extension of morphological operators to multivariate images. For an in-depth study of the theory behind multivariate MM the reader can refer to Refs. [6,9].

2.1. Orderings

As the concept of order plays a central role in this paper, we start by recalling the relative definitions. A binary relation $\mathcal{R}$ on a set $\mathcal{S}$ is called:

- *reflexive* if $x\mathcal{R}x$, $\forall x \in \mathcal{S}$,
- *anti-symmetric* if $x\mathcal{R}y$ and $y\mathcal{R}x \Rightarrow x = y$, $\forall x, y \in \mathcal{S}$,
- *transitive* if $x\mathcal{R}y$ and $y\mathcal{R}w \Rightarrow x\mathcal{R}w$, $\forall x, y, w \in \mathcal{S}$,
- *total* if $x\mathcal{R}y$ or $y\mathcal{R}x$, $\forall x, y \in \mathcal{S}$.

A binary relation $<$ that is reflexive and transitive is called a *pre-ordering* (or *quasi-ordering*); if the anti-symmetry constraint is also met, it becomes an *ordering*. If additionally the totality statement holds for $<$, it is denoted as *total*, if not *partial*.

2.2. Complete lattices and multivariate morphology

The complete lattice theory is widely accepted as the appropriate algebraic basis for MM. Besides unifying the approaches previously employed in binary and greyscale morphology, it also makes it possible to generalise the fundamental concepts of morphological operators to a wider variety of image types and situations. Extensive details on the lattice based description of the MM theory can be found in Refs. [3,9–13]. As a remark, we should also add that the scope of MM has also been extended to complete semilattices, that are more general than complete lattices [14–16].

Specifically, a *complete lattice* $\mathcal{L}$ is a non-empty set equipped with a partial ordering $<$, such that every non-empty subset $\mathcal{P}$ of $\mathcal{L}$ has a greatest lower bound $\bigwedge \mathcal{P}$, called

infimum, and a least upper bound $\bigvee \mathcal{P}$, called *supremum*. In this context, images are modelled by functions mapping their domain space $\mathcal{E}$, an arbitrary non-empty set that is an abelian group with respect to $+$, into a complete lattice $\mathcal{T}$, defining the set of possible “grey values”. Moreover, if $\mathcal{F}$ represents the set of functions $f : \mathcal{E} \rightarrow \mathcal{T}$, then for the partial ordering:

$$f, g : \mathcal{E} \rightarrow \mathcal{T}, \quad f < g \Leftrightarrow \forall x \in \mathcal{E}, \quad f(x) < g(x). \quad (1)$$

$\mathcal{F}$ also forms a complete lattice, where “ $f(x) < g(x)$ ” refers to the partial ordering in $\mathcal{T}$. In other words a complete lattice structure is imposed on the pixel intensity range. Usually $\mathcal{E}$ (the space of pixels) is taken to be either $\mathbb{R}^d$ (d -dimensional Euclidean space) or $\mathbb{Z}^d$ (d -dimensional discrete space), hence $\mathcal{F}$ corresponds, respectively, to the set of continuous or discrete images. Likewise various choices are available for $\mathcal{T}$, such as $\mathcal{T} = \mathbb{R}^n$ and $\mathcal{T} = \mathbb{Z}^n$. The case of $n > 1$ corresponds to the so-called multivalued images [17]. Namely, in the case of a multivalued image with n components, $\mathcal{F} = \mathcal{T}_1 \times \dots \times \mathcal{T}_n$ is considered as the Cartesian product of n complete lattices, and each mapping $f_i : \mathcal{E} \rightarrow \mathcal{T}_i$, $i \in \{1, \dots, n\}$ is called a *channel* or *band* of the multivalued image.

Within this model, morphological operators are represented by mappings between complete lattices (i.e. the input and output images) with some additional properties such as increasingness and translation invariance. They are employed in combination with matching patterns, called *structuring elements* (SE), that are usually subsets of $\mathcal{E}$ (i.e. *flat* SE). Particularly, *erosion* and *dilation* constitute the fundamental blocks of MM, from the combinations of which several sophisticated operators can be derived. More precisely, given two complete lattices $\mathcal{L}$ and $\mathcal{M}$, from an algebraic point of view, an operator $\varepsilon : \mathcal{L} \rightarrow \mathcal{M}$ is called an erosion, if it is distributive over infima, i.e. $\varepsilon(\bigwedge_i P_i) = \bigwedge_i \varepsilon(P_i)$ for every collection $\{P_i\}$ of elements of $\mathcal{L}$. Similarly, $\delta : \mathcal{L} \rightarrow \mathcal{M}$ is called a dilation, if it is distributive over suprema, i.e. $\delta(\bigvee_i P_i) = \bigvee_i \delta(P_i)$ for every collection $\{P_i\}$ of elements of $\mathcal{L}$. As suggested in Ref. [2], dilation and erosion basically rely on three concepts: a ranking scheme, the extrema derived from this ranking and finally the possibility of admitting an infinity of operands. Yet, the first two are missing from multivalued images.

For example, if we apply the preceding notions to the case of continuous multidimensional greyscale images ($f : \mathbb{R}^d \rightarrow \mathbb{R}$), it suffices to replace the partial ordering $<$ of $\mathcal{T}$ with the usual comparison operator in $\mathbb{R}$, in order to induce a complete lattice structure on $\mathcal{F}$ and subsequently on $\mathcal{F}$ by means of Eq. (1), which will make the computation of extrema possible during erosion and dilation. Likewise, the inclusion operator “ $\subset$ ” can be used with binary images ($f : \mathbb{R}^d \rightarrow \{0, 1\}$). However, if we now consider multivalued images ($f : \mathbb{R}^d \rightarrow \mathbb{R}^n$, $n > 1$), it becomes problematic to find an ordering relation for the vectors of $\mathbb{R}^n$, due to the fact that there is no universal method for ordering multivariate data.

In order to remedy this inconvenience, in the classic paper of Goutsias et al. [6] it is proposed to employ an adequate surjective mapping h to transform the image data into a more “suitable” space for morphological operators. More precisely, the idea of using a surjective mapping $h : \mathcal{F} \rightarrow \mathcal{L}$, where

$\mathcal{T}$ is a non-empty set and $\mathcal{L}$ a complete lattice, constitutes the theoretical support upon which several of the present multivariate morphological frameworks are based. Specifically, its importance lies in the fact that $\mathcal{T}$ is no longer required to be a complete lattice, since the ordering of $\mathcal{L}$ can be induced upon $\mathcal{T}$ by means of h :

$$\forall t, t' \in \mathcal{T}, \quad t \leq_h t' \Leftrightarrow h(t) \leq h(t') \quad (2)$$

hence making it possible to construct h -morphological operators on $\mathcal{T}$. Consequently, one can deal with multivalued images, $f: \mathcal{E} \rightarrow \mathbb{R}^n$, through the use of a well chosen mapping $h: \mathbb{R}^n \rightarrow \mathcal{L}$, where $\mathcal{L}$ is a new, more suitable space for lattice based operations [6,17].

Besides, given an adequate vector ranking scheme, the vectorial erosion (ε_b) and dilation (δ_b) of a multivalued image $\mathbf{f}$ by a flat SE b , can be expressed immediately by means of the vectorial extrema operators $\sup_v$ and $\inf_v$ based on the given ordering:

$$\varepsilon_b(\mathbf{f})(\mathbf{x}) = \inf_v \{ \mathbf{f}(\mathbf{x} + \mathbf{s}) \}, \quad (3)$$

$$\delta_b(\mathbf{f})(\mathbf{x}) = \sup_v \{ \mathbf{f}(\mathbf{x} - \mathbf{s}) \}. \quad (4)$$

Therefore, the main obstacle preventing the extension of morphological operators to multivalued images, consists in defining an ordering relation that will induce a complete lattice structure on the set of vectorial pixel intensities.

2.3. Vector orderings

Especially in the last few decades, a lot of effort has been put in engineering a way of ordering vectors. Although numerous techniques for ordering multivariate data can be found in the literature [18–23], according to the classical paper of Barnett [24], they can be classified into one of the following groups.

Marginal ordering (M-ordering): Which corresponds to univariate orderings realised on every component of the given vectors:

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{R}^n, \quad \mathbf{v} \leq \mathbf{v}' \Leftrightarrow \forall i \in \{1, \dots, n\}, \quad v_i \leq v'_i. \quad (5)$$

Data is ordered along each one of its channels independently from others. Hence also the name *componentwise* ordering.

Conditional (sequential) ordering (C-ordering): In which vectors are ordered by means of some of their marginal components, selected sequentially according to different conditions. Whereas the sequentially not participating in the comparison process are listed according to the position of their ranked counterparts. Hence, the ordering of the vectors is *conditioned* upon the particular marginal set of ranked components. Lexicographical ordering constitutes a widely known example of C-ordering employing potentially all the available components of the given vectors:

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{R}^n, \quad \mathbf{v} \leq_L \mathbf{v}' \Leftrightarrow \exists i \in \{1, \dots, n\}, \quad (\forall j < i, \quad v_j = v'_j) \wedge (v_i \leq v'_i). \quad (6)$$

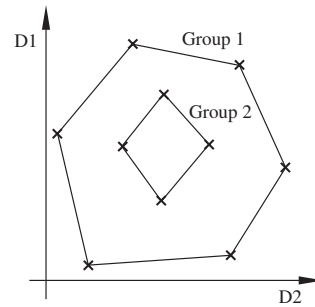


Fig. 1. A P-ordering example in a bi-dimensional space, based on the “peeling” principle.

But of course one can also restrict the comparison process to use only a subset of the available components, as in Ref. [18]. C-orderings are most suitable for cases where one can establish a priority among the image channels.

Partial ordering (P-ordering): In this case, “partial” is an abuse of terminology (Section 2.1), since there are total orderings belonging to this particular class (Section 3.3). That is why we will use the term P-ordering. P-orderings consist simply of pre-orderings that partition the given vectors into groups of equivalence with respect to order, rank or extremeness [25]. They are generally geometric in nature and account well for the inter-relations between components.

A simple example of P-ordering is the one based on the “peeling” of a multivariate sample (Fig. 1). First the convex hull of the entire sample is calculated. Consequently, the points on the border of the convex hull constitute the first group. The points belonging to the border of the convex hull of the interior constitute the second, and so forth. As a result, the entire sample becomes partitioned and ranked according to their group number, but of course no internal distinction is in place for the contents of the groups.

Reduced ordering (R-ordering): In which vectors are first reduced to scalar values and then ranked according to their natural scalar order. A further categorisation of R-orderings consists in classifying them as distance orderings and projection orderings [26]. For instance, a R-ordering on $\mathbb{R}^n$ could consist in defining first a transformation $h: \mathbb{R}^n \rightarrow \mathbb{R}$, and then ordering the vectors of $\mathbb{R}^n$ with respect to the scalar order of their projection on $\mathbb{R}$ by h :

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{R}^n, \quad \mathbf{v} \leq \mathbf{v}' \Leftrightarrow h(\mathbf{v}) \leq h(\mathbf{v}'). \quad (7)$$

According to the chosen transformation it is possible to obtain a total pre-ordering (h non-injective) or even a total ordering (h injective) [27]. An additional advantage of R-orderings lies in the fact that with an adequately chosen h , they can attribute equal priority to all components, unlike C-orderings.

As the aforementioned ordering groups are not mutually exclusive, their combinations as well as numerous implementational variants have led to several morphological frameworks with diverse properties. The next section will elaborate on the different approaches.

3. Approaches to multivariate MM

Despite the rich variety of multivariate morphological frameworks, there are in fact two main variables that are modified at each case, the extrema calculation method and the transformation (e.g. domain space change, etc.) if any, that takes place on the image data before ranking; undoubtedly both influence the properties of the resulting operators.

In this section, we review in relative detail the approaches used for implementing multivariate morphological operators, primarily according to their vector ordering scheme, along with additional comments on their use with different preprocessing methods. However, instead of the conventional ordering categorisation presented in Section 2.3, it was decided to adopt the scheme employed by Chanussot in Ref. [27] where they are classified according to their algebraic properties, as given in Section 2.1. This choice was made with the aim of underlining the effect that these basic properties have on the end result. Additionally, the vectors of the ordering relations that are mentioned in the sequel, are considered in $\mathbb{R}^n$, unless otherwise specified.

3.1. Processing strategies

Given a multivalued image, in practice there are two general methods of morphological processing: *marginal* (or *component-wise/scalar*) and *vectorial*.

Marginal processing: It consists in processing separately each channel of the image. The inter-channel correlation is totally ignored, along with all information that could be potentially used in order to improve the quality of the result. Furthermore, the repetition of the processing procedure for each channel renders it expensive in terms of computational complexity. On the other hand, the marginal approach makes it possible to employ directly all methods offered by greyscale morphology (Fig. 2 left).

Vectorial processing: This main alternative of the marginal approach, as its name implies, processes all available channels globally and simultaneously. Given that the vectorial pixels are considered as the new processing units, the correlation among the different channels is no longer ignored (Fig. 2 right). However, when compared to its marginal counterpart, the most important inconvenience of the vectorial approach appears to be primarily the need for adapting the existing algorithms in order to accommodate vectorial data; thus leading often to slower implementations than their scalar versions. The rest of this section elaborates on each approach as well as on their variations.

3.2. Partial ordering based approach (marginal)

The so-called *marginal* processing strategy, despite being presented usually as an alternative to *vectorial*, is as a matter of fact no more than just its variation, as it employs the partial ordering defined in Eq. (5). Obviously, there can be vectors that may not be comparable under this ordering relation, for instance $\mathbf{a} = [7, 2]^T$ and $\mathbf{b} = [3, 4]^T$. Nevertheless, this does not prevent the definition of valid morphological operators based on extrema computed by means of this ordering [9]. Furthermore, it makes it possible to employ all tools offered by greyscale morphology with no need for special adaptation steps. For instance, the erosion and dilation expressions given in Section 2.2 become equivalent to:

$$\varepsilon_b(\mathbf{f})(\mathbf{x}) = [\varepsilon_b(f_1)(\mathbf{x}), \dots, \varepsilon_b(f_n)(\mathbf{x})]^T, \quad (8)$$

$$\delta_b(\mathbf{f})(\mathbf{x}) = [\delta_b(f_1)(\mathbf{x}), \dots, \delta_b(f_n)(\mathbf{x})]^T, \quad (9)$$

where ε_b and δ_b denote, respectively, the scalar erosion and dilation operators with a SE b . For instance, Gu in Ref. [28] has employed the marginal approach for establishing multivalued morphological operators applied to moving object segmentation and tracking. Another example is given by Aptoula et al. in Ref. [29], where marginal morphological operators are employed with the purpose of galaxy detection from multispectral data. Moreover, a generalisation of the marginal approach was given by the matrix morphology theory of Wilson [30], which was based on the work of Heijmans and Ronse [11].

Despite its implementational simplicity, marginal ordering suffers mainly from two disadvantages: not accounting for inter-channel information as well as the risk of altering the spectral composition of its input. More precisely, as each component is processed independently, any eventual correlation among them is totally ignored, hence rendering this approach unsuitable for images with highly correlated components (e.g. RGB colour images) [31]. A possible solution to this problem, as proposed in Ref. [6], consists in applying a decorrelating transformation (e.g. maximum noise fraction transform (MNF), principal component analysis (PCA), discrete cosine transform, a proper colour space, etc.) prior to ordering.

An application of this idea to the detection of land mines, in combination with the MNF transform, can be found in Ref. [32]. Or in Ref. [33], where hyperspectral remote sensing images are classified by means of morphological profiles computed marginally from the first two principal components of the input. Nevertheless, these decorrelating transformations also introduce an additional computational burden.

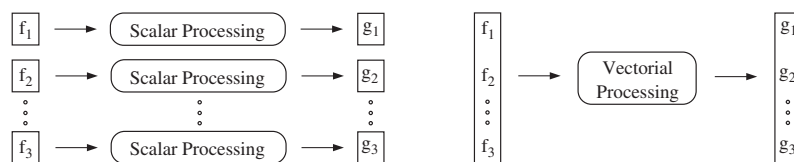


Fig. 2. Marginal (left) and vectorial processing (right) strategies.

Furthermore, there is absolutely no guarantee that marginally processed vectors belong to the input image. The lack of vector preservation constitutes an undesirable effect for several applications. For example, in the case of colour image processing this would lead to the appearance of new colours (also known as *false colours*) and thus deteriorate the visual quality of the result, and in particular the colour balance and the object boundaries; whereas the effect near the spatial edges, leads to the so-called *edge jitter*. Similarly, it can be harmful for the spectral signature based classification of remote sensing images, where an altered spectral composition can damage the entire process.

A detailed study of the problem of vector preservation can be found in Refs. [34,35]. According to the last reference, the only way to use morphological operators without generating new vectors is to impose an order on the vector space by means of an (pre-)ordering verifying the totality constraint; besides, this is why the majority of the published articles on multivalued morphology deal with total (pre-)orderings. Yet, vector preserving approaches are in the same time limited by this property, due to the restriction imposed on their output (i.e. the output must be a vector from the input set); whereas the marginal approach has access to a much broader range of output values, an advantage which becomes most valuable during noise reduction [5].

Of course, there are also situations such as the one presented in Ref. [36], where the inter-channel correlation level is low (near and mid infra-red bands) and vector preservation is not a necessity; in which case the marginal operators used for detecting the dim point targets on maritime backgrounds perform adequately. Additional attempts motivated by the intuitiveness of the marginal approach have resulted in improvements. Namely, Serra in Ref. [9] presents an intermediate form between a M-ordering and a C-ordering, while an application of marginal ordering in the HSV colour space is introduced by Weber and Acton in Ref. [37], where the aforementioned undesirable effects are reduced by shifting the hue origin. Moreover, Al-Otun in Ref. [38], proposes the *corrected componentwise morphological algorithm* with the end of preventing the appearance of new vectors. It consists in replacing each new vector of the output with its closest vector from the original image, chosen with the help of a Mahalanobis distance based error function.

In brief, either accompanied by additional transformations or not, the marginal processing strategy uses conventional greyscale operators and pixels are still treated as scalar values.

3.3. Total pre-ordering based approach

Contrarily to the marginal approach, by means of the additional property of totality, all vectors become comparable and as a result it is possible to construct a totally ordered lattice structure. Hence pixels can be manipulated as whole vectors, and consequently the risk of altering the pixel composition of the input is eliminated [35].

All pre-orderings, however, share a common drawback, which is the relaxation of the anti-symmetry constraint. Thus distinct vectors can eventually end up being equivalent. That is why additional measures become necessary, in order to resolve

the ambiguity of eventually multiple extrema. For instance a partial, but considerably useful in practice solution proposed by Comer and Delp in Refs. [5,39] against this problem, consists in selecting the output vector according to its position in the SE.

Reduced total pre-orderings: Total pre-orderings can be obtained with R-orderings employing a non-injective reduction transformation [27]. Distance measures are typical examples of such transformations; in fact they account for the majority of the proposed R-orderings.

This first variant, ranks a family $\{\mathbf{v}_j\}$ of vectors according to their distance from a reference vector $\mathbf{v}_{ref}$:

$$\forall \mathbf{v}_k, \mathbf{v}_l \in \{\mathbf{v}_j\}, \quad \mathbf{v}_k \leq \mathbf{v}_l \Leftrightarrow d(\mathbf{v}_k, \mathbf{v}_{ref}) \leq d(\mathbf{v}_l, \mathbf{v}_{ref}), \quad (10)$$

where $d(\cdot, \cdot)$ represents a distance measure. An example of this type of R-ordering, as applied to colour images, can be found in Ref. [40]. In case the reference vector is the origin, expression (10) becomes equivalent to using the norms of the vectors. A variation of expression (10), which eliminates the need for a reference vector, consists in associating each vector with the sum of its distances from the other vectors:

$$\forall \mathbf{v}_k, \mathbf{v}_l \in \{\mathbf{v}_j\}, \quad \mathbf{v}_k \leq \mathbf{v}_l \Leftrightarrow \sum_j d(\mathbf{v}_k, \mathbf{v}_j) \leq \sum_j d(\mathbf{v}_l, \mathbf{v}_j). \quad (11)$$

This additional advantage, however, comes at a high cost, in terms of both computational complexity and theoretical properties. Specifically, although this approach can be used for computing multivariate extrema, it does not constitute an ordering in the sense of a binary relation, hence leading to pseudo-morphological operators. Furthermore, the infimum of a family of vectors calculated in this way, corresponds to the notion of “median vector” and consequently does not carry the significance of a “minimum” in the numerical sense. As a remark, we should also add that in situations where the directions of vectors gain more importance than their magnitude, Eqs. (10) and (11) may also be used by replacing the distance d with its angular counterpart, as proposed in Ref. [41].

There is also another property worth mentioning, referenced in [42], concerning Eq. (11) and implementations of Eq. (10), if the reference vector is chosen to be “in the middle” (e.g. the median or the average). It concerns the instability of the supremum computed by means of these orderings. Despite the remarkable stability of the infimum under the same conditions, even a slightly varied input can radically change the least upper bound. Consequently, the dilation operator that makes use of the supremum becomes unsound from a practical point of view.

The choice of the distance measure is of course another key topic. Theoretically, any kind of metric can be used, and so it is in practice. For instance, in an application of Eq. (11) to vector area morphology in Ref. [43], the city-block and Euclidean distances are considered. Whereas, in Ref. [44] the same reduction scheme is employed in combination with spectral angle distances to the end of creating morphological profiles for hyperspectral images. Moreover, Al-Otun [38,45] uses Eq. (10) coupled with the Mahalanobis distance for ordering

the components of colour images in different colour spaces, with “black” being the reference vector. An example of Euclidean norm based ordering can be found in Ref. [5], as applied to RGB colour images.

Moreover, an original approach implemented in the polar version of the C–Y colour space appears in Ref. [46], where a saturation based combination of the hue and intensity planes is proposed. Specifically, the vectorial pixels contained within the SE are reduced into scalars by means of a weighted combination of vector angle (d_a) and Euclidean distances (d_E) from the vector $\mathbf{v}_c$ at the centre of the SE:

$$h(\mathbf{v}) = w(\mathbf{v}, \mathbf{v}_c) \cdot d_a(\mathbf{v}, \mathbf{v}_c) + (1 - w(\mathbf{v}, \mathbf{v}_c)) \cdot d_E(\mathbf{v}, \mathbf{v}_c), \quad (12)$$

where $w(\cdot, \cdot)$ denotes the weight computed based on the saturation levels of the given vectorial pixels. The underlying idea is to favour hue differences (d_a) when saturation levels are high, while intensity differences (d_E) gain more importance with achromatic pixels.

Of course the non-injective transformation choices are by no means limited by these few variants. For instance, in Ref. [47], vectors are ordered based on the number of times they appear in the input image. Nevertheless, distance based R-orderings hold the potential of accounting for all dimensions without privileging any of them, a property which becomes particularly useful in case there is no predefined order of importance among the available channels (e.g. RGB colour images). Otherwise any transformation capable of realising the necessary reduction may be used. For example in Ref. [5], an additional R-ordering is presented, applied to RGB colour images, where vectors are reduced to scalars with the help of a linear weighted combination:

$$h(\mathbf{v}) = w_1 \cdot v_1 + \dots + w_n \cdot v_n. \quad (13)$$

Hence making possible the arbitrary prioritisation of colour channels by means of the $w_i \in \mathbb{R}$, $i \in \{1, \dots, n\}$ coefficients. Finally, a more complicated ordering scheme takes place in Ref. [48], where a new R-ordering based on ordinal first PCA is introduced and the derived operators are applied to edge detection on RGB colour images.

Conditional total pre-orderings: As previously mentioned in Section 2.3 C-orderings restrict the ordering process to only one or more components of the given vectors, while the others are conditioned upon them. That is why C-orderings are suitable for situations where certain channels are more “privileged” than others. Besides, unless all vector components participate in the ordering process, the resulting C-ordering is bound to be a total pre-ordering, thus sharing their aforementioned inconveniences. For example, in the case where only the first component is employed [18]:

$$\mathbf{v} \leq \mathbf{v}' \Leftrightarrow v_1 \leq v'_1. \quad (14)$$

Hence, the two distinct vectors $\mathbf{a} = [7, 2]^T$ and $\mathbf{b} = [7, 3]^T$ would be considered equivalent according to Eq. (14). The main problem of C-orderings concerns of course the choice of the ordered components. Obviously, ordering vectors along only some of their components is practically justifiable only if the given components represent sufficiently the vectors.

In the general case, this approach attributes far too much significance to the selected components, disregarding all others. While usually this constitutes a severe problem, there are cases where it becomes invaluable. For instance, when it is a priori known that certain channels are noise free. In fact, C-orderings are almost always used in combination with a suitable domain space change, so that the data of interest, in its majority, will lie in only some of the channels.

An application example can be found in Ref. [49], where a C-ordering on the HSV colour space is employed, ignoring the hue component. In fact, this also constitutes a fine example of a combined ordering; more precisely, the result can also be considered as a P-ordering where the ordered groups contain the colours of equal value and saturation with no internal distinction whatsoever as hue is not taken into account. An application of the same ordering on colour image skeletonisation can be found in Ref. [50].

3.4. Total ordering based approach

Total orderings, from a theoretical point of view, have two main advantages that render them more suitable for vector ordering, as far as multivariate MM is concerned. First, thanks to their totality, they are vector preserving, and contrarily to pre-orderings, as they verify the anti-symmetry constraint the computed extrema are unique. That is why the majority of the attempts concentrated on extending morphological operators to multivalued images are based on total orderings. In particular, the lexicographical ordering (C-ordering) along with its variants is among the most implemented choices.

However, the uniqueness of extrema takes a serious toll, since the prioritisation of certain vector components becomes inevitable [27]. That is why they are almost always used in combination with a suitable domain transformation (e.g. HSV, $L^*a^*b^*$, PCA, etc.) that will place the “interesting” part of the data in the first few channels. Nevertheless, as it will be subsequently presented some implementations tend to be more “symmetric” than others.

Lexicographical ordering: Lexicographical ordering, introduced in Eq. (6), is undoubtedly the most widely employed total ordering within this context. As a conditional ordering, it is most suitable to situations where an order of “importance” exists on the available channels, either inherently or artificially created by means of an appropriate transformation.

Fig. 3 provides an example of the priority attributed to the first component during lexicographical ordering. More precisely, a vectorial dilation is applied on a RGB colour image (Fig. 3, left) and as red is the head component (dark grey), it dominates visibly over green (light grey) (Fig. 3, middle). Whereas if we permute the channels as GRB, the effect is reversed in favour of green (Fig. 3, right).

Lexicographical ordering has been mostly employed in the domain of colour MM, where several colour spaces exhibit an inherent prioritisation among their channels as far as human observers are concerned. Specifically, as the human vision system is more sensitive to brightness changes than to chromatic changes, brightness components tend to be privileged [51]. Hue

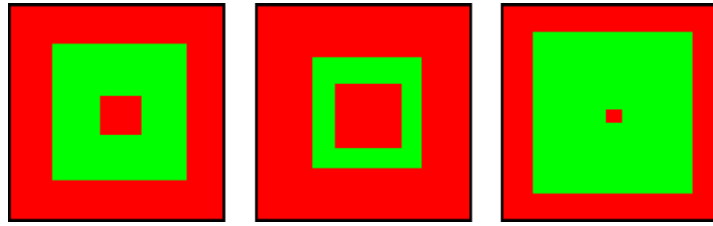


Fig. 3. Original image (left), results of applying a vectorial dilation based on a lexicographical ordering (RGB—middle) (GRB—right) with a 21×21 square SE.

based spaces in particular have been extensively studied in the literature as they are closely related to the way which humans perceive colour. It should be additionally noted that the hue component poses considerable problems as it is a 2π periodical value and thus cannot be intuitively ordered simply from smallest to largest [52,53]. An example of luminance based lexicographical ordering on the HSV colour space can be found in Ref. [54], whereas Refs. [55,56] use the same ordering, respectively, for median filtering and granulometry calculation on colour images. Of course there can also be specific situations where chromatic information is more significant, for instance in Ref. [57], hence hue is compared first in the lexicographical cascade.

Other applications of lexicographical ordering in the HSI space include Ref. [58], where reconstruction based morphological operators are developed with the end of brightness elimination, using an intensity based ordering scheme. The same principle is additionally employed in Ref. [59] with the purpose of noise elimination. The perceptually uniform $L^*a^*b^*$ [60] and $L^*u^*v^*$ [53] spaces have also been used together with a lexicographical ordering. Moreover, a thorough study of the potential of this ordering in the HLS space is provided in Ref. [61]; while in Ref. [62], the use of lexicographical ordering in the improved HLS (IHLS) space is explored. And finally, the results of a comparative application of this ordering in different hue based colour spaces can be found in Ref. [8].

As the majority of lexicographical comparisons are determined by the first components [60], variations of the classical lexicographical ordering were proposed, with the end of better tuning the priority as well as degree of influence of each component. A non-exhaustive list of such extensions follows. A first group of variations is based on the use of additional components during comparison. Rivest in Refs. [63,64] has exploited the symmetrical nature of a norm based R-ordering by placing it to the first position of the lexicographical cascade, with the end of developing morphological operators for complex signals:

$$\mathbf{v} \leqslant \mathbf{v}' \Leftrightarrow [\|\mathbf{v}\|, v_1, \dots, v_n]^T \leqslant_L [\|\mathbf{v}'\|, v'_1, \dots, v'_n]^T, \quad (15)$$

where $\leqslant_L$ is the lexicographical ordering. Similar variants to Eq. (15) have been proposed in Refs. [40,65], with the difference of placing a reference vector based distance measure in the position of the first component. Of course, there is no limit to the number or type of functions that can be used

according to this principle. Other examples include the use of the maximum and minimum of the compared components in the case of RGB colour images, as well as their weighted combinations [66]. Another type of extension to the classical lexicographical ordering consists in using of a user defined parameter α in such a way that it can modify the degree of influence of the first component. For instance, the α -modulus lexicographical ordering, introduced by Angulo [67,68]:

$$\begin{aligned} \forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \\ \mathbf{v} \leqslant \mathbf{v}' \Leftrightarrow [\lceil v_1/\alpha \rceil, v_2, \dots, v_n]^T \leqslant_L [\lceil v'_1/\alpha \rceil, v'_2, \dots, v'_n]^T \end{aligned} \quad (16)$$

was implemented on a variation of the HLS colour space. Specifically, by dividing the first component with a parameter α , and then rounding it off to the next closest integer, a sub-quantisation of the first vector dimension is realised, hence forming larger equality groups within this dimension. Consequently, a higher number of comparisons is expected to reach the second dimension.

Bit mixing based ordering: Bit interlacing (or mixing) constitutes an innovative R-ordering, aiming mainly to eliminate the unavoidable asymmetry, which results from the application of total orderings. Specifically it employs an injective transformation exploiting the binary representation of each component in order to impose a total order on the vector space [27,69,70].

Given a vector $\mathbf{v}$, with each component coded in k bits, the corresponding reduction transformation $h: \mathbb{Z}^n \rightarrow \mathbb{Z}$ is formulated as

$$h(\mathbf{v}) = \sum_{m=1}^k \left\{ 2^{n \cdot (k-m)} \cdot \sum_{i=1}^n 2^{n-i} \cdot v_{i,m} \right\}, \quad (17)$$

where $v_{i,m}$ denotes the m th bit of the i th component of $\mathbf{v}$. Hence, the resulting binary representation of $h(\mathbf{v})$ becomes

$$v_{1,1}v_{2,1} \dots v_{n,1}v_{1,2}v_{2,2} \dots v_{n,2} \dots v_{1,k}v_{2,k} \dots v_{n,k}. \quad (18)$$

Besides being endowed with all the qualities of a total ordering, bit mixing provides a more symmetrical approach than its lexicographical counterpart as dimensions are mixed in bit level. Of course some dimensions continue to be more privileged than others, with the degree of that privilege being proportional to the significance of the bit position that they occupy. Furthermore, a finer grained symmetry can be obtained by

Table 1

Synoptic table of references developing multivariate morphological operators, organised according to the properties of the employed ordering scheme and its corresponding implementation space (VP: vector preserving, PB: prioritisation of bands)

Data space	Properties		
	$\neg$ VP	VP $\wedge$ PB	VP $\wedge \neg$ PB
Generic	[5,6,28,85]	[65,69,70,83,85–87]	[5,38,43,44,47,48,65,72,76,82,84]
HSV	[37]	[49,50,54–57,77,88]	
L*a*b*		[60]	
L*u*v*		[53]	
(I)HLS		[61,62,66,68,81]	
HSI		[58,59,89]	
C–Y		[40,46]	
Complex		[63,64]	[81]
Matrix			[78–80]

modifying the original mix order [27]. Then again, there are several situations where data channels need to be processed with a certain priority. Bit mixing can easily respond to this requirement by placing the important vector components to more significant bit positions. Additionally, bit interlacing can also be used as a means of fusing the channels, hence allowing for instance the use of the watershed transformation on colour images [71].

From a theoretical point of view, this approach aims to fill a given multi-dimensional space using a “balanced” *space filling curve* (SFC) with respect to the available dimensions. In the case of total orderings, these curves (e.g. Peano curve) pass through all vector coordinates of the space under consideration, hence vectors can be ordered according to their position on it. An additional approach developed based on SFC can be found in Ref. [72]. On the other hand, according to Ref. [73] the main inconvenience of a total ordering obtained in this fashion is its lack of physical interpretation.

3.5. Other approaches

Besides the previously presented ordering methodologies, there have been also some rather “unconventional” orderings or extremum calculation approaches, in the sense that they either do not consist in simply using a standard vector ordering scheme, or they are developed for a particular form of image data, or even combine additional theories with the end of achieving an efficient solution.

For example, following the relatively recent success of fuzzy morphology with greyscale images [74,75], a couple of attempts to apply the fuzziness concepts to multivariate images have already been carried out. Köppen et al. in Ref. [76], introduced *fuzzy Pareto morphology*, a means of computing multivariate extrema based on the notion of Pareto sets, a concept belonging to the field of multi-criteria optimisation, combined with fuzzy subethood. A second attempt was made by Louverdis et al. [77], who employ a direct fuzzification of the vectorial pixels in the HSV colour space, that are subsequently ranked according to a lexicographical ordering scheme. The morphological operators based on this total pre-ordering have been compared in the original article with their counterparts in

Ref. [54] resulting from the use of a similar lexicographic ordering also in the HSV space. According to their experiments, the fuzzy version has slightly better noise reduction capabilities as well as less sensitivity to distortions.

Moreover, the domain of medical imaging has recently introduced matrix-valued images, where a positive definite matrix is associated to each voxel. Pioneering work on the application of MM theory to this type of images has been made by Burgeth et al. [78–80], using both analytic (i.e. Loewner ordering) and geometric approaches (i.e. modelling matrices with ellipsoids).

Some further original approaches aiming to compute multivariate extrema, include the one proposed in Ref. [81], developed with the purpose of colour object detection using vector projection measures in combination with vectorial SEs, as well as the graph based methodology introduced by Lezoray et al. [82], employing the minimum spanning tree algorithm among the pixels under the SE. An ordering based on labels has been additionally proposed by Ronse and Agnus [83]. Moreover, Zaharescu et al. have explored in Ref. [84] the potential of the triangle representation of colours, which leads to a R-ordering, while Mojsilovic and Soljanin [85] have employed a method based on a quantisation using Fibonacci lattices, in order to obtain a partial ordering. And finally Gibson et al. [86] have relied on local convex hull computations for locating multivariate extrema, whereas Busch and Eberle [87] have proposed a conditional ordering based on semantic principles.

3.6. Synopsis

This section presented the different approaches that have appeared so far in the literature, with the end of extending morphological operators to multivalued images. A summary of the related references is given in Table 1.

4. Experiment results

Having provided an insight into several multivariate morphological frameworks, in this section a brief series of comparative tests will be carried out with the aim of measuring their relative performances in three image processing tasks.

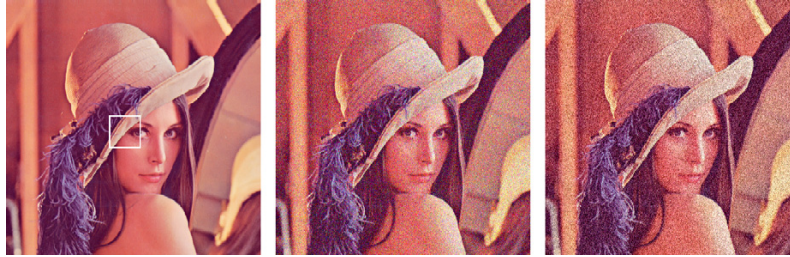


Fig. 4. From left to right: the original image “Lenna”, corrupted with uncorrelated Gaussian noise ($\sigma = 32$, $\rho = 0.0$) and corrupted with correlated Gaussian noise ($\sigma = 32$, $\rho = 0.9$). The white square represents the enlarged results in Fig. 5.

So far, the comparative studies of multivalued morphological processing approaches [5,7,8,90,91] have been restricted to a particular application domain (e.g. colour morphology) or vector space. The reason of these restrictions is the overwhelming number of crucial variables that need to be taken into account in order to achieve a fully objective performance measure (e.g. number of images, type of images, number of channels, correlation of channels, vector space transformations, etc.). That is why the aim of this section is *not* to realise a “benchmark” of the different approaches, but rather to provide indications on their suitability for particular input properties, as well as to assert the remarks made in the previous sections with experimental results.

For the sake of simplicity only certain approaches have been considered during the tests, chosen subjectively as representative members of the classification given in Section 2.1. More precisely, the ordering schemes participating in the tests are: marginal (M), bit mixing (B), lexicographical (L), a variation of lexicographical ordering (rL), Eq. (15), reference vector based distance orderings using the median (R) and the origin (N) as reference vectors, as well as the method based on cumulative distances (C), although it does not qualify as an ordering from an algebraic point of view. The Euclidean distance and L_2 norm are used where necessary. As far as their implementation is concerned, vector comparisons were hard coded without the use of any indexes [35], within a generic Java based framework. The subjects are chosen from the two primary application areas of multivariate morphology, colour and remote sensing imagery. In particular, noise reduction and texture classification tests are realised with the first, while the latter are used for a pixel based classification.

4.1. Noise reduction

During this first test, the relative performances of the vector ordering schemes are examined in terms of noise reduction quality. As a quantitative measure, the *normalised mean squared error* (NMSE) is used:

$$\text{NMSE} = \frac{\sum_{i=1}^N \sum_{j=1}^M \|\mathbf{f}(i, j) - \mathbf{f}'(i, j)\|^2}{\sum_{i=1}^N \sum_{j=1}^M \|\mathbf{f}(i, j)\|^2}, \quad (19)$$

Table 2

NMSE $\times 100$ values for the “Lenna” image obtained against uncorrelated Gaussian noise

Colour spaces	Vector orderings						
	M	L	N	B	C	R	rL
RGB	0.78	2.29	2.10	2.21	3.10	3.07	2.10
GBR	0.78	2.33	2.10	2.19	3.10	3.07	2.10
L*a*b*	0.92	2.23	2.50	2.05	3.13	3.02	2.50
YUV	0.79	2.09	2.12	2.03	2.64	2.59	2.12

Table 3

NMSE $\times 100$ values for the “Lenna” image obtained against highly correlated Gaussian noise

Colour spaces	Vector orderings						
	M	L	N	B	C	R	rL
RGB	0.79	0.97	0.87	0.97	1.35	1.34	0.87
L*a*b*	0.85	0.92	1.34	0.88	1.43	1.38	1.34
YUV	0.79	0.89	0.90	0.88	1.65	1.67	0.90

where N and M represent the image dimensions while $\mathbf{f}(i, j)$ and $\mathbf{f}'(i, j)$ denote, respectively, the vectorial pixels at position (i, j) for the original and filtered images. The tests have been repeated with a number of RGB colour images of various content, however, here we exhibit the results obtained for the “Lenna” image (Fig. 4). The subjects, of size 512×512 pixels and 24 bits per pixel, are first contaminated with zero-mean additive Gaussian noise, $\sigma = 32$ and correlation factor $\rho = 0.0$ (Table 2) and $\rho = 0.9$ (Table 3). Other noise distributions that have also been tested, include double-exponential and uniform, however, as the influence of the type of noise distribution on the resulting performances was observed to be minimal, they were omitted from Tables 2 and 3 for the sake of clarity.

The filter employed for smoothing is open-close close-open (OCCO). More precisely, let first γ_b and ϕ_b denote, respectively, the vectorial opening and closing operators:

$$\gamma_b(\mathbf{f})(x, y) = \delta_b(\varepsilon_b(\mathbf{f}))(x, y), \quad (20)$$

$$\phi_b(\mathbf{f})(x, y) = \varepsilon_b(\delta_b(\mathbf{f}))(x, y). \quad (21)$$

In which case, OCCO is defined as the pixelwise average of open-close and close-open:

$$\text{OCCO}_b(\mathbf{f})(x, y) = \frac{1}{2}\gamma_b(\phi_b(\mathbf{f}))(x, y) + \frac{1}{2}\phi_b(\gamma_b(\mathbf{f}))(x, y) \quad (22)$$

given the size of the input, the SE b is chosen as a cross of size 3×3 . Larger sizes have been also tested but the relative performances remained almost the same. The OCCO filter is chosen primarily due to its effective combination of the basic morphological operators as well as for its suitability against the chosen noise type. On the other hand, taking the average of the intermediate results, obviously prevents the preservation of the initial vectors. Nevertheless, this is a desired property in the context of noise reduction as it makes it possible to better approximate the original noise free values. Different colour spaces have been also employed during smoothing for mainly two reasons; first in order to simulate various correlation levels among the channels as well as for testing the effect that uneven distributions of the intensity information among the channels have on the underlying ordering schemes. Moreover, as hue based colour spaces usually require a reference value, they have been excluded from the experimentation process with the end of avoiding any further parametrisation.

Table 2 shows the NMSE results of the orderings within four colour spaces against uncorrelated Gaussian noise. The entries corresponding to the best performance of their row (colour space) are in bold. According to the obtained values, the overall superiority of the marginal approach over its vectorial counterparts is remarkable.

As stated earlier the reduced smoothing capacity of vectorial approaches is part of the trade-off between noise reduction capability and vector preservation. In brief, even with a maximised processing symmetry among the available channels the final result of vectorial openings and closings in Eq. (22) will necessarily be one of the input vectors. And as such it is natural for the marginal strategy to outperform, by having access to a much broader range of output values, that are not necessarily included, however, in the original image.

A second remark concerns the relatively high error rates of the total orderings in the highly correlated RGB space, which is simply due to the fact that both bit mixing and lexicographical orderings inevitably prioritise the vector components. Consequently, the green and especially blue channels influence the outcome of vector comparisons much less than red, thus resulting in poor smoothing quality in the last two channels. The effect of prioritisation is particularly visible with lexicographical ordering. An exception to this remark is rL, which although total, results always in a NMSE value identical to norm based reduced ordering (N), due to the vector norm that occupies the first position, as indicated in Eq. (15), during the lexicographical comparison of vectors. Therefore, as the vast majority of comparisons is decided by means of the first vector component, from a practical point view it appears to behave exactly as a norm based R-ordering.

As far as the three distance based approaches are concerned (R, C, N), the norm based ordering takes the lead while its distance based counterparts provide quite unsatisfactory results.

Table 4

Relative durations for the application of the OCCO filter on the Lenna image with a 3×3 cross shaped SE

M	L	N	B	C	R	rL
1.36	1.00	1.49	1.79	2.03	2.22	1.54

Thanks to the non-injective scalarisation that takes place, no distinction is made among the channels during comparison. This last important property becomes more evident within the GBR space, a permutation of RGB, where they achieve the exact same performances thus underlining their robustness against situations where it is not a priori known in which channel takes place the majority of the “important” information or even which channel is the most corrupted.

Furthermore, the $L^*a^*b^*$ colour space causes significant changes in the results as it is less correlated than RGB and the brightness information is moved exclusively to the first channel. As expected, the bit mixing and lexicographical orderings improve their smoothing rates substantially, since their prioritisation works this time in their favour by rendering the brightness dimension more “exploitable”. It is also presumed that the finer grained symmetry of the vector components during comparison is what makes it possible for the bit mixing ordering to surpass its lexicographical counterpart. Nevertheless, the marginal ordering still provides the best results.

On the other hand, Table 3 shows the NMSE results obtained against highly correlated Gaussian noise. Although the relative performances of vectorial approaches have stayed the same, this time they are clearly much closer to that of marginal processing.

As far as the computation times are concerned, Table 4 illustrates the relative duration of a single application of the OCCO filter for each ordering. Apparently the lexicographical ordering is the fastest among them, closely followed by the marginal approach, while the R-orderings suffer from the extensive distance calculations.

In conclusion, the main points of this discussion include the overall superiority of the marginal approach, particularly against uncorrelated noise. On the other hand, both marginal and vectorial approaches exhibit similar performances in the case of correlated noise (Fig. 5). Although vectorial approaches tend to be more complex from an implementational point of view, they offer the possibility of smoothing without the risk of introducing new vectors. Additionally, the results have also underlined the importance of the colour space, as it highly influences the performance of the employed ordering.

4.2. Texture classification

As far as texture classification is concerned, here we employ the colour textures of Outex13 (Fig. 6) [93]. The question whether colour should be processed separately or jointly from texture is still an open problem [92], and vectorial morphological feature extraction operators represent in their majority the latter case [94,95]. As colour texture descriptor we employ the morphological version of the autocorrelation operator, namely

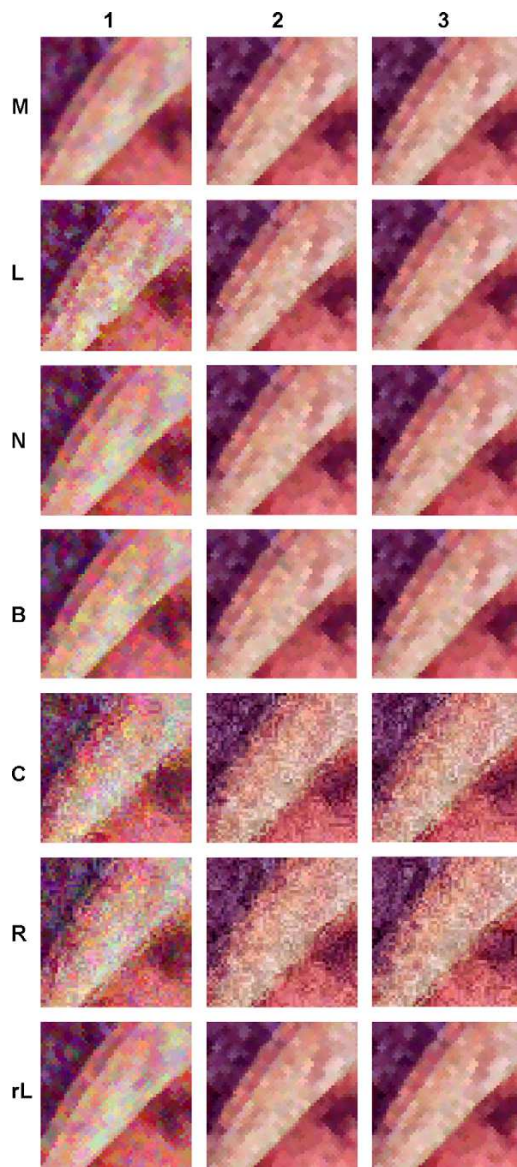


Fig. 5. Enlarged results of the area contained within the white square of Fig. 4, obtained on the RGB colour space against uncorrelated Gaussian noise (column 1), on the RGB colour space against correlated Gaussian noise (column 2), and on the YUV colour space against correlated Gaussian noise (column 3).

morphological covariance. The vectorial version of morphological covariance K' of an image f , is defined as the volume Vol of the image (i.e. sum of pixel values), eroded by a pair of points $P_{2,v}$ separated by a vector $\mathbf{v}$:

$$K'(f; P_{2,v}) = \text{Vol}(\varepsilon_{P_{2,v}}(f)). \quad (23)$$

In practice, K' is computed for varying lengths of $\mathbf{v}$, and most often as also here the normalised version K is used for



Fig. 6. Examples of the 68 textures of Outex 13 [92].

Table 5

Classification rates in % for the textures of Outex13, using vectorial erosion based covariance

Colour spaces	Vector orderings				
	M	L	N	B	rL
RGB	77.65	70.74	71.62	70.86	71.85
L*a*b*	77.76	80.03	77.89	80.13	78.02
YUV	77.66	79.23	77.75	79.43	77.36

measurements:

$$K(f) = \text{Vol}(\varepsilon_{P_{2,v}}(f)) / \text{Vol}(f). \quad (24)$$

Given the resulting K series, one can gain insight into the structure of a texture [96]. In particular, the periodic nature of covariance is strongly related to that of its input, whereas the thinness and coarseness of the textures can be evaluated by studying the sharpness of the periodic patterns within this series. Additional information concerning the anisotropy of the textures can be obtained by plotting against not only different lengths of $\mathbf{v}$, but orientations as well.

Since the SE contains only 2 pixels, the use of R and C approaches is of no practical interest. The covariance based feature vectors have been calculated using four directions for the point pairs (0° , 45° , 90° , 135°), each along with distances ranging from 1 to 49 pixels in steps of size two. Consequently 25 values are available for each direction, making a total of 100 values for every image channel after concatenation. Furthermore, as far as the conversion from RGB to L*a*b* is concerned, the proper transformation matrix calibrated to the CIE A white point of the acquisition apparatus has been used [92]. The classification process is realised using a kNN classifier with $k = 1$.

Table 5 contains the accuracy rates that have been obtained. In general, the relative performances can be considered similar to those obtained during noise reduction. More precisely, marginal processing continues to outperform its alternatives in RGB, once more followed by norm based ordering. Since all channels are equally important in this colour space, both prioritisation attempts with L and B fail to surpass it. However,

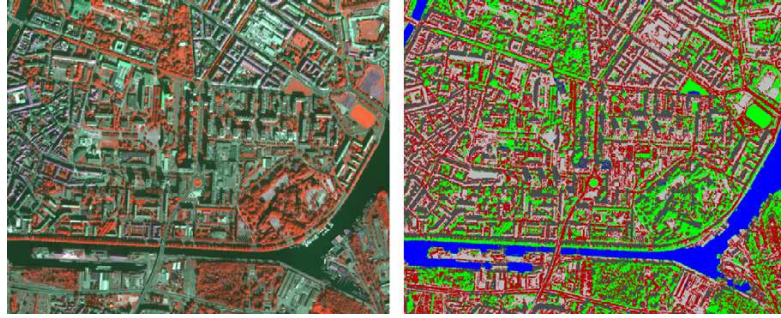


Fig. 7. The original multispectral image of Strasbourg, France in false colours (left), and the classification result using the bit mixing ordering based DMP features (right).

the situation is radically modified with the other two colour spaces, where all approaches increase their performances, a result showing the effect of the colour space choice on their behaviour. Specifically, lexicographical and bitmixing orderings take the lead, though with a small margin, while they provide the overall best results in $L^*a^*b^*$.

Since luminance/luminosity alone is considered sufficient for the recognition of most image variations, its prioritisation by means of total orderings has a positive impact on the end results. Contrarily to the previous experiment however, luminosity (L^*) provides better results with respect to luminance (Y), possibly due to the use of camera specific conversion matrices. Further improvements are probably possible by decreasing the perturbations caused by the last two channels (i.e. a^* , b^* and U , V), for instance by processing the last two marginally, since vector preservation is of no particular interest in this case.

4.3. Pixel classification

In this section we compare the different vector orderings in terms of their effect on the pertinence of features extracted from high resolution spatial data. As to the subject of the experiment, it is a multispectral image of an urban area of Strasbourg, France (Fig. 7, left). It comprises mainly dense residential and commercial areas along with a few open areas as well. The image has a spatial resolution of 1.3 m and a size of 1100×900 pixels with three channels (green, red and near infra-red) common to most commercial satellites (e.g. LANDSAT, IKONOS, Quickbird, etc.). Its high multispectral resolution, equivalent to the spatial resolution of panchromatic data, is due to the use of SPOT 5 sensors during a simulation flight. As a result, several small scale details and objects (e.g. cars) are present as well as long shadows, rendering the image relatively hard to classify.

Feature extraction is realised with the help of *differential morphological profiles* (DMP), an approach employed successfully in the field of remote sensing [97] for classifying panchromatic urban data. It consists of a granulometry (*opening profile*) and of an antigranulometry (*closing profile*) [96]. Thus, if we extend this principle to vectorial pixels, for each $\mathbf{f}(\mathbf{x}) = [x_1, x_2, x_3]^T$, two bi-dimensional $((n+1) \times 3)$ series are obtained, respectively, the *vectorial opening and closing*

profiles:

$$\Pi_\gamma(\mathbf{f})(\mathbf{x}) = \{\Pi_{\gamma_\lambda}(\mathbf{f})(\mathbf{x}) | \Pi_{\gamma_\lambda}(\mathbf{f})(\mathbf{x}) = \gamma_\lambda(\mathbf{f})(\mathbf{x}), \forall \lambda \in \{0, \dots, n\}\}, \quad (25)$$

$$\Pi_\phi(\mathbf{f})(\mathbf{x}) = \{\Pi_{\phi_\lambda}(\mathbf{f})(\mathbf{x}) | \Pi_{\phi_\lambda}(\mathbf{f})(\mathbf{x}) = \phi_\lambda(\mathbf{f})(\mathbf{x}), \forall \lambda \in \{0, \dots, n\}\}, \quad (26)$$

where $\gamma_0(\mathbf{f})(x, y) = \phi_0(\mathbf{f})(x, y) = \mathbf{f}(x, y)$ and λ denotes the size of the SE, chosen as a square of $2\lambda + 1$ pixels wide, whereas n being odd represents the size of the profile. Next the derivatives of each profile are calculated by means of componentwise subtractions for every vectorial pixel. Thus the resulting multi-scale feature set represents the slopes of each profile for every step of the monotonically varying SE size series:

$$\Delta_\gamma(\mathbf{f})(\mathbf{x}) = \{\Delta_{\gamma_\lambda}(\mathbf{f})(\mathbf{x}) | \Delta_{\gamma_\lambda}(\mathbf{f})(\mathbf{x}) = |\Pi_{\gamma_\lambda}(\mathbf{f})(\mathbf{x}) - \Pi_{\gamma_{\lambda-1}}(\mathbf{f})(\mathbf{x})|, \forall \lambda \in \{1, \dots, n\}\}, \quad (27)$$

$$\begin{aligned} \Delta_\phi(\mathbf{f})(\mathbf{x}) &= \{\Delta_{\phi_\lambda}(\mathbf{f})(\mathbf{x}) | \Delta_{\phi_\lambda}(\mathbf{f})(\mathbf{x}) \\ &= |\Pi_{\phi_\lambda}(\mathbf{f})(\mathbf{x}) - \Pi_{\phi_{\lambda-1}}(\mathbf{f})(\mathbf{x})|, \\ &\forall \lambda \in \{1, \dots, n\}\}. \end{aligned} \quad (28)$$

As a last step, the result is transformed into a uni-dimensional series by separating the channel components in order to form the vectorial DMP of size $6n$:

$$\text{DMP}(\mathbf{f})(\mathbf{x}) = \left\{ \begin{aligned} &\Delta_c(\mathbf{f})(\mathbf{x}) | \Delta_c(\mathbf{f})(\mathbf{x}) \\ &= \left\{ \begin{aligned} &\Delta_{\gamma_{\lambda=c-n(2i+1)}}^i(\mathbf{f})(\mathbf{x}), \\ &\forall c \in \{2ni+1, 2ni+n\}, \\ &\forall i \in \{0, 1, 2\} \\ &\Delta_{\phi_{\lambda=c-n(2i+1)}}^i(\mathbf{f})(\mathbf{x}), \\ &\forall c \in \{n(2i+1)+1, 2n(1+i)\}, \\ &\forall i \in \{0, 1, 2\} \end{aligned} \right\} \end{aligned} \right\}, \quad (29)$$

where $\Delta_{\gamma_\lambda}^i$ and $\Delta_{\phi_\lambda}^i$ represent, respectively, the i th component of Δ_{γ_λ} and Δ_{ϕ_λ} . The size of the profile was fixed as $n = 5$ thus providing $6n = 30$ features per vectorial pixel.

Table 6
Average and overall accuracies in percentage computed by means of a 10-fold cross validation

Class	Vector orderings						
	M	L	N	B	C	R	rL
Roads	83.3	82.4	85.5	82.5	85.5	85.4	85.6
Buildings	69.3	72.3	68.6	72.4	68.0	68.3	68.4
Water	91.3	96.8	89.5	97.2	89.7	90.1	89.8
Vegetation	94.0	94.4	93.1	94.4	93.0	92.8	93.4
Shadow	94.2	92.9	95.1	92.7	94.7	94.6	95.2
Average	86.4	87.8	86.4	87.8	86.2	86.2	86.5
Overall	87.0	88.2	86.9	88.3	86.8	86.9	87.0

For the classification process a naive Bayes classifier is employed, which despite its extended independence assumptions, is known for its robustness. Approximately 1% of the total number of pixels is included in the training set, composed of the following five classes: buildings, roads, water, vegetation and shadows. The feature sets are extracted independently and with the exact same arguments for each ordering while the evaluation step is realised as a 10-fold cross-validation. Table 6 shows the resulting classification accuracies for each class; as before the entries corresponding to the best performance of their row (pixel class) are in bold.

Given the values of Table 6, the overall accuracies appear to be in the magnitude of other reported uses of DMP for classification of panchromatic (monospectral) data ($\approx 78\%$ for IRS-1C and $\approx 95\%$ for IKONOS in Ref. [98]). Possible justifications of the difference include primarily the large size of the feature set. In addition, the components of the resulting vectorial features could have been combined in more sophisticated ways. Other influential factors include of course the complexity of the image as well as the efficiency of the classifier. Nevertheless, these choices were made arbitrarily because they do not bear importance in the context of this experiment as it focuses only on the relative performances of the approaches with respect to each other.

Moreover, the overall accuracies are remarkably close (L and B at $\approx 88\%$ and M, R, C and rL at $\approx 87\%$). In other words the total orderings have a slight advantage, with the exception of rL, which once more behaves almost identically to N as the outcome of most comparisons is determined by the first component. The order of the channels (green, red, near infra-red) apparently works in favour of total orderings, with the chosen classes. This remark becomes more evident if we study the classification rates of each class individually, where contrarily to the overall rates, differences of more than 7% take place.

For instance, it is widely known that clear water appears dark-bluish (i.e. higher green band reflectance), which when combined with the prioritisation attributed by L and B to the first channel (i.e. green) results in more pertinent features, equivalent to visibly higher classification rates for the same two orderings with the class “water”. The same logic can also be applied to the class “vegetation” which has high reflectance rates in

green and in near infra-red. On the other hand, shadows, roads and buildings may appear in various shades of blue or grey, depending on their composition, hence orderings taking into account all components equally (i.e. N, C, R, M and partially rL) are more effective in obtaining discriminating features, as asserted by Table 6.

In brief, despite the expectation that the restricted output of vectorial operators would damage the pertinence of the extracted features, the results show that this deficiency is compensated by their capacity to exploit the correlation of channels and thus they perform at least as well as the marginal approach. Additionally, pixel classes with distinctive spectral signatures, are more efficiently processed by orderings prioritising the corresponding channels, hence surpassing the accuracy rates of marginal ordering.

5. Conclusion

In this paper, we have presented an exhaustive and up to date review of the methods proposed in the literature with the end of extending the MM framework to multivalued images. We aimed especially to obtain a panorama of the existing approaches that, besides the MM community, is also accessible to those interested in multivalued image processing but not yet acquainted with the potential of multivariate MM. That is why the presentation of the theoretical foundation of the subject, that has been hardly modified since the milestone work of Goutsias et al. [6], was limited with the essential points, whereas abundant references to applications from various domains of multivalued morphological processing (i.e. colour, remote sensing, radar data, etc.) were provided.

The key point of constructing morphological operators for multivalued images consists in defining an ordering scheme for vectorial data. Several ordering methods as well as other specialised approaches were examined in this context, with each of them equipped with their own properties (e.g. channel prioritisation, vector preservation, etc.), thus rendering them suitable for certain tasks and unsuitable for others. In conclusion, there is still no single ordering scheme appropriate for all kinds of multivalued input, since their suitability is directly related to not only the overall semantic meaning of the individual channels but to the particular task under consideration as well.

The experiments that have been carried out in the second part aimed to assert the remarks concerning the theoretical properties of each ordering, and although not exhaustive, provided some interesting results. First the selected approaches have been tested in terms of noise reduction quality, where the overall superiority of the marginal processing strategy became clear, especially against uncorrelated noise, while in the case of correlated noise both marginal and vectorial orderings have produced similar smoothing rates, hence asserting the results of Comer and Delp [5]. Moreover, the same orderings have been also employed to the end of colour texture classification by means of the morphological covariance operator, where total orderings were shown to surpass their marginal counterpart in combination with a suitable colour space.

In the last experiment, multispectral remote sensing data has been used with the purpose of pixel classification. Although the so far reported cases in the literature of morphologically extracted features from this kind of images make exclusive use of the marginal approach, in this process we have showed that vector orderings chosen in relation to the properties of the desired classes hold the potential of improving the classification rates.

Acknowledgements

The authors wish to thank Prof. Ch. Ronse and the anonymous reviewers for their valuable comments, as well as the members of the FoDoMuST¹ project, for the remote sensing data and technical support during the classification experiments.

References

- [1] G. Matheron, Random Sets and Integral Geometry, Wiley, New York, 1975.
- [2] J. Serra, Image Analysis and Mathematical Morphology, vol. I, Academic Press, London, 1982.
- [3] J. Serra, Image Analysis and Mathematical Morphology, vol. II: Theoretical Advances, Academic Press, London, 1988.
- [4] P. Soille, Applications of morphological operators, in: B. Jähne, H. Haußecker, P. Geißler (Eds.), Handbook of Computer Vision and Applications, vol. 3, Academic Press, San Diego, 1999, pp. 283–296, (Chapter 12).
- [5] M. Comer, E. Delp, Morphological operations for color image processing, J. Electron. Imaging 8 (3) (1999) 279–289.
- [6] J. Goutsias, H.J.A.M. Heijmans, K. Sivakumar, Morphological operators for image sequences, Comput. Vision Image Understanding 62 (3) (1995) 326–346.
- [7] M.C. d'Ornellas, R.v.d. Boomgaard, J. Geusebroek, Morphological algorithms for color images based on a generic programming approach, in: Proceedings of the Brazilian Conference on Image Processing and Computer Graphics, Rio de Janeiro, Brazil, 1998, pp. 220–228.
- [8] F. Ortiz, F. Torres, J. Angulo, S. Puente, Comparative study of vectorial morphological operations in different color spaces, in: Proceedings of Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision, vol. 4572, 2001, pp. 259–268.
- [9] J. Serra, Anamorphoses and function lattices, in: E.R. Dougherty (Ed.), Mathematical Morphology in Image Processing, Marcel Dekker, New York, 1993, pp. 483–523, (Chapter 13).
- [10] H.J.A.M. Heijmans, Morphological Image Operators, Advances in Electronics and Electron Physics Series, Academic Press, Boston, 1994.
- [11] H.J.A.M. Heijmans, C. Ronse, The algebraic basis of mathematical morphology, part I: dilations and erosions, Comput. Vision Graphics Image Process. 50 (3) (1990) 245–295.
- [12] C. Ronse, Why mathematical morphology needs complete lattices, Signal Process. 21 (2) (1990) 129–154.
- [13] C. Ronse, H.J.A.M. Heijmans, The algebraic basis of mathematical morphology, part II: openings and closings, Comput. Vision Graphics Image Process. 54 (1) (1991) 74–97.
- [14] H.J.A.M. Heijmans, R. Keshet, Inf-semilattice approach to self-dual morphology, J. Math. Imaging Vision 17 (1) (2002) 55–80.
- [15] R. Keshet, Extensions of morphological operations to complete semilattices and its applications to image and video processing, in: H. Heijmans, J. Roerdink (Eds.), Mathematical Morphology and its Application to Image and Signal Processing, Kluwer Academic Publishers, Dordrecht, Germany, 1998, pp. 35–42.
- [16] R. Keshet, Mathematical morphology on complete semilattices and its applications to image processing, Fundam. Inform. 41 (1/2) (2000) 33–56.
- [17] J. Goutsias, H.J.A.M. Heijmans, Fundamenta morphologicae mathematicae, Fundam. Inform. 41 (1–2) (2000) 1–31.
- [18] R.C. Hardie, G.R. Arce, Ranking in R^p and its use in multivariate image estimation, IEEE Trans. Circuits Syst. Video Technol. 1 (2) (1991) 197–209.
- [19] R. Lukac, B. Smolka, K. Martin, K.N. Plataniotis, A.N. Venetsanopoulos, Vector filtering for color imaging, IEEE Signal Process. Mag. 22 (1) (2005) 74–86.
- [20] I. Pitas, P. Tsakalides, Multivariate ordering in color image filtering, IEEE Trans. Circuits Syst. Video Technol. 1 (3) (1991) 247–259.
- [21] K. Tang, J. Astola, Y. Neuvo, Multivariate ordering and filtering with application to color image processing, in: Proceedings of the Sixth European Signal Processing Conference, Brussels, Belgium, 1992, pp. 1481–1484.
- [22] P.E. Trahanias, A.N. Venetsanopoulos, Color edge detection based on multivariate ordering, in: SPIE Conference, Visual Communications and Image Processing, Boston, USA, 1992, pp. 1396–1407.
- [23] P.E. Trahanias, A.N. Venetsanopoulos, Multichannel image processing using vector-angle ranking, in: E.R. Dougherty, J. Astola, H.G. Longbotham (Eds.), SPIE Conference, Nonlinear Image Processing IV, vol. 1902, San Jose, USA, 1993, pp. 116–127.
- [24] V. Barnett, The ordering of multivariate data, J. Stat. Soc. A 139 (3) (1976) 318–355.
- [25] D.M. Titterton, Estimation of correlation coefficients by ellipsoid trimming, Appl. Stat. 27 (3) (1978) 227–234.
- [26] K.V. Mardia, J. Kent, J. Bibby, Multivariate Analysis, Academic Press, London, 1976.
- [27] J. Chanussot, Approches vectorielles ou marginales pour le traitement d'images multi-composantes, Ph.D. Thesis, Université de Savoie, France, 1998.
- [28] C. Gu, Multivalued morphology and its application in moving object segmentation and tracking, in: P. Maragos, R.W. Schafer, M.A. Butt (Eds.), Mathematical Morphology and its Application to Image and Signal Processing, Kluwer Academic Publishers, Boston, USA, 1996, pp. 345–352.
- [29] E. Aptoula, S. Lefèvre, C. Collet, Mathematical morphology applied to the segmentation and classification of galaxies in multispectral images, in: Proceedings of the 14th European Signal Processing Conference, Florence, Italy, 2006.
- [30] S.S. Wilson, Theory of matrix morphology, IEEE Trans. Pattern Anal. Mach. Intell. 14 (6) (1992) 636–652.
- [31] J. Astola, P. Haavisto, Y. Neuvo, Vector median filters, IEEE Proc. 78 (4) (1990) 678–689.
- [32] S. Batman, J. Goutsias, Unsupervised iterative detection of land mines in highly cluttered environments, IEEE Trans. Image Process. 12 (5) (2003) 509–523.
- [33] J.A. Benediktsson, J.A. Palmason, J.R. Sveinsson, Classification of hyperspectral data from urban areas based on extended morphological profiles, IEEE Trans. Geosci. Remote Sensing 43 (3) (2005) 480–491.
- [34] R. Jones, H. Talbot, Morphological filtering for colour images with no new colours, in: Proceedings of the Image and Vision Computing New Zealand (IVCNZ'96), Lower Hutt, New Zealand, 1996, pp. 149–154.
- [35] H. Talbot, C. Evans, R. Jones, Complete ordering and multivariate mathematical morphology, in: H. Heijmans, J. Roerdink (Eds.), Mathematical Morphology and Its Applications to Image and Signal Processing, Kluwer Academic Press, Amsterdam, 1998, pp. 27–34.
- [36] A. Toet, Detection of dim point targets in cluttered maritime backgrounds through multisensor image fusion, in: W.R. Watkins, D. Clement, W.R. Reynolds (Eds.), SPIE Conference, Targets and Backgrounds VIII: Characterization and Representation, vol. 4718, Orlando, USA, 2002, pp. 118–129.
- [37] K.R. Weber, S.T. Acton, On connected filters in color image processing, J. Electron. Imaging 13 (3) (2004) 619–629.
- [38] H.M. Al-Otum, Morphological operators for color image processing based on Mahalanobis distance measure, Soc. Photo-Opt. Instrum. Eng. 42 (9) (2003) 2595–2606.

¹ <http://lsiit.u-strasbg.fr/afd/sites/fodomust/fr-accueil.php>

- [39] M. Comer, E. Delp, An empirical study of morphological operators in color image enhancement, in: *Proceedings of the SPIE Conference on Image Processing Algorithms and Techniques III*, vol. 1657, San Jose, USA, 1992, pp. 314–325.
- [40] L.J. Sartor, A.R. Weeks, Morphological operations on color images, *J. Electron. Imaging* 10 (2) (2001) 548–559.
- [41] P.E. Trahanias, A.N. Venetsanopoulos, Vector directional filters—a new class of multichannel image processing filters, *IEEE Trans. Image Process.* 2 (4) (1993) 528–534.
- [42] A. Trémeau, C. Fernandez-Maloigne, P. Bonton, *Image Numérique Couleur*, Dunod, Paris, 2004.
- [43] A.N. Evans, Vector area morphology for motion field smoothing and interpretation, *IEE Proc. Vision Image Signal Process.* 150 (4) (2003) 219–226.
- [44] A. Plaza, P. Martinez, R. Perez, J. Plaza, A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles, *Pattern Recognition* 37 (6) (2004) 1097–1116.
- [45] M.T. Ureikat, H.M. Al-Otum, Angle-valued Mahalanobis-based image morphological operators, in: *Proceedings of the Fourth WSEAS International Conference on Signal, Speech and Image Processing*, Izmir, Turkey, 2004.
- [46] H.M. Al-Otum, M.T. Ureikat, Color image morphology using an adaptive saturation-based technique, *Opt. Eng.* 43 (6) (2004) 1280–1292.
- [47] A. Ledda, W. Phillips, Majority ordering for colour mathematical morphology, in: *Proceedings of the 13th European Signal Processing Conference*, Antalya, Turkey, 2005.
- [48] J. Li, Y. Li, Multivariate mathematical morphology based on principal component analysis: initial results in building extraction, in: *Proceedings of the 20th ISPRS*, vol. 35, Istanbul, Turkey, 2004, pp. 1168–1173.
- [49] M.I. Vardavoulia, I. Andreadis, P. Tsalides, Vector ordering and morphological operations for colour image processing: fundamentals and applications, *Pattern Anal. Appl.* 5 (3) (2002) 271–287.
- [50] I. Andreadis, M. Vardavoulia, G. Louverdis, N. Papamarkos, Colour image skeletonisation, in: *Proceedings of the 10th European Signal Processing Conference*, vol. 4, Tampere, Finland, 2000, pp. 2389–2392.
- [51] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, second ed., Addison-Wesley, New York, 1992.
- [52] A. Hanbury, J. Serra, Morphological operators on the unit circle, *IEEE Trans. Image Process.* 10 (12) (2001) 1842–1850.
- [53] A. Peters, Mathematical morphology for angle-valued images, in: *Proceedings of SPIE, Nonlinear Image Processing VIII*, vol. 3026, 1997, pp. 84–94.
- [54] G. Louverdis, M. Vardavoulia, I. Andreadis, P. Tsalides, A new approach to morphological colour image processing, *Pattern Recognition* 35 (8) (2002) 1733–1741.
- [55] G. Louverdis, I. Andreadis, P. Tsalides, Morphological granulometries for color images, in: *Second Hellenic Conference on Artificial Intelligence*, Thessaloniki, Greece, 2002, pp. 333–342.
- [56] M.I. Vardavoulia, I. Andreadis, P. Tsalides, A new vector median filter for colour image processing, *Pattern Recognition Lett.* 22 (6–7) (2001) 675–689.
- [57] F. Ortiz, Colour mathematical morphology for neural image analysis, *Real-Time Imaging* 8 (5) (2002) 455–465.
- [58] F. Ortiz, F. Torres, Vectorial morphological reconstruction for brightness elimination in colour images, *Real-Time Imaging* 10 (6) (2004) 379–387.
- [59] F. Ortiz, F. Torres, P. Gil, Gaussian noise elimination in colour images by vector-connected filters, in: *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 4, Cambridge, UK, 2004, pp. 807–810.
- [60] A. Hanbury, J. Serra, Mathematical morphology in the CIELAB space, *Image Anal. Stereology* 21 (3) (2002) 201–206.
- [61] A. Hanbury, J. Serra, Mathematical morphology in the HLS colour space, in: T. Cootes, C. Taylor (Eds.), *12th British Machine Vision Conference*, Manchester, UK, 2001, pp. 451–460.
- [62] J. Angulo, J. Serra, Mathematical morphology in color spaces applied to the analysis of cartographic images, in: S. Levachkine, J. Serra, M. Egenhofer (Eds.), *Proceedings of GEOPRO '03, International Workshop of Semantic Processing of Spatial Data*, Mexico City, Mexico, 2003, pp. 56–66.
- [63] J. Rivest, Morphological operators on complex signals, *Signal Process.* 84 (1) (2004) 133–139.
- [64] J. Rivest, Granulometries and pattern spectra for radar signals, *Signal Process.* 86 (5) (2006) 1094–1103.
- [65] J. Angulo, Morphological color processing based on distances. Application to color denoising and enhancement by centre and contrast operators, in: *Proceedings of the IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP'2005)*, Benidorm, Spain, 2005, pp. 314–319.
- [66] J. Angulo, J. Serra, Morphological coding of color images by vector connected filters, in: *IEEE Proceedings of the Seventh International Symposium on Signal Processing and Its Applications (ISSPA'2003)*, vol. 1, Paris, France, 2003, pp. 69–72.
- [67] J. Angulo, *Morphologie mathématique et indexation d'images couleur. application à la microscopie en biomedicine*, Ph.D. Thesis, Ecoles des Mines, Paris, France, 2003.
- [68] J. Angulo, Unified morphological color processing framework in a lum/sat/hue representation, in: C. Ronse, L. Najman, E. Decencière (Eds.), *Proceedings of the Seventh ISMM, Computational Imaging and Vision*, vol. 30, Springer, Dordrecht, Germany, 2005, pp. 387–396.
- [69] J. Chanussot, P. Lambert, Total ordering based on space filling curves for multivalued morphology, in: *Proceedings of the International Symposium on Mathematical Morphology (ISMM'98)*, Amsterdam, Netherlands, 1998, pp. 51–58.
- [70] J. Chanussot, P. Lambert, Extending mathematical morphology to color image processing, in: *International Conference on Color in Graphics and Image Processing*, Saint-Etienne, France, 2000, pp. 158–163.
- [71] J. Chanussot, P. Lambert, Watershed approaches for color image segmentation, in: *Proceedings of the IEEE Workshop on Nonlinear Signal and Image Processing*, Antalya, Turkey, 1999, pp. 129–133.
- [72] E. Stringa, A. Teschioni, C.S. Regazzoni, A classical morphological approach to color image filtering based on space filling surveys, in: *Proceedings of the IEEE-Eurasip Workshop on Nonlinear Signal and Image Processing*, Antalya, Turkey, 1999, pp. 351–354.
- [73] P. Soille, M. Pesaresi, Advances in mathematical morphology applied to geoscience and remote sensing, *IEEE Trans. Geosci. Remote Sensing* 40 (9) (2002) 2042–2055.
- [74] I. Bloch, H. Maitre, Fuzzy mathematical morphologies: a comparative study, *Pattern Recognition* 28 (9) (1995) 1341–1387.
- [75] T.Q. Deng, H.J.A.M. Heijmans, Grey-scale morphology based on fuzzy logic, *J. Math. Imaging Vision* 16 (2) (2002) 155–171.
- [76] M. Köppen, C. Nowack, G. Rösel, Pareto-morphology for color image processing, in: *Proceedings of the 11th Scandinavian Conference on Image Analysis*, vol. 1, Kangerlussuaq, Greenland, 1999, pp. 195–202.
- [77] G. Louverdis, I. Andreadis, P. Tsalides, A new fuzzy model for morphological colour image processing, *IEE Proc.—Vision, Image Signal Process.* 149 (3) (2002) 129–139.
- [78] B. Burgeth, A. Bruhn, N. Papenberg, M. Welk, J. Weickert, Mathematical morphology for matrix fields induced by the loewner ordering in higher dimensions, *Signal Process.* 87 (2) (2007) 277–290.
- [79] B. Burgeth, N. Papenberg, A. Bruhn, M. Welk, C. Feddern, J. Weickert, Morphology for higher-dimensional tensor data via Loewner ordering, in: C. Ronse, L. Najman, E. Decencière (Eds.), *Proceedings of the Seventh ISMM, Computational Imaging and Vision*, vol. 30, Springer, Dordrecht, Germany, 2005, pp. 407–416.
- [80] B. Burgeth, M. Welk, C. Feddern, J. Weickert, Morphological operations on matrix-valued images, in: *European Conference on Computer Vision*, vol. 4, Prague, Czech Republic, 2004, pp. 155–167.
- [81] M. Wheeler, M.A. Zmuda, Processing color and complex data using mathematical morphology, in: *IEEE Proceedings of the National Aerospace and Electronics Conference*, vol. 1, Dayton, USA, 2000, pp. 618–624.
- [82] O. Lezoray, C. Meurie, A. Elmoataz, A graph approach to color mathematical morphology, in: *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology*, Athens, Greece, 2005.

- [83] C. Ronse, V. Agnus, Morphology on label images: flat-type operators and connections, *J. Math. Imaging Vision* 22 (2/3) (2005) 283–307.
- [84] E. Zaharescu, M. Zamfir, C. Vertan, Color morphology-like operators based on color geometric shape characteristics, in: *Proceedings of the International Symposium on Signals, Circuits and Systems*, vol. 1, Bucharest, Romania, 2003, pp. 145–148.
- [85] A. Mojsilovic, E. Soljanin, Quantization of color spaces and processing of color images by Fibonacci lattices, in: *Proceedings of Human Vision and Electronic Imaging V*, vol. SPIE 3959, San Jose, USA, 2000, pp. 345–355.
- [86] S. Gibson, J.A. Bangham, R. Harvey, Evaluating a colour morphological scale-space, in: *Proceedings of the British Machine Vision Conference*, Norwich, UK, 2003.
- [87] C. Busch, M. Eberle, Morphological operations for color-coded images, in: *Proceedings of EUROGRAPHICS'95*, Computer Graphics Forum, vol. 14, Maastricht, Netherlands, 1995, pp. 193–204.
- [88] A. Hanbury, The morphological top-hat operator generalised to multi-channel images, in: *Proceedings of the 17th International Conference on Pattern Recognition*, Cambridge, UK, 2004, pp. 672–675.
- [89] F. Ortiz, F. Torres, P. Gil, J. Pomares, S. Puente, F. Candelas, Ordering by distance for hsi mathematical morphology, in: *Proceedings of the Ninth Spanish Symposium on Pattern Recognition and Image Analysis*, vol. 2, Castellón, Spain, 2001, pp. 379–384.
- [90] C. Vertan, M. Malciu, V. Buzuloiu, V. Popescu, Median filtering techniques for vector valued signals, in: *Proceedings of the Third IEEE International Conference on Image Processing*, vol. 1, Lausanne, Switzerland, 1996, pp. 977–980.
- [91] C. Vertan, V. Popescu, V. Buzuloiu, Morphological like operators for color images, in: *Proceedings of the Eighth European Signal Processing Conference*, Trieste, Italy, 1996, pp. 165–168.
- [92] T. Mäenpää, M. Pietikäinen, Classification with color and texture: jointly or separately?, *Pattern Recognition* 37 (8) (2004) 1629–1640.
- [93] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, S. Huovinen, Outex: new framework for empirical evaluation of texture analysis algorithms, in: *Proceedings of the 16th ICPR*, vol. 1, Quebec City, Canada, 2002, pp. 701–706.
- [94] E. Aptoula, S. Lefèvre, Spatial morphological covariance applied to texture classification, in: *Proceedings of the International Workshop on Multimedia Content Representation, Classification and Security*, vol. 4105, Istanbul, Turkey, 2006, pp. 522–529.
- [95] A. Hanbury, U. Kandaswamy, D.A. Adjeroh, Illumination-invariant morphological texture classification, in: C. Ronse, L. Najman, E. Decencière (Eds.), *Proceedings of the Seventh ISMM, Computational Imaging and Vision*, vol. 30, Springer, Dordrecht, Netherlands, 2005, pp. 377–386.
- [96] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer, Berlin, 2003.
- [97] M. Pesaresi, J. Benediktsson, A new approach for the morphological segmentation of high resolution satellite imagery, *IEEE Trans. Geosci. Remote Sensing* 39 (2) (2001) 309–320.
- [98] J.A. Benediktsson, M. Pesaresi, K. Arnason, Classification and feature extraction for remote sensing images from urban areas based on morphological transformations, *IEEE Trans. Geosci. Remote Sensing* 41 (9) (2003) 1940–1949.

About the Author—ERCHAN APTOULA received the B.Sc. degree in Computer Engineering from Galatasaray University, Istanbul, Turkey in 2004 and the M.Sc. in Computer Science from Louis Pasteur University, Strasbourg, France in 2005, where he is currently pursuing a Ph.D. degree in Computer Science with a thesis on colour mathematical morphology applied to content-based image retrieval. Apart from morphological analysis, his interests include multispectral image processing, pattern recognition and machine learning.

About the Author—SÉBASTIEN LEFÈVRE received in 1999 the M.S. and Eng. degrees from the University of Compiègne in Computer Engineering, and in 2002 the Ph.D. degree from the University of Tours in Computer Science. He is currently an Assistant Professor in the Department of Computer Science and the LSIIT, University Louis Pasteur, Strasbourg. From 1999 to 2002 he was with AtosOrigin as a Research and Development Engineer. In 2003, he was with the Polytechnical School of the University of Tours as an Assistant Professor. His research interests are in multispectral image and colour video processing, multimedia analysis and indexing, and mathematical morphology.

Available online at www.sciencedirect.com

Pattern Recognition Letters 29 (2008) 109–118

Pattern Recognition
Letterswww.elsevier.com/locate/patrec

On lexicographical ordering in multivariate mathematical morphology

E. Aptoula ^{*}, S. Lefèvre

*Image Sciences, Computer Sciences and Remote Sensing Laboratory, LSIT-UMR-7005-CNRS Louis Pasteur University, Pôle API,
Blvd Sébastien Brant, P.O. Box 10413, 67412 Illkirch Cedex, France*

Received 28 January 2007; received in revised form 2 September 2007
Available online 1 October 2007

Communicated by G. Borgefors

Abstract

Since mathematical morphology is based on complete lattice theory, a vector ordering method becomes indispensable for its extension to multivariate images. Among the several approaches developed with this purpose, lexicographical orderings are by far the most frequent, as they possess certain desirable theoretical properties. However, their main drawback consists of the excessive priority attributed to the first vector dimension. In this paper, the existing solutions to solving this problem are recalled and two new approaches are presented. First, a generalisation of α -modulus lexicographical ordering is introduced, making it possible to accommodate any quantisation function. Additionally, an input specific method is suggested, based on the use of a marker image. Comparative application results on colour noise reduction and texture classification are also provided.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Multivariate mathematical morphology; Lexicographical ordering; Colour morphology; Dimension prioritisation

1. Introduction

With the acceptance of complete lattice theory as the appropriate theoretical framework for mathematical morphology (MM) (Goutsias et al., 1995; Ronse, 1990; Serra, 1993), it became possible to define morphological operators for any type of image data, as long as a complete lattice structure can be introduced on the pixel intensity range. However, as far as colour, and more generally multivariate images are concerned, the ambiguity of ordering vectorial pixels constitutes the main obstacle in this regard. Several vector ordering mechanisms have been reported in the literature with this purpose, of which none has yet been widely accepted. A comprehensive survey on the different approaches to multivariate MM can be found in (Aptoula and Lefèvre).

As confirmed by the last reference, among the rich variety of ordering methodologies, lexicographical ordering has been by far the most frequently employed solution with the end of extending morphological operators to multivariate image data. Its relative popularity is due mainly to its theoretical properties, since it makes it possible to totally order the underlying pixel data while also providing unique extrema. Nevertheless, lexicographical ordering suffers from a major drawback, which consists of the excessive prioritisation of the first image channel during the lexicographical cascade, hence leading to an inefficient exploitation of inter-channel relations.

In this paper, we recall the existing solutions, and present two approaches that aim to counter this inconvenience. First, a generalisation of α -modulus lexicographical ordering is introduced, that makes it possible to quantise the first vector dimension based on an arbitrary model. Moreover, a colour morphology related example of this method is provided, using luminance and saturation dependencies. Additionally, an input specific approach is suggested, that orders the first vector components with respect to a marker image,

^{*} Corresponding author. Tel.: +33 (0)3 90 24 44 99; fax: +33 (0)3 90 24 44 55.

E-mail address: aptoula@lsit.u-strasbg.fr (E. Aptoula).

thus providing a means of controlling the channel prioritisation of lexicographical ordering based on topological information. Both approaches are tested in the context of colour morphology, against state-of-the-art methods, with noise reduction and texture classification applications, where they provide superior results.

The rest of the paper is organised as follows. Section 2 introduces briefly the theoretical background of extending morphological operators to multivariate images as well as the problems related to lexicographical ordering and its variations. In Section 3, a generalisation of α -modulus lexicographical ordering is presented, along with an application in colour morphology. Then, in Section 4, the marker based ordering is introduced and its properties are elaborated. Next, in Section 5, the results of a series of comparative tests are provided. Finally Section 6 is devoted to concluding remarks.

2. Lexicographical ordering in morphological processing

In this section we review briefly the main issues concerning the extension of MM to multivariate images, and discuss the properties of lexicographical ordering that contributed to its widespread use, along with its inconveniences and subsequent variations. For an in-depth study of multivariate MM the reader can refer to the fundamental references (Goutsias et al., 1995; Serra, 1993).

2.1. Theoretical background

Complete lattice theory (Birkhoff, 1967), is nowadays considered as the right mathematical framework for morphology (Ronse, 1990). The (minimal) conditions that need to be satisfied by morphological operators have been formulated by means of this theory. Specifically, introducing a complete lattice structure on the “grey levels” or intensity range of an image is sufficient for the definition of valid morphological operators.

More precisely, given an image $f: \mathcal{E} \rightarrow \mathcal{T}$ with $\mathcal{E}$ an arbitrary non empty set, a complete lattice structure is imposed on the intensity range $\mathcal{T}$. In other words, $\mathcal{T}$ must be a non empty set equipped with a partial ordering such that every non empty subset $\mathcal{P} \subseteq \mathcal{T}$ has a greatest lower bound $\bigwedge \mathcal{P}$ (infimum) and a least upper bound $\bigvee \mathcal{P}$ (supremum). Consequently, the set of images $\mathcal{F}(\mathcal{E}, \mathcal{T})$ is also a complete lattice.

For instance, in the case of grey-scale images, where usually $\mathcal{T} = \mathbb{Z}$, or $\mathcal{T} = \mathbb{R}$ or a finite interval $[a, b]$, the scalar order is employed for ordering the pixel values in order to obtain the lower and greater bounds. However, this becomes a much more challenging task with multivariate images, where usually $\mathcal{T} = \mathbb{Z}^n$ or $\mathcal{T} = \mathbb{R}^n$ with $n > 1$, since there is no natural ordering relation for multivariate data.

Indeed, given a vector ranking mechanism $<$, the vectorial versions of the two fundamental morphological operators, *erosion* (ε_b) and *dilation* (δ_b) of a multivalued image f , by a flat structuring element (SE) b , can be immediately

derived by means of the vectorial extrema operators $\sup_v$ and $\inf_v$ based on the given ordering

$$\varepsilon_b(f)(x) = \inf_{s \in b} \{f(x + s)\} \quad (1)$$

$$\delta_b(f)(x) = \sup_{s \in b} \{f(x - s)\} \quad (2)$$

Therefore, the main obstacle preventing the extension of morphological operators to multivalued images, consists in defining a vector ordering relation that will induce a complete lattice structure on the pixel intensity range.

2.2. Lexicographical ordering

The problem of ordering multivariate data has been extensively studied in the past, especially in the field of statistics (Barnett, 1976). According to the ordering categorisation given in the last reference, *lexicographical ordering* ($<_L$) is a type of conditional ordering (c-ordering), where the vector components are ordered conditionally based on some of their marginal components, and is defined as

$$\forall v, v' \in \mathbb{R}^n, \quad v <_L v' \iff \exists i \in \{1, \dots, n\}, \\ (\forall j < i, v_j = v'_j) \wedge (v_i < v'_i) \quad (3)$$

For vectors in a colour space VWX , where we compare first dimension V , then W and finally X , we adopt the notation $V \rightarrow W \rightarrow X$.

Lexicographical ordering is particularly suitable for ordering vectors in the context of multivariate MM, in combination with image data where a natural or artificial priority order exists among the different bands, as it possesses desirable theoretical as well as practical properties. First, as all c-orderings, it is a total ordering, hence preserving the input vectors, in the contrary of marginal processing. In other words, it does not introduce any new vectors within the processing results, and thus prevents the apparition of *false colours* in colour morphology (Talbot et al., 1998), as well as the negative effect of mixed pixel signatures during the morphological classification of multispectral images (Aptoula and Lefèvre). It should also be noted, however, that a total ordering is not always desirable within the morphological context (Soille, 2005).

Moreover, as it satisfies the anti-symmetry constraint, a property lacking from most reduced orderings, it makes it possible to compute unique vectorial extrema, hence effectively avoiding ambiguities during vector ordering. Furthermore, by modifying the order of comparison during the lexicographical cascade, one can easily establish different priorities among the image channels. For instance, if used in the RGB colour space, the red channel would be inevitably prioritised, whereas it would be sufficient to permute the bands as GRB to shift this priority to the green¹ channel (Fig. 1).

¹ For interpretation of colour in Fig. 1 the reader is referred to the web version of this article.

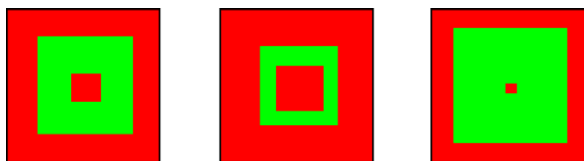


Fig. 1. Original image (left), results of applying a vectorial dilation based on a lexicographical ordering ($R \rightarrow G \rightarrow B$, middle) ($G \rightarrow R \rightarrow B$, right) with a 21×21 square SE.

In the case of multispectral or even hyperspectral images, where usually no a priori priority order exists among the tens or even hundreds of bands, the use of lexicographical ordering is practically justified only after the application of a proper transform (e.g. maximum noise fraction transform, principal components analysis, discrete cosine transform, etc.) which will redistribute the total variational information contained within the bands with a monotonic order, and thus artificially create the prioritised band environment (Aptoula and Lefèvre; Plaza et al., 2004). Hence making it possible to lexicographically exploit the information concentrated on the first few bands.

These invaluable properties have led to the widespread use of lexicographical ordering with the purpose of establishing multivariate morphological operators, especially in the field of colour morphology (Angulo, 2005; Aptoula and Lefèvre; Hanbury and Serra, 2001; Louverdis et al., 2002; Rivest, 2004).

2.3. Variations of lexicographical ordering

However, besides being equipped with several desirable properties, lexicographical ordering also suffers from a serious drawback. More precisely, the outcome of the vast majority of lexicographical comparisons, is decided based only on the first few vector components that are compared, while the remaining dimensions' contribution can be considered negligible. This property is illustrated in Table 1, where the percentages of comparisons determined by the three channels of the three RGB colour images of Fig. 2, during a vectorial dilation based on a lexicographical ordering are shown. The channel occupying the first position of the lexicographical cascade (i.e. red²) is obviously responsible for the vast majority of comparisons.

Of course this might be a desired behaviour in cases where the first image channel contains the majority of the total variational information, for instance after applying a PCA transform. Nevertheless, most often it leads to an insufficient exploitation of the image channels and inter-channel relations. This effect is most aggravated in the case of hyperspectral images, where despite the availability of hundreds of channels, only at most the first few participate in the overall process. That is why, variations of expression (3) were reported, with the end of better tuning the priority

² For interpretation of colour in Fig. 2 the reader is referred to the web version of this article.

Table 1

The percentages of comparisons that have led to equalities, and those that have been determined at each channel, during the dilation of the images of Fig. 2, with a lexicographical ordering ($R \rightarrow G \rightarrow B$) using a square shaped SE of size 5×5 pixels

Images	Equalities (%)	Red (%)	Green (%)	Blue (%)
Lenna	0.12	93.02	6.3	0.56
Macaws	8.89	83.3	5.71	2.1
Mandrill	0.54	95.28	3.51	0.67



Fig. 2. The test images, from left to right, Lenna, Macaws and Mandrill.

as well as degree of influence of each vector component on the comparison outcome, by means of an user defined argument α . A list of such extensions follows.

The first attempt aiming to decrease the priority attributed to the first vector component during lexicographical comparison was made by Ortiz et al. (2001), that proposed the α -lexicographical ordering

$$\forall \mathbf{v}, \quad \mathbf{v}' \in \mathbb{R}^n, \mathbf{v} < \mathbf{v}' \\ \iff \begin{cases} v_1 + \alpha < v'_1, & \text{or} \\ v_1 + \alpha \geq v'_1 & \text{and } [v_2, \dots, v_n]^T <_L [v'_2, \dots, v'_n]^T \end{cases} \quad (4)$$

where $\alpha \in \mathbb{R}^+$. The α argument is thus used to the end of increasing the occurrence of equivalences within the first vector dimension, since a scalar value v_1 becomes “equal” to all values contained in the interval $[v_1 - \alpha, v_1 + \alpha]$, hence allowing comparisons to reach more frequently the second dimension. Nevertheless, expression (4) is not transitive, and consequently does not represent an ordering from an algebraic point of view.

A further attempt was made by Aptoula and Lefèvre (2007), that introduced α -trimmed lexicographical extrema. Specifically, it can be considered as an iterative approach based on the principle of α -trimming, where given a set V of k vectors, in the case of the maximum, starting from the first dimension, the contents of V are sorted according to this dimension, and then the $\lceil \alpha \times k \rceil$ ($\alpha \in]0, 1[$) greatest vectors are kept with respect to the sorting order and considered as the new set V . By repeating this process for each dimension, at each step the initial set of vectors gets smaller, leading eventually to the desired extremum. Thus, each dimension can contribute to the computation of the extrema by means of the trimming step. For the formal description of this procedure the reader can refer to the original paper (Aptoula and Lefèvre, 2007). On the other hand, as this methodology is not based on a binary relation among vectors, but rather relies on a “collective”

extremum computation, it does not constitute an ordering, hence leading to pseudo-morphological operators.

A theoretically sounder approach was proposed by Angulo and Serra (2003) and was named α -modulus lexicographical ordering

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \quad \mathbf{v} < \mathbf{v}' \iff \left[\lceil v_1/\alpha \rceil, v_2, \dots, v_n \right]^T <_L \left[\lceil v'_1/\alpha \rceil, v'_2, \dots, v'_n \right]^T \quad (5)$$

which aims to create equivalence groups within the first dimension. It relies on a quantisation through division by a constant α followed by a rounding off, which reduces the dynamic margin of the first dimension, thus allowing a greater number of comparisons to reach the second. On the contrary of the aforementioned two approaches, it is an ordering and leads to valid morphological operators.

The effect of the operation realised on the first dimension, becomes more clear by studying the *space filling curves* (SFC), that travel through points of multi-dimensional space. Since the search for a total vector ordering can be formulated as a search for an injective function mapping all the points of a multi-dimensional space onto an uni-dimensional space, SFC satisfy this requirement and make it possible to model different solutions, where vectors are ordered according to the position of their coordinates on the SFC. Since lexicographical ordering corresponds to a bijection (Chanussot and Lambert, 2000), its SFC will pass once from all points of a bi-dimensional discrete space $D1 \times D2$ as illustrated in Fig. 3 (left), where the high priority attributed to the first dimension ($D1$) is represented by the high frequency of horizontal curves. Fig. 3 (right) shows the quantised form $D1'$ of dimension $D1$ with $\alpha = 4$, which leads to the creation of equivalence groups. For instance the points $\{4, 5, 6, 7\}$ of $D1$ now belong to the same group 1 of $D1'$, and thus given two coordinates $(4, 1)$ and $(7, 0)$ in $D1 \times D2$, the first components are considered equal in $D1' \times D2$ and the outcome of the comparison is determined by the second dimension, in other words $(4, 1) > (7, 0)$.

Nevertheless, special attention is required with the use of this ordering, since the resulting equivalence groups obviously eliminate the anti-symmetry property of lexicographical ordering. Consequently, equivalence between vectors does not occur in the usual sense within the quantised space

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \quad \mathbf{v} = \mathbf{v}' \iff (\mathbf{v} \leq \mathbf{v}') \wedge (\mathbf{v} \geq \mathbf{v}') \quad (6)$$

but rather between vectors using their original unquantised values

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n, \quad \mathbf{v} = \mathbf{v}' \iff [v_1, v_2, \dots, v_n]^T = [v'_1, v'_2, \dots, v'_n]^T \quad (7)$$

In conclusion, among the solutions that have been so far reported with the purpose of fine-tuning the dimension prioritisation of lexicographical ordering, only expression (5) satisfies the theoretical requirements of an ordering and hence leads to valid morphological operators. The practical use of α -modulus lexicographical ordering however is limited, as it relies on an implicit assumption on the dimension to be quantised. The next sections introduce a generalisation of this approach that makes it possible to employ a wider range of quantisation functions, as well as a method that relies on the, eventually morphological, smoothing of the first image channel, thus leading to equivalence groups based on spatial relations.

3. Quantisation based α -lexicographical ordering

As far as its theoretical properties are concerned, α -modulus lexicographical ordering, appears to be the most pertinent among the variations presented in Section 2.3, as it provides an effective means of shifting priority away from the first vector dimension, while preserving the desirable characteristics of lexicographical ordering that have made it popular within the morphological context (Angulo, 2005; Aptoula and Lefèvre; Hanbury and Serra, 2001; Louverdis et al., 2002; Rivest, 2004).

However, its practical use relies on an important implicit assumption. More precisely, as shown in Fig. 4a, the

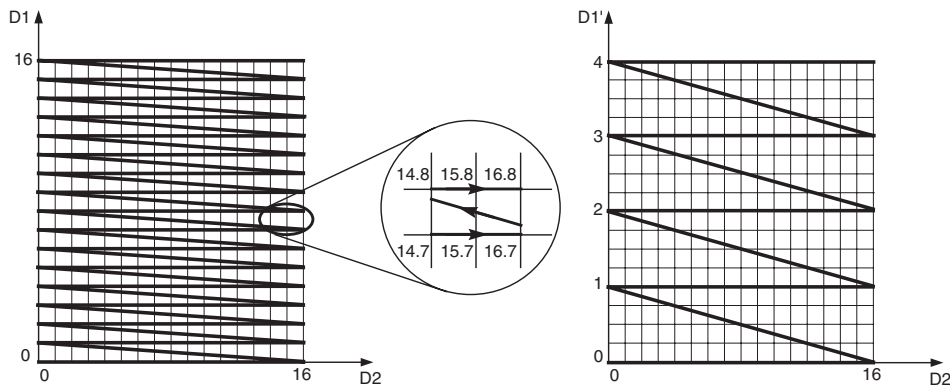


Fig. 3. Space filling curves in a bi-dimensional space for a lexicographical ordering (left), where the arrows denote the direction of increasing vectorial coordinates, and for the α -modulus lexicographical ordering with $\alpha = 4$ (right).

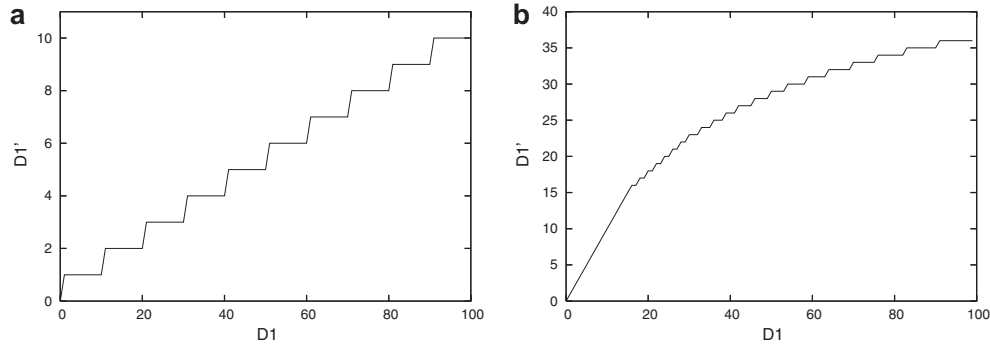


Fig. 4. Example of quantisation applied on a dimension D1 using $\alpha = 10$ and expression (5) (a), and using the proposed approach with the same α and an exponential distribution (b).

dimension under consideration is processed uniformly, based on the assumption that all of its subsets are equally “unimportant” with respect to the second dimension in the lexicographical cascade. In practice, often complicated relations are present among the image channels, and the need to shift priority to the next dimensions arises only partially. Consider for instance the case of colour morphology in the HSL colour space, where the importance of saturation is minimised for extreme levels of luminance.

In the light of these remarks, we propose the generalisation described in Algorithm 1, as a replacement for the division followed by rounding off, of expression (5). Based on a simple principle, this algorithm realises a quantisation of a given discrete “pixel range” interval, in practice often $[0, 255]$, by associating to each value, a group of equivalence, the size of which is computed using a user defined function f , and is limited by the value of $\alpha \in \mathbb{N}^*$. Hence, it is possible to reformulate expression (5) as

Algorithm 1. The algorithm for computing a quantised discrete dimension based on an arbitrary priority distribution

Input: $I = [a, \dots, b] \subseteq \mathbb{Z}$, an array containing the discrete dimension to have its dynamic margin reduced
 $\alpha \in \mathbb{N}^*$, a parameter setting the maximum allowed size of equivalence groups within I
 $f : \mathbb{N} \rightarrow [0, 1]$, a function modelling the desired priority distribution within I
Output: $J \subseteq \mathbb{N}$, the array containing the new quantised dimension
 $tmp \leftarrow 0$
for $i \leftarrow a$ **to** b **do**
 $k \leftarrow \lceil \alpha \times f(I[i - a]) \rceil$
 for $j \leftarrow i$ **to** $i + k$ **do**
 $J[j - a] \leftarrow tmp$
 end for
 $tmp \leftarrow tmp + 1$
 $i \leftarrow i + k$
end for

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n,$$

$$\mathbf{v} < \mathbf{v}' \iff [w_1, v_2, \dots, v_n]^T <_L [w'_1, v'_2, \dots, v'_n]^T \quad (8)$$

where w_1 and w'_1 represent, respectively, the equivalence group of v_1 and v'_1 , obtained through Algorithm 1. Of course one is by no means limited with applying this procedure only to the first dimension. In fact, complicated inter-channel relations often require the repeated use of such an approach in more than one dimensions in order to be effectively modelled.

An example of this algorithm is given in Fig. 4b, where it is applied on the integer interval $[0, 100]$ using an exponential priority distribution. In other words, values close to zero are of high importance and need to be processed with a fine precision whereas the second dimension may be used with values approaching 100. The increase in the size of equivalence groups can be easily observed as the values approach the upper interval bound. Consequently, in order to obtain the uniform distribution of Fig. 4a, corresponding to the quantisation of α -modulus lexicographical ordering, with this approach, it is sufficient to employ the constant function $\forall n \in \mathbb{N}, f(n) = 1$ within Algorithm 1. By means of the function f , one can thus model arbitrary priority distributions within image channels. Furthermore, image specific ordering approaches may be developed using the histogram of the dimension under consideration, as shown in Fig. 5, hence leading to adaptive vector orderings.

As a more concrete example, let us return to the aforementioned case of colour morphology in the HSL colour space. Polar colour spaces exhibit an inherent prioritisation of channels, as they model the higher levels of human colour vision, and thus luminance is most often used in the first position of the lexicographical cascade when developing morphological operators (Angulo, 2005; Aptoula and Lefèvre; Louverdis et al., 2002; Ortiz et al., 2001). However, non trivial relations are present among the three channels (luminance, saturation and hue), that are totally ignored using expression (5). Specifically, the importance of saturation is maximised for “medium” levels of luminance, whereas

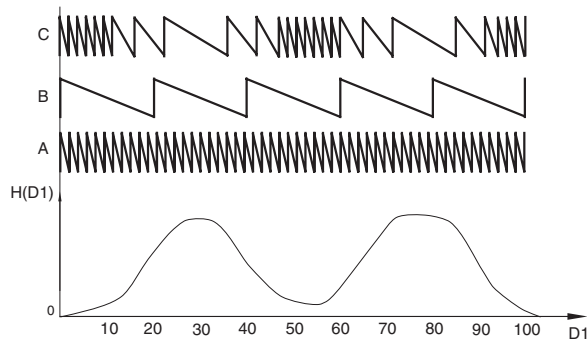


Fig. 5. The probability distribution of dimension D1 (bottom) for an arbitrary image, general appearance of the space filling curve frequency corresponding to its lexicographical ordering (A), α -modulus lexicographical ordering with $\alpha = 20$ (B), and quantisation based α -lexicographical ordering using the dimension histogram as priority distribution function (C).

hue is of practical importance only for relatively high levels of saturation.

Since luminance alone is sufficient for the recognition of most objects, it occupies almost always the first position of lexicographical cascades during the comparison of HSL colour vectors. For the sake of simplicity let us limit all subsequent operations on this colour space to the last two channels ($L \rightarrow S$), since the hue, being an angular value (Hanbury and Serra, 2001) requires a reference value which would further parameterise the procedure, and besides most comparisons seldom reach the third dimension. Moreover, the same principles if desired can also be applied between saturation and hue in a similar way. The standard way for decreasing the excessive priority attributed to luminance, using α -modulus lexicographical ordering, would lead to the quantised result shown in Fig. 4a.

For illustrational purposes we choose to model the importance of saturation $f(L)$ with respect to luminance L using a double sigmoidal distribution (Fig. 6a). Consequently, by means of $f(L)$ combined with the proposed quantisation scheme, we obtain the corresponding quantised

form of luminance as depicted in Fig. 6b. Observe the presence of the largest equivalence groups for medium luminance values. Hence, at this region, the number of comparisons determined by the luminance component will be minimised, accordingly to the given model, and instead will be determined by the following vector dimension, in other words saturation.

As illustrated by this relatively simple example, the proposed priority reduction scheme constitutes a generalisation of the quantisation step in expression (5), that provides a high degree of flexibility, making it possible for the user to accommodate within the resulting ordering relation practically any particular model of priority distribution. An application example is given in Fig. 7, which shows the results obtained after dilation and erosion with orderings 5 and 8 using the same value of α . The differences between the two approaches are fairly visible, in particular at the bleak and head region, where more smoothed variations are obtained.

4. Marker based lexicographical ordering

Creating artificial equivalence groups within the dimension to have its lexicographical priority reduced, is an effective means of shifting priority to the remaining vector dimensions. However, this is an operation realised independently from the images to be processed (except if the priority distribution is image specific, e.g. based on its histogram), and no a priori information is available on where the, eventually abrupt, equivalence group limits correspond within the input image. This undesirable situation is illustrated in Fig. 8 where the equivalence groups of a uni-dimensional discrete signal, obtained with expression (5), lead to artificial edges/value variations during comparison.

An alternative way of achieving a priority shift that avoids this inconvenience, consists in assuming an image specific approach and forming these equivalence groups based on the content of the image to be processed. Specifically, one can preprocess the input image's channel that is to have its priority reduced, for instance the luminance channel in the example of Section 3, so that it is "flattened",

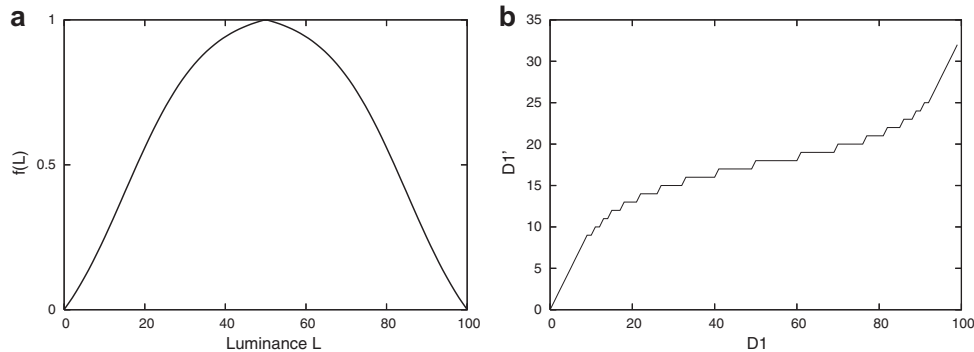


Fig. 6. The "importance" of saturation with respect to luminance in the HSL colour space with a double sigmoidal distribution (a), and the quantised form of the luminance channel using this relation (b) with $\alpha = 10$ and Algorithm 1.

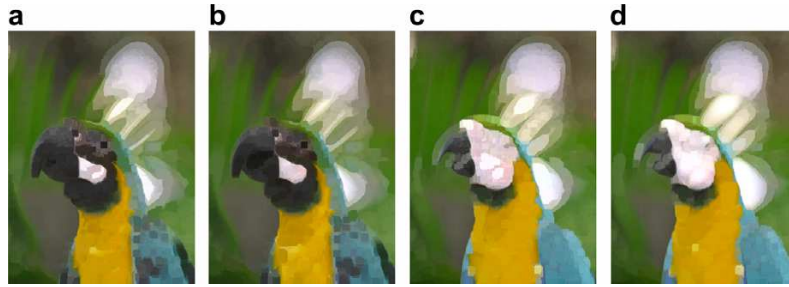


Fig. 7. The erosion and dilation results of a part of the macaws image (Fig. 10a) with a square shaped SE of size 9×9 pixels, using α -modulus lexicographical ordering 5 (a, b) and the proposed quantisation based lexicographical ordering 8 along with the priority distribution illustrated in Fig. 6a and both with $\alpha = 50$ (c, d).

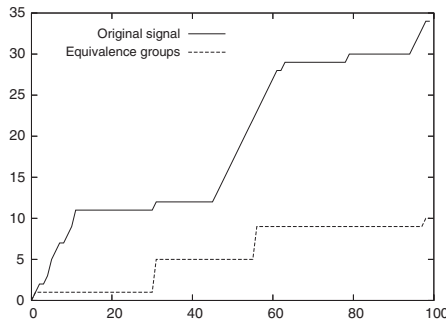


Fig. 8. Example of an uni-dimensional discrete signal with its corresponding equivalence groups obtained with expression (5) and $\alpha = 11$.

or in other words heavily smoothed. If we call this image a “marker” $m : \mathbb{Z}^2 \rightarrow \mathbb{R}$, then the ordering of vectors of image g can be realised as follows:

$$\begin{aligned} \forall x, y, s, t \in \mathbb{Z}, \quad g(x, y) < g(s, t) \\ \iff [m(x, y), g_2(x, y), \dots, g_n(x, y)]^T \\ <_L [m(s, t), g_2(s, t), \dots, g_n(s, t)]^T \end{aligned} \quad (9)$$

In other words, the formation of equivalence groups is now controlled only by the marker image. If a couple of pixels have the same value in the marker image, then they are considered equal for the dimension that this marker image represents, and the comparison outcome is determined by the following dimensions. Hence equivalence groups are now formed at the flat regions of m (Fig. 9). The “flattening” process may be achieved using a variety of filters such as large median filters, alternating sequential filters (Soille, 2003), morphological levelings (Meyer, 2004), etc. As an application example, let us return to the HLS colour space, and consider a lexicographical ordering of type $L \rightarrow S$. As such the luminance component would determine the outcome of the vast majority of comparisons. Shifting priority to the second dimension (i.e. saturation) can be achieved using for instance the quantisation example of the previous section. However, since edge related information is represented in its majority at the luminance component, it is obviously pertinent to employ

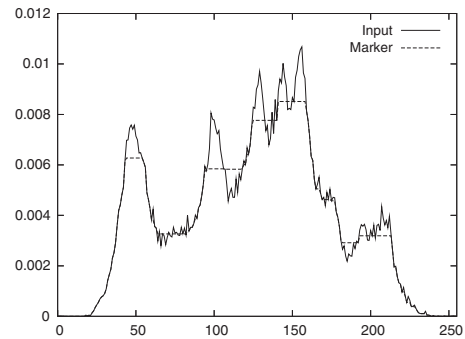


Fig. 9. Example of an uni-dimensional signal and its marker, obtained with an alternating sequential filter using line shaped SEs of length 5 pixels.

this component at object edges and borders, while using saturation for homogeneous or flat image regions. This relatively constrained situation can be resolved with expression 9 as shown in Fig. 10.

More precisely, the priority shift towards saturation can be effectively limited with only the relatively homogeneous regions of the input, by flattening them using a morphological levelling applied to the luminance component. The “priority map” shown in Fig. 10d asserts this statement. Specifically, this map associates to each pixel the percentage of comparisons within the SE that have been determined using the first dimension (i.e. luminance). Hence dark areas correspond to pixels where the outcome of vector comparisons has been determined using the saturation component, whereas bright areas represent pixels where luminance, the first vector dimension, has been used to order the vectors.

Consequently, although the creation of equivalence groups is image specific, and needs to be realised independently for each image to be processed, marker based ordering provides nonetheless a means of avoiding the artificial edge related pitfalls of the aforementioned approaches, by taking into account the spatial relationships of pixels, thus making it possible to accommodate topological restrictions during ordering.

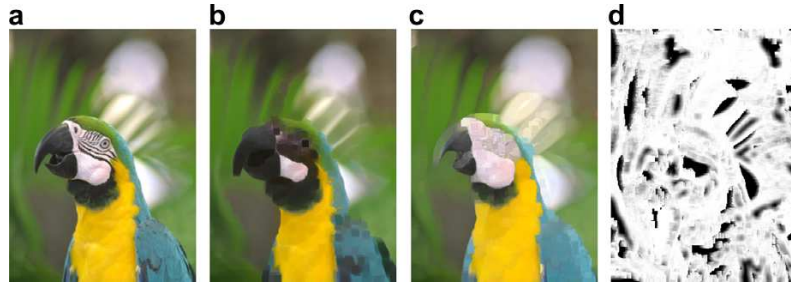


Fig. 10. The original image (a), its erosion (b) and dilation (c) results with a square shaped SE of size 9×9 pixels, using the marker based ordering 9 along as marker its morphological levelling. The marker of the levelling has been obtained by means of alternating sequential filters using the same SE. And rightmost, is the priority map (d).

5. Applications

In this section, the proposed ordering approaches 8 (QuantLex) and 9 (MarkerLex) are compared against the α -modulus (AmodLex) 5 as well as the standard lexicographical ordering (Lexico) 3. For the sake of objectivity two different applications are employed, both using colour images, noise reduction and texture classification. The other variations (i.e. α -trimmed and α -lexicographical) were omitted since they are not orderings and hence do not verify fundamental theoretical requirements. Both AmodLex and QuantLex are set with $\alpha = 10$, the image data being in the HSL colour space and in the integer interval $[0, 255]$. Once more, for the reasons given in Section 3 an ordering ignoring the hue ($L \rightarrow S$) is employed. Specifically, the priority distribution depicted in Fig. 6a is used for QuantLex. As to MarkerLex, the marker image is obtained by levelling ($\lambda = 3$) the opening–closing result of the luminance band employing a square shaped SE of size 7×7 pixels.

5.1. Colour noise reduction

For the first application the 100 test images of the Berkeley segmentation dataset have been employed (Martin et al., 2001). The image data has been corrupted in the RGB colour space with uncorrelated ($\sigma = 32$, $\rho = 0.0$) zero-mean Gaussian noise, and in order to quantify the relative performances, the relative normalized mean squared error (RNMSE) has been calculated, also in RGB

$$\text{RNMSE} = \frac{\sum_{x,y} \|f(x,y) - f_f(x,y)\|^2}{\sum_{x,y} \|f(x,y) - f_n(x,y)\|^2} \quad (10)$$

where $f(x,y)$, $f_n(x,y)$, $f_f(x,y)$ denote the vector pixels at position (x,y) , respectively, of the original, the noisy and filtered images, while $\|\cdot\|$ represents the Euclidean norm. The images have been filtered using a square shaped SE of size 3×3 pixels and the open–close close–open filter (OCCO)

$$\text{OCCO}(f) = \frac{1}{2} \phi(\gamma(f)) + \frac{1}{2} \gamma(\phi(f)) \quad (11)$$

Table 2

The average $1000 \times \text{RNMSE}$ errors against uncorrelated Gaussian noise ($\sigma = 32$, $\rho = 0.0$) using the test images of the Berkeley segmentation dataset

Luminance	Lexico	AmodLex	QuantLex	MarkerLex
565.63	565.59	568.65	564.19	651.47

where ϕ and γ denote, respectively, the vectorial closing and opening operators.

The average error rates are given in Table 2. One can easily remark the extremely similar values obtained by using only the luminance channel for comparisons and the standard lexicographical ordering. A result which asserts the highly asymmetrical priority attributed to the first vector component. Moreover, the transition of this priority to the saturation band by means of the AmodLex ordering does not result in an improvement, leading us to assume that the luminance channel is more pertinent for this task. However, the proposed QuantLex approach gives the overall best error ratio, though only slightly superior than its lexicographical counterpart, by exploiting the inner relationships among luminance and saturation. As to MarkerLex, it is obviously not suitable for the task of noise reduction, since its highly smoothed marker image prevents it from accessing the fine differences among the corrupted pixels of the luminance band.

5.2. Colour texture classification

As to texture classification, it has been chosen to employ the publicly available Outex13 colour texture database (Fig. 11) (Mäenpää and Pietikäinen, 2004; Ojala et al., 2002). The descriptors have been obtained by means of the morphological version of the autocorrelation operator, namely morphological covariance, which represents the family of joint methods for the processing of colour and textural information (Aptoula and Lefèvre, 2006; Epifanio and Soille, 2007). The vectorial version of morphological covariance K' of an image f , is defined as the volume Vol of the image (i.e. sum of pixel values), eroded by a pair of points $P_{2,v}$ separated by a vector v



Fig. 11. Examples of the 68 colour textures of Outex13 (Mäenpää and Pietikäinen, 2004).

Table 3
Classification rates in % for the textures of Outex13, using vectorial erosion based covariance

Luminance	Lexico	AmodLex	QuantLex	MarkerLex
76.03	76.05	76.16	77.17	79.86

$$K'(f; P_{2,v}) = \text{Vol}(\epsilon_{P_{2,v}}(f)) \quad (12)$$

In practice, K' is computed for varying lengths of v , and most often as also here the normalised version K is used for measurements

$$K(f) = \text{Vol}(\epsilon_{P_{2,v}}(f)) / \text{Vol}(f) \quad (13)$$

The resulting K series provides information on the periodicity and coarseness of its input (Aptoula and Lefèvre, 2007; Soille, 2003). Whereas additional information concerning the anisotropy of the textures can be obtained by plotting against not only different lengths of v , but orientations as well.

The covariance based feature vectors have been calculated using four directions for the point pairs (0° , 45° , 90° , 135°), each along with distances ranging from 1 to 49 pixels in steps of size two. Consequently 25 values are available for each direction, making a total of 100 values for every image channel after concatenation. The classification process has been realised using a kNN classifier ($k = 1$) along with the Euclidean distance.

The classification rates are shown in Table 3. As before, one can immediately remark the almost identical results of Lexico and Luminance. Once more the quantisation based lexicographical ordering leads to an improvement over AmodLex, as it controls the transitions among the image channels in a finer fashion, the difference being however almost negligible (1.01 points). Contrarily to noise reduction, the marker based approach provides the best results in this case, and outperforms the α -modulus lexicographical ordering by 3.7 points. Apparently, the use of saturation with relatively “flat” regions, and that of luminance with transitions within the image, lead to more pertinent descriptors by better combining colour and textural information.

6. Conclusion

Lexicographical ordering has proven to be an invaluable vector ordering methodology for MM, as it possesses multiple desirable theoretical properties and makes it possible to easily customize the order of comparison among image channels. Its widespread use however is mostly hindered by the usually excessive priority attributed to the first vector dimension. Among the existing approaches countering this problem that have been recalled, α -modulus lexicographical ordering stands out as the most theoretically viable solution.

In this paper, two additional solutions have been presented. First, a generalisation to α -modulus lexicographical ordering has been proposed, avoiding the implicit assumption of a constant priority distribution, hence making it possible to model arbitrarily complicated inter-channel priority relations. The inconvenience of quantisation based approaches however, is the introduction of artificial edges at equivalence group borders, as shown in Section 4. One possible way of avoiding this effect is to follow an image specific approach, where the priority shift occurs according to topological and spatial criteria.

The interest of the proposed approaches has been illustrated by means of comparative applications in the context of colour morphology. The quantisation based ordering has systematically outperformed α -modulus lexicographical ordering in noise reduction, by using a finer tuned priority transition among the image channels. As to the marker based ordering, it has provided the overall best results in terms of texture classification accuracy by limiting the use of saturation to homogeneous image regions while using luminance for edges and object transitions, hence leading to more pertinent descriptors.

Future work will concentrate on the study of the theoretical properties of the proposed approaches as well as on their more application oriented use.

Acknowledgement

The authors wish to thank the anonymous reviewers for their valuable comments.

References

- Angulo, J., 2005. Unified morphological color processing framework in a lum/sat/hue representation. In: Ronse, C., Najman, L., Decencière, E. (Eds.), Proc. 7th ISMM, Comput. Imag. Vision, vol. 30. Springer-Verlag, Dordrecht, Germany.
- Angulo, J., Serra, J., 2003. Morphological coding of color images by vector connected filters. In: IEEE Proc. 7th Internat. Symposium on Signal Process. and Its Applications (ISSPA'2003), vol. 1. Paris, France.
- Aptoula, E., Lefèvre, S., 2006. Spatial morphological covariance applied to texture classification. In: Proc. Internat. Workshop on Multimedia Content Representation. Classification and Security, vol. 4105. Istanbul, Turkey.
- Aptoula, E., Lefèvre, S., 2007. On morphological color texture characterization. In: Internat. Sympos. on Mathematical Morphology. Rio de Janeiro, Brazil.

- Aptoula, E., Lefèvre, S., 2007. Pseudo multivariate morphological operators based on α -trimmed lexicographical extrema. In: IEEE Proc. Internat. Symposium on Image Signal Processing and Analysis. Istanbul, Turkey.
- Aptoula, E., Lefèvre, S. A comparative study on multivariate mathematical morphology, *Pattern Recognition* (doi:doi:10.1016/j.patcog.2007.02.004).
- Barnett, V., 1976. The ordering of multivariate data. *J. Statist. Soc. A* 139 (3), 318–355.
- Birkhoff, G., 1967. *Lattice Theory*. American Mathematical Society, Providence.
- Chanussot, J., Lambert, P., 2000. Extending mathematical morphology to color image processing. In: *Internat. Conf. on Color in Graphics and Image Processing*, Saint-Etienne, France.
- Epifanio, I., Soille, P., 2007. Morphological texture features for unsupervised and supervised segmentations of natural landscapes. *IEEE Trans. Geosci. Remote Sens.* 45 (4), 1074–1083.
- Goutsias, J., Heijmans, H.J.A.M., Sivakumar, K., 1995. Morphological operators for image sequences. *Computer Vision and Image Understanding* 62 (3), 326–346.
- Hanbury, A., Serra, J., 2001. Mathematical morphology in the HLS colour space, In: Cootes, T., Taylor, C., (Eds.). *12th British Machine Vision Conf.* Manchester, UK.
- Hanbury, A., Serra, J., 2001. Morphological operators on the unit circle. *IEEE Trans. Image Process.* 10 (12), 1842–1850.
- Louverdis, G., Vardavoulia, M., Andreadis, I., Tsalides, P., 2002. A new approach to morphological color image processing. *Pattern Recognition* 35 (8), 1733–1741.
- Mäenpää, T., Pietikäinen, M., 2004. Classification with color and texture: jointly or separately? *Pattern Recognition* 37 (8), 1629–1640.
- Martin, D., Fowlkes, C., Tal, D., Malik, J., 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: *Proc. 8th Internat. Conf. on Computer Vision*, vol. 2. Vancouver, Canada. URL <<http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>>.
- Meyer, F., 2004. Levelings, image simplification filters for segmentation. *J. Math. Imag. Vision* 20, 59–72.
- Ojala, T., Mäenpää, T., Pietikäinen, M., Viirtola, J., Kyllönen, J., Huovinen, S., 2002. Outex: New framework for empirical evaluation of texture analysis algorithms. In: *Proc. 16th ICPR*, vol. 1. Quebec City, Canada.
- Ortiz, F., Torres, F., Angulo, J., Puente, S., 2001. Comparative study of vectorial morphological operations in different color spaces. *Proc. Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision* 4572, 259–268.
- Plaza, A., Martínez, P., Perez, R., Plaza, J., 2004. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles. *Pattern Recognition* 37 (6), 1097–1116.
- Rivest, J., 2004. Morphological operators on complex signals. *Signal Process.* 84 (1), 133–139.
- Ronse, C., 1990. Why mathematical morphology needs complete lattices. *Signal Process.* 21 (2), 129–154.
- Serra, J., 1993. Anamorphoses and function lattices. In: Dougherty, E.R. (Ed.), *Mathematical Morphology in Image Processing*. Marcel Dekker, New York, pp. 483–523, chapter 13.
- Soille, P., 2003. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, Berlin.
- Soille, P., 2005. Beyond self-duality in morphological image analysis. *Image Vision Comput.* 23 (2), 249–257.
- Talbot, H., Evans, C., Jones, R., 1998. Complete ordering and multivariate mathematical morphology. In: Heijmans, H., Roerdink, J. (Eds.), *Mathematical Morphology and Its Applications to Image and Signal Processing*. Kluwer Academic Press, Amsterdam.

Segmentation par Ligne de Partage des Eaux avec Marqueurs Spatiaux et Spectraux

Sébastien LEFÈVRE

Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (UMR 7005 CNRS-UDS)
Université de Strasbourg, Pôle API, Boulevard Brant, BP 10413, 67412 Illkirch Cedex, France
lefevre@unistra.fr

Résumé – Du fait de son rôle majeur dans les systèmes d'analyse et de traitement des images, le problème de la segmentation a fait l'objet de nombreux travaux. Cependant il n'existe pas encore de méthode automatique apte à traiter correctement tout type d'image. Les méthodes semi-automatiques sont donc une solution appropriée pour résoudre ce problème lorsqu'une faible intervention de l'utilisateur est possible. Parmi ces méthodes, l'approche par ligne de partage des eaux avec marqueurs est certainement l'une des plus employées. Dans cet algorithme, l'utilisateur fournit des marqueurs, qui ne sont utilisés qu'en tant qu'initialisations spatiales des régions produites par le processus de segmentation. Nous proposons ici d'exploiter également le contenu (spectral) des marqueurs, au travers d'une classification supervisée floue. Les résultats obtenus dans différents domaines applicatifs (imagerie couleur, télédétection) démontrent l'intérêt et la pertinence de l'approche proposée, qui ne nécessite pas d'autre paramètre que les marqueurs fournis par l'utilisateur.

Abstract – Due to its great influence in image analysis and processing systems, the problem of image segmentation has been studied for years. However there is not yet any automatic method able to correctly process any kind of image. Semi-automatic methods are then an appropriate solution to solve this problem when a limited user intervention is possible. Among these methods, the marker-based watershed approach is certainly one of the most widely used. In this algorithm, the user provides some markers, which are used only as spatial initialisations of the regions to be produced by the segmentation process. We propose here to also rely on the (spectral) content of the markers, through a fuzzy supervised classification. The results we have obtained in several application domains (color imaging, remote sensing) underline the relevance of the proposed approach, which does not need any parameter apart from the markers provided by the user.

1 Introduction

L'analyse, le traitement et l'interprétation d'images nécessitent souvent l'emploi de différents algorithmes. Parmi ceux-ci, la segmentation, qui a pour but d'extraire d'une image un ensemble de régions pertinentes, est particulièrement critique puisque la qualité des régions produites aura une incidence majeure sur le résultat du processus global. Le problème de la segmentation d'image a donc fait l'objet de nombreux travaux, cherchant à proposer des méthodes les plus automatiques ou les plus génériques possibles (concevoir une méthode de segmentation apte à traiter correctement tout type d'image de façon automatique est aujourd'hui encore impossible, et le restera sans doute longtemps puisque le problème est mal posé).

Les méthodes de segmentation semi-automatique, impliquant une (faible) intervention de l'utilisateur, permettent de traiter une variété importante d'images. Cette intervention de l'utilisateur peut prendre la forme d'une initialisation spatiale de l'algorithme, d'une définition des classes d'intérêt dans l'image, d'une pondération des différentes caractéristiques (couleur, texture, forme) à prendre en compte, etc. Parmi ces méthodes, l'algorithme de segmentation par ligne de partage des eaux (LPE) avec marqueurs [11] (une extension très communément utilisée de la méthode de segmentation LPE [12]) peut faire appel

à l'utilisateur au travers d'une initialisation spatiale définie par la position des marqueurs. Dans cet algorithme, seule la position spatiale des marqueurs est utilisée, le contenu (spectral) des marqueurs étant complètement ignoré.

Nous proposons ici d'améliorer la performance de la segmentation LPE avec marqueurs en exploitant le contenu des marqueurs. Pour cela nous utilisons les marqueurs comme des ensembles d'apprentissage dans une classification supervisée floue, le processus de segmentation s'appuyant alors sur les cartes de probabilités d'appartenance des pixels aux classes.

2 Segmentation LPE avec marqueurs

Pour décrire notre méthode qui est une extension de l'algorithme de LPE avec marqueurs, nous rappelons tout d'abord ce dernier [11] et donnons sa formulation algorithmique selon le principe de l'immersion [12] utilisé ici. Nous détaillons ensuite quelques travaux qui cherchent à mieux exploiter l'information disponible au sein du processus de segmentation.

2.1 Algorithme original de LPE avec marqueurs

Notons f l'image initiale à niveaux de gris définis dans l'intervalle $[h_{\min}, h_{\max}]$ que l'on va représenter comme un relief

topographique, et M l'ensemble des marqueurs fourni au titre de l'initialisation spatiale. Ces marqueurs vont permettre de localiser les minima que l'on souhaite imposer à la fonction f , et on définit alors le nouveau relief g comme :

$$g(p) = \begin{cases} h_{\min-1} & \text{si } p \in M \\ f(p) & \text{sinon} \end{cases} \quad (1)$$

où p représente les coordonnées d'un pixel, et $h_{\min-1}$ une nouvelle valeur spécifique aux marqueurs initiaux. La segmentation de l'image correspond à l'ensemble $X_{h_{\max}}$ des bassins de rétention de la fonction g , calculé d'après le schéma de récursion suivant :

$$\begin{aligned} X_{h_{\min-1}} &= T_{h_{\min-1}}(g) \\ X_{h+1} &= IZ_{T_{h+1}(g)}(X_h), \quad h_{\min-1} \leq h < h_{\max} \end{aligned} \quad (2)$$

où T_h est la fonction seuil de niveau h (c'est-à-dire $T_h(f) = \{p \mid f(p) \geq h\}$), et $IZ_A(B)$ est l'union des zones d'influence géodésique des composantes connexes de B avec $B \subseteq A$. Nous faisons donc appel également aux définitions suivantes :

$$IZ_A(B) = \bigcup_{j=1}^l iz_A(B_j) \quad (3)$$

$$iz_A(B_j) = \{p \in A \mid \forall k \in [1, l] \setminus \{j\} : d_A(p, B_j) < d_A(p, B_k)\} \quad (4)$$

$$d_A(a, B) = \min_{b \in B} d_A(a, b) \quad (5)$$

où $d_A(a, b)$ représente la distance géodésique entre a et b dans A . Une définition complète est donnée dans l'article original [12].

2.2 Approches LPE exploitant des informations a priori

Beare [3] exploite des connaissances spatiales disponibles a priori et introduit un moyen pour contraindre la croissance des marqueurs via l'utilisation de fonctions de distances basées sur des éléments structurants. La méthode proposée est capable de traiter des frontières d'objets incomplètes ou bruitées.

Li et Hamarneh [6] considèrent un ensemble d'images d'apprentissage avec des segmentations de références pour construire des histogrammes de forme et des descripteurs d'apparence (la moyenne et la variance des intensités des pixels des objets). Leur segmentation guidée par des connaissances s'appuie sur un algorithme de LPE classique suivi par un algorithme de clustering k-means, mais la forme et l'apparence ne sont pas prises en compte dans l'étape de segmentation par la LPE.

Dans un travail précédent [4], nous avons proposé d'introduire la connaissance sous forme de pixels étiquetés lors d'une segmentation par LPE d'images multispectrales en télédétection. Plus précisément, un ensemble de c classes prédéfinies (par exemple les toits des bâtiments, la végétation, les routes) est considéré et des pixels exemples sont donnés pour chaque classe. Une classification floue basée pixels est alors appliquée en utilisant les signatures spectrales afin de générer des cartes de probabilité représentées sous la forme d'une seule image à c bandes. Un gradient morphologique est appliqué sur cette

image et la norme euclidienne permet ensuite d'obtenir une image à niveaux de gris, sur laquelle est finalement appliquée l'algorithme de LPE. La connaissance est donc ici uniquement liée à l'information spectrale, ne tenant pas compte de la position spatiale des pixels donnés en exemple.

A la connaissance de l'auteur, la seule tentative d'utilisation des marqueurs comme source de connaissances dans la segmentation par LPE a été effectuée par Grau et al dans [5], où une hypothèse très restrictive est faite quant à la distribution normale des intensités au sein des objets de l'image. Chaque marqueur ou classe étant représenté par la moyenne et la variance des valeurs de ses pixels, les auteurs considèrent un cadre bayésien et modélisent les corrélations locales entre pixels à l'aide de champs de Markov. Le processus de segmentation est donc relativement coûteux en temps de calcul.

3 Méthode proposée

Dans la définition de l'algorithme de segmentation par LPE avec marqueurs donnée précédemment, l'ensemble des marqueurs M a été utilisé seulement pour générer l'ensemble initial $X_{h_{\min-1}}$ impliqué dans l'algorithme récursif. Nous proposons ici d'exploiter non seulement la position mais aussi le contenu des marqueurs. Notre méthode offre l'avantage de ne pas nécessiter d'autre paramètre à régler outre ces marqueurs.

Nous proposons de modifier la définition des marqueurs et considérons maintenant une collection $M = \{M_i\}_{1 \leq i \leq c}$ de c marqueurs. Chaque marqueur est un ensemble de pixels $M_i = \{p\}$, donnant ainsi lieu à une ou plusieurs composantes connexes. Ces pixels peuvent être décrits par différentes caractéristiques (intensité, couleur, signature spectrale, texture, etc). Nous associons à chaque marqueur M_i une classe C_i et appliquons une classification supervisée floue des pixels, en utilisant M_i comme ensemble d'apprentissage pour la classe C_i . La procédure de classification supervisée retourne un ensemble de probabilités $\{w_i(p)\}$ où $w_i(p)$ représente la probabilité pour que le pixel p appartienne à la classe i , avec la contrainte suivante : $\sum_{1 \leq i \leq c} w_i(p) = 1$. Le contenu d'un marqueur donné M_i permet de générer une nouvelle image w_i où les valeurs élevées représentent les pixels qui appartiennent le plus probablement à M_i . Comme le principe de LPE considère un niveau croissant h , nous définissons les fonctions

$$f_i = (1 - w_i) \cdot f \quad (6)$$

où les pixels de probabilité élevée w_i auront leur altitude initiale f réduite, alors que les pixels de faible probabilité resteront inchangés. Il est aussi possible de considérer seulement les cartes d'appartenance (soit $f_i = 1 - w_i$), voire de choisir un modèle de pondération, avec κ un coefficient à fixer :

$$f_i = \kappa \cdot (1 - w_i) + (1 - \kappa) \cdot f \quad (7)$$

Nous discuterons ces possibilités dans la prochaine section et avons opté ici pour le modèle de fusion multiplicative donné par l'équation (6). Les fonctions f_i vont être considérées comme les reliefs dans le processus de LPE par immersion.

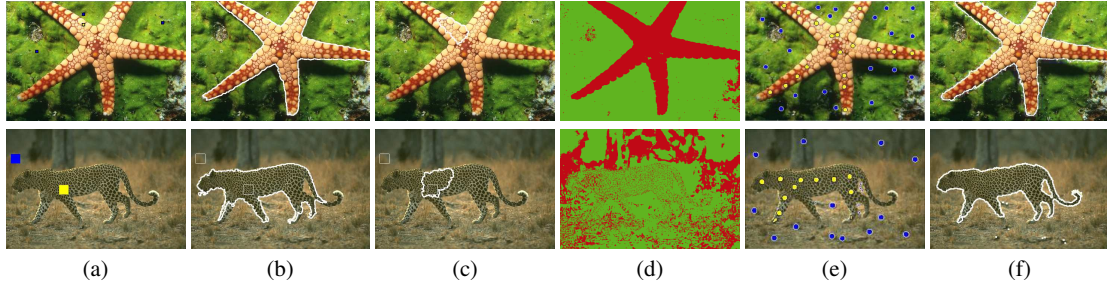


FIG. 1 – Segmentation d’images couleur : (a) image initiale avec marqueurs, (b) segmentation proposée, (c) segmentation par LPE avec marqueurs, (d) classification supervisée, (e) marqueurs et (f) résultats présentés dans [8].

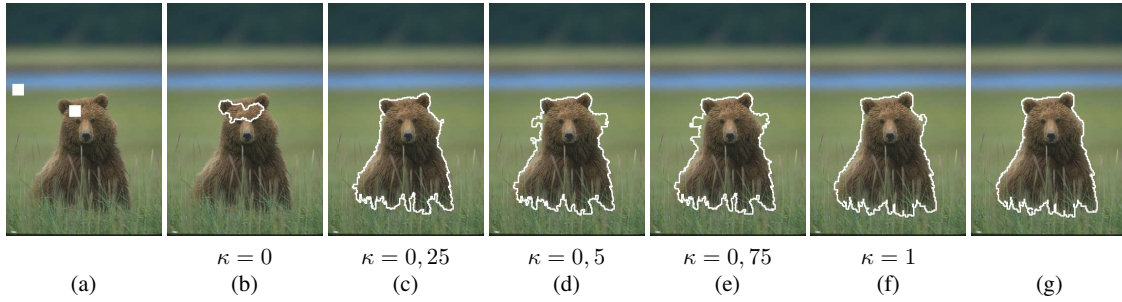


FIG. 2 – Influence du modèle de fusion entre relief initial f et probabilités d’appartenance w_i : (a) image initiale et ses deux marqueurs, (b) à (f) pondération des informations spatiales et spectrales selon l’équation (7) avec différentes valeurs de κ , et (g) fusion multiplicative adoptée ici.

L’algorithme initial de LPE nécessite une image à niveaux de gris f , tandis qu’ici la classification supervisée produit un ensemble de c images f_i . Au lieu d’imposer une étape de fusion pour combiner ces différentes images, nous considérons plutôt une image différente f_i pour chaque marqueur M_i , ce qui nécessite d’adapter l’algorithme original décrit précédemment. Nous définissons alors les fonctions g_i comme

$$g_i(p) = \begin{cases} h_{\min-1} & \text{si } p \in M_i \\ f_i(p) & \text{sinon} \end{cases} \quad (8)$$

et l’ensemble $X = \{X^i\}_{1 \leq i \leq c}$, modifiant ainsi la récursion :

$$\begin{aligned} X_{h_{\min-1}}^i &= T_{h_{\min-1}}(g_i) \\ X_{h+1}^i &= IZ_{\{T_{h+1}(g_i)\}_{1 \leq i \leq c}}(X_h), \quad h_{\min-1} \leq h < h_{\max} \end{aligned} \quad (9)$$

et considérant une définition adaptée des zones d’influence :

$$IZ_A^i(B) = \bigcup_{j=1}^l iz_A^i(B_j^i) \quad (10)$$

$$\begin{aligned} iz_A^i(B_j^i) &= iz_A(B_j^i) \cup \bigcap_{m=1}^c \{p \in A^i \cap A^m \mid \\ &\forall k \in [1, l] : d_{A^i}(p, B_j^i) < d_{A^m}(p, B_k^m)\} \end{aligned} \quad (11)$$

Autrement dit, chaque bassin de rétention est initialement défini par un marqueur donné et va s’étendre selon le relief construit à partir de ce marqueur. Plusieurs fonctions de relief f_i seront impliquées seulement dans le cas des pixels de bord qui pourraient être affectés à plusieurs bassins.

4 Résultats

Nous avons évalué notre méthode dans différents contextes applicatifs dont quelques exemples sont donnés ici. Nous avons calculé le relief initial par la norme euclidienne d’un gradient morphologique marginal sans changement d’espace couleur, et exploité les marqueurs au travers d’une classification par 5-plus proches voisins avec pondération de distances.

Nous présentons en figure 1 des exemples de résultats obtenus en quelques secondes sur la base d’images naturelles de Berkeley [7]. Nous avons naturellement comparé notre approche avec une classification supervisée et avec une segmentation LPE avec marqueurs classique, afin d’évaluer l’intérêt de notre approche sur ces deux techniques sur lesquelles elle s’appuie. D’autre part nous avons également comparé nos résultats avec ceux fournis par d’autres méthodes de segmentation, comme l’approche de Micusik et Hanbury [8] qui nécessite un plus grand nombre et un meilleur positionnement des marqueurs et qui s’appuie sur des caractéristiques plus évoluées (intensité, couleur et texture) que les seules valeurs RVB utilisées ici.

Nous pouvons observer dans la figure 2 l’intérêt du modèle de fusion multiplicative (équation (6)) par rapport à une fusion pondérée (équation (7)) et différentes valeurs de κ des informations spatiales et spectrales. La fusion proposée permet de conserver des frontières précises qui ne peuvent être obtenues même dans le cas d’une image où les objets sont spectralement

bien différenciés, et évite le problème délicat du réglage de κ .

La figure 3 présente quant à elle un exemple de résultat obtenu en télédétection à très haute résolution spatiale, où l'objectif est de fournir un outil d'aide à la localisation d'objets d'intérêt (trait de côte ici, mais aussi bâtiments par exemple). Les résultats obtenus ont convaincu des équipes de géographes d'utiliser cet outil dans leurs travaux en photo-interprétation [10].



FIG. 3 – Segmentation avec 2 marqueurs (terre et mer) d'une image multispectrale pour l'extraction du trait de côte.

5 Conclusion

Dans cet article nous nous sommes intéressés au problème de la segmentation d'image, pour lequel l'algorithme de ligne de partage des eaux avec marqueurs est une des solutions les plus fréquemment utilisées. Cet algorithme nécessite une initialisation spatiale (automatique ou manuelle) pour définir les positions initiales des bassins de rétention (i.e. les marqueurs). Cependant, aussi pertinentes soient-elles, ces informations ne sont que partiellement utilisées.

Nous avons donc proposé ici d'utiliser les marqueurs comme des ensembles d'apprentissage dans un processus de classification supervisée floue des pixels. Pour chaque marqueur ou classe, nous obtenons alors en chaque pixel la probabilité d'appartenance à la classe donnée au travers d'une carte de probabilités. Il est ensuite possible de construire à partir de cette carte le relief qui sera nécessaire à l'algorithme de ligne de partage des eaux avec marqueurs, et d'utiliser à la fois les positions spatiales et le contenu (spectral) des marqueurs. Nous avons présentés des résultats prometteurs pour différentes applications (images couleur, télédétection).

Afin de mieux évaluer les résultats de segmentation produits par notre méthode, nous souhaitons mesurer leur pertinence et précision en utilisant la base d'images segmentées de Berkeley [7]. De plus, nous étudions l'intégration d'autres caractéristiques, liées par exemple à la texture [2] ou la couleur au travers d'autres espaces de représentation [1], et d'autres classifieurs (par exemple les SVM). La robustesse de notre méthode à la position initiale des marqueurs doit également être évaluée afin de déterminer la taille minimale des marqueurs.

Enfin, nous envisageons d'étudier l'automatisation de la méthode proposée. Pour cela, nous pourrions nous inspirer des travaux de Micusik et Hanbury qui ont construit une solution de segmentation complètement automatique [9] en itérant leur mé-

thode de segmentation semi-automatique [8], supprimant ainsi le besoin du réglage manuel des marqueurs.

Références

- [1] E. Aptoula and S. Lefèvre. A comparative study on multivariate mathematical morphology. *Pattern Recognition*, 40(11) :2914–2929, Novembre 2007.
- [2] E. Aptoula and S. Lefèvre. On morphological color texture characterization. In *International Symposium on Mathematical Morphology (ISMM)*, pages 153–164, Rio de Janeiro, Brésil, Octobre 2007.
- [3] R. Beare. A locally constrained watershed transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7) :1063–1074, Juillet 2006.
- [4] S. Derivaux, S. Lefèvre, C. Wemmert, and J.J. Korczak. Segmentation par ligne de partage des eaux basées sur des connaissances texturales. In *Colloque GRETSI sur le Traitement du Signal et des Images*, Troyes, France, Septembre 2007.
- [5] V. Grau, A.U.J. Mewes, M. Alcaniz, R. Kikinis, and S.K. Warfield. Improved watershed transform for medical image segmentation using prior information. *IEEE Transactions on Medical Imaging*, 23(4) :447–458, Avril 2004.
- [6] X. Li and G. Hamarmeh. Modeling prior shape and appearance knowledge in watershed segmentation. In *Canadian Conference on Computer Vision*, 2005.
- [7] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *IEEE International Conference on Computer Vision (ICCV)*, pages 416–423, Juillet 2001.
- [8] B. Micusik and A. Hanbury. Steerable semi-automatic segmentation of textured images. In *Scandinavian Conference on Image Analysis (SCIA)*, 2005.
- [9] B. Micusik and A. Hanbury. Automatic image segmentation by positioning a seed. In *European Conference on Computer Vision (ECCV)*, 2006.
- [10] A. Puissant, S. Lefèvre, R. Desguée, and F. Levoy. Cartographie et suivi de l'évolution des schorres de la baie du mont-saint-michel (1986–2006) à partir d'images satellitaires haute résolution. *Photo-Interprétation*, 2009. A paraître, 17 pages.
- [11] J.F. Rivest, S. Beucher, and J. Delhomme. Marker-controlled segmentation : an application to electrical borehole imaging. *Journal of Electronic Imaging*, 1(2) :136–142, Avril 1992.
- [12] L. Vincent and P. Soille. Watersheds in digital spaces : An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6) :583–598, Juin 1991.

A new approach for unsupervised classification in image segmentation

Sébastien Lefèvre

Abstract Image segmentation is a fundamental problem in image analysis and understanding. Among the existing approaches proposed to solve this problem, unsupervised classification or clustering is often involved in an early step to partition the space of pixel intensities (i.e. either grey levels, colours or spectral signatures). Since it completely ignores pixel neighbourhoods, a second step is then necessary to ensure spatial analysis (e.g. with a connected component labeling) in order to identify the regions built from the segmentation process. The lack of spatial information is a major drawback of the classification-based segmentation approaches, thus many solutions (where classification is used together with other techniques) have been proposed in the literature. In this paper, we propose a new formulation of the unsupervised classification which is able to perform image segmentation without requiring the need for some additional techniques. More precisely, we introduce a *k*means-like method where data to be clustered are pixels themselves (and not anymore their intensities or colours) and where distances between points and class centres are not anymore Euclidean but topographical. Segmentation is then an iterative process, where at each iteration resulting classes can be seen as influence zones in the context of mathematical morphology. This comparison provides some efficient algorithms proposed in this field (such as hierarchical queue-based solutions), while adding the iterative property of unsupervised classification methods considered here. Finally, we illustrate the potential of our approach by some segmentation results obtained on artificial and natural images.

Sébastien Lefèvre
LSIIT – CNRS / University of Strasbourg
Pôle API, Bd Brant, BP 10413, 67412 Illkirch Cedex, France, e-mail: lefevre@unistra.fr

1 Introduction

Classification methods, either supervised or unsupervised, have always been very useful tools in the field of digital image analysis and processing, especially with the aim of image segmentation or image understanding. Conversely, images can be seen as semi-structured complex data which offer new perspectives and new challenges to the field of data mining and knowledge discovery.

We study here the existing link between these two fields, i.e. image processing and data mining. More precisely we focus on the role of (unsupervised) classification in solving one of the main problems of image processing, i.e. image segmentation. A segmentation, or a partition of an image into regions, can indeed be obtained from an unsupervised classification applied on all pixel values (e.g. grey levels, colours or spectral signatures) followed by a spatial analysis (e.g. connected component labeling). Despite being very commonly used, this segmentation strategy is not perfect since spatial information is taken into account only in a second step and is completely ignored during the classification phase. Thus numerous ad-hoc solutions have been proposed to solve this problem, as we will see in Sec. 2.2.

Instead of proposing yet another ad-hoc technique for classification-based image segmentation, we consider here the following question: is it possible to apply such a classification *directly* in the (spatial) space of image pixels ? If so, it would provide a straight way to use classification methods to solve the image segmentation problem. Of course, such a strategy would require some adaptations to let the classification method be able to produce directly a relevant segmentation result without any help from additional postprocessing. Main concepts within the classification paradigm (e.g. data space, similarity or distance measure) have to be rethought in the context of image segmentation. Moreover underlying algorithms may benefit from advances in the image processing field. Our contribution concerns such an approach, and we focus more precisely on one of the most famous unsupervised classification algorithm, i.e. the *k*means method. Thus we reformulate here the *k*means algorithm in the context of image segmentation, which does not require any additional step.

This chapter is organised as follows. Sec. 2 recalls how unsupervised classification (and especially partitional clustering) is used to solve the problem of image segmentation. After pointing out the major drawbacks of the classical segmentation-by-clustering paradigm, we then review the main solutions proposed in the literature. Our proposal is explained in more details in Sec. 3. Basically it consists in performing classification directly in the pixel space rather than in the intensity space. This requires avoiding the use of Euclidean distance and gravity centres classically used in the *k*means algorithm. In Sec. 4, we study the link with the field of mathematical morphology, in order to benefit from efficient algorithms. Thus the proposed scheme of unsupervised classification for image segmentation can also be seen as an iterative morphological segmentation process. Finally, we illustrate in Sec. 5 our approach by some first segmentation results obtained on artificial and natural images.

2 On the limitations of classification for image segmentation

In this section, we introduce the main notations and explain why image segmentation cannot be achieved with classification. We then draw up an overview of existing approaches proposed with the aim of adapting classification methods to the problem of image segmentation.

2.1 Segmentation versus classification

An image is usually defined as a function $f: \mathcal{E} \rightarrow \mathcal{T}$ which assigns to each pixel $p = (x, y)$ taken in space $\mathcal{E} \subset \mathbb{N}^2$, a value $v = f(p)$ in $\mathcal{T}$. This value may be for instance a grey level ($\mathcal{T} = \{0, \dots, 255\}$), a colour described by its tristimulus chromatic representation ($\mathcal{T} = \{0, \dots, 255\}^3$ or $\mathcal{T} = [0, 1]^3$ after normalisation) or a spectral signature ($\mathcal{T} \subset \mathbb{R}^n$ for an image with n spectral bands).

Segmentation aims at partitioning the pixel space $\mathcal{E}$ of an image f into a set of K regions $\{R_k\}_{1 \leq k \leq K}$ which are homogeneous according to a given criterion (e.g. the values v of pixels composing each region). Thus it is a function $\pi: \mathcal{E} \rightarrow \mathcal{C}$ which assigns to each pixel p the index k of component or region R_k to which it belongs. Each region R_k is built as a connected component (see Fig. 1(b)), i.e. a set of adjacent pixels (or neighbours two by two) with value k . More formally, we define a discrete path P_{pq} from pixel p to pixel q as the set of pixels $P_{pq} = \{p_i\}_{0 \leq i \leq m}$ with $p_0 = p$, $p_m = q$, and $\forall i \in \{0, \dots, m-1\}$, p_i is adjacent to p_{i+1} . Two pixels $p = (x_p, y_p)$ and $q = (x_q, y_q)$ are considered λ -adjacent if their λ -distance equals 1, that is $d_\lambda(p, q) = 1$. We define 4-distance, 8-distance and Euclidean distance respectively by:

$$d_4(p, q) = |x_p - x_q| + |y_p - y_q| \quad (1)$$

$$d_8(p, q) = \max(|x_p - x_q|, |y_p - y_q|) \quad (2)$$

$$d_E(p, q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} \quad (3)$$

but these terms used in discrete geometry / digital image processing are equivalent to more general notions of city-block / Manhattan / boxcar / absolute value (for d_4) and Chebyshev / maximum value (for d_8) distances. The neighbourhood of a pixel p is written $\mathcal{N}(p)$ and defined as the set of pixels q adjacent to p . Each path P_{pq} is associated with a cost $\omega(P_{pq})$, which may be defined as the number of pixels it contains, or relying on distances between pixels such as:

$$\omega(P_{pq}) = \sum_{i=1}^m d(p_{i-1}, p_i) \quad (4)$$

with d a given distance measure, e.g. Eqs. (1) to (3). Finally, a region R_k , as a connected component, verifies $\forall p, q \in R_k, \exists P_{pq}$ such that $\forall p_i \in P_{pq}, p_i \in R_k$ or in other words $\pi(p_i) = k$.

Unsupervised classification or data clustering aims at gathering data into homogeneous sets called classes or clusters. Applied on a digital image, the classification may be represented as a function $\pi : \mathcal{E} \rightarrow \mathcal{C}$ which assigns to each pixel p the index k of the class C_k to which it belongs. Similarly to the case of regions R_k , the content of classes C_k is expected to be homogeneous (e.g. pixels composing a given class share similar v values). However, contrary to regions R_k produced by a segmentation, classes C_k do not have the property of connexity, i.e. they are not assumed to be connected components, as illustrated by Fig. 1(c).

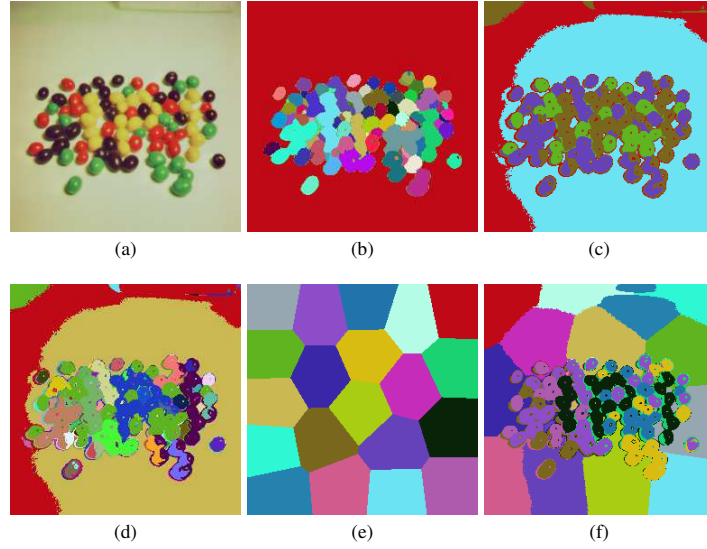


Fig. 1 Difference between segmentation and classification: (a) input image, (b) region segmentation, (c) spectral classification, (d) spatially refined classification, (e) spatial classification, (f) spatial/spectral classification.

Here we focus more precisely on the case of unsupervised classification using a k means-like approach [Kaufman and Rousseeuw, 1990]. In this algorithm, the number K of classes $\{C_k\}_{1 \leq k \leq K}$ is a priori defined (even if there exists some other partitional algorithms which overcome this condition). Each class C_k is characterized by its centre noted c_k . Initialisation of the class centres may rely on some assumptions available on the data, or may be performed randomly if no such knowledge is available. The algorithm consists of two iterative steps performed until convergence. First, for each pixel the distance to the different class centres are computed and the pixel is assigned to the class with the closest centre. Then class centres are updated by using the new data partition.

When involved to solve the problem of image segmentation, the *kmeans* algorithm is applied in the pixel value space $\mathcal{T}$ rather than the pixel space $\mathcal{E}$, i.e. data to be classified are the pixel values v rather than directly the pixels p . Alg. 1 describes this process, which tends to minimise a global cost function J , e.g. defined by:

$$J = \sum_{k=1}^K \sum_{p \in C_k} (d(f(p), c_k))^2 \quad (5)$$

Algorithm 1: original *kmeans* algorithm for image segmentation.

Input: Image $f : \mathcal{E} \rightarrow \mathcal{T} : p \mapsto f(p)$
Input: Number K of classes
Output: Set of classes $\{C\}_K$ or classification map $\pi : \mathcal{E} \rightarrow \mathcal{C} : p \mapsto \pi(p)$

```

1 /* initialisation of classification map and class centres */
2 foreach pixel  $p$  do  $\pi^0(p) \leftarrow \emptyset$ 
3 foreach class centre  $c_k$  do  $c_k^0 \leftarrow \text{RANDOM}(\mathcal{T})$ 
4 /* iterative assignment-update process */
5 repeat
6    $l \leftarrow l + 1$ 
7   foreach pixel  $p$  do /* pixel-to-class assignment */
8     computation of distances to the different class centres, i.e.  $d(f(p), c_k^{l-1})$ 
9     assignment to the class with closest centre, i.e.  $\pi^l(p) = \arg \min_k d(f(p), c_k^{l-1})$ 
10  end
11  foreach class centre  $c_k$  do /* update of class centres */
12     $c_k^l = \text{avg}\{f(p_i) \mid \pi^l(p_i) = k\}$  with avg the average fonction
13  end
14 until  $\pi^l = \pi^{l-1}$  /* stability partition as convergence criterion */
```

The distance d involved in the algorithm and in Eq. (5) is the Euclidean distance computed in the n -dimensional space $\mathcal{T}$, that is $d_E(v, w) = \sqrt{\sum_{i=1}^n (v_i - w_i)^2}$ with $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$. The convergence criterion (partition stability) may be somewhat relaxed and replaced by a convergence of the cost function, or even a finite number of iterations.

In the previous algorithm (see also Fig. 1(c)), it is clear that the location of a pixel p (i.e. its coordinates (x, y)) has been completely ignored, contrary to its value $f(p)$. Two pixels p and q with close values $f(p)$ and $f(q)$ will then be most probably assigned to the same class C_k even if they form two disjoint connected components in the classification map π . Thus some additional processings are required to obtain a segmentation (i.e. a partition of the image into homogeneous, connected classes).

2.2 Classification relying on spatial information

In order to obtain a segmentation from a classification, it is necessary to involve spatial information [Haralick and Shapiro, 1985]. We give here a brief survey of main approaches addressing this problem in the literature. In this survey, we do not consider approaches such as Markov Random Fields which combine classification and spatial regularisation, even if these approaches can be seen as generalisations of classification methods such as *kmeans* [Pappas, 1992]. We focus instead on approaches such as *kmeans* or fuzzy *cmeans*.

The most commonly adopted solution consists of a postprocessing, where a connected component labeling is applied on the classification result (see Fig. 1(d)). It aims at gathering into a single region R_k all adjacent pixels p which have been assigned to the same class C_j . A region R_k is then a connected component verifying $\forall p \in R_k, \pi(p) = j$ in the classification map. A class C_j is thus split into several regions R_k (unless if it already ensures the connexity property).

However, applying a classification at the pixel level and then gathering pixels of similar class into connected components often results in a segmentation map where many regions contain only a single isolated pixel (instead of a pixel aggregate). So a filtering step can be involved in order to remove these isolated pixels and to reassign them to neighbouring regions. In the case of a fuzzy classification, the postprocessing to be applied can be even a segmentation algorithm such as region growing [Eum et al., 1996], region merging [Chen and Lu, 2002], or marker-based watershed [Lezoray and Cardot, 2002].

Involving spatial information can also be done through a preprocessing step. This can be a segmentation into regions, which result (i.e. regions) will be further processed with a subsequent classification (but in this case the result is rather a region-based classification than a segmentation). Moreover, one can describe each pixel p by its value v but also by its coordinates (x, y) [Krishnapuram and Freg, 1992] as illustrated in Figure 1(f). It is also possible to perform an image interpolation, where the estimated value $\hat{f}(p)$ in each pixel p is computed from its 4 or 8 neighbours (e.g. through a mean or median). Classification is then applied on pixels using values $\hat{f}$ alone or in complement with initial image f [Turi, 2001, Chen and Zhang, 2004].

Another way is to consider spatial information directly within the classification algorithm. In [Ilea and Whelan, 2008], textural information can be associated with colour information. While the former is measured by a gradient, the latter is obtained by applying a *kmeans* on a smoothed image (resulting from an anisotropic filtering) and represented in two different colour spaces. The dissimilarity measure between a pixel and a class relies on these information, which are related to pixels but also to classes, and updated at each iteration in order to track the evolution of class content.

In [Liew et al., 2000], the dissimilarity between a pixel and a class centre within the fuzzy *cmeans* algorithm is defined using the neighbourhood of the pixels: contribution of each neighbour (computed as a distance to the corresponding class centre) is proportional to the similarity between this neighbour and the considered pixel (the similarity being measured in the feature space). This approach is inspired from [Tolias and Panas, 1998] where similarity between the pixel to be classified and its

neighbour modifies the class membership probability of this pixel (by adding or removing of a constant). This similarity between neighbours can be directly computed in the class membership space [Noordam et al., 2000, Xia et al., 2007]. A similar approach using a probabilistic classification based on neighbour kriging indicators is proposed in [Pham, 2001].

Applying a classifier in a multiscale framework, through a hierarchical method, offers another mean to take into account spatial context. In [Luo et al., 2003], the authors propose a new algorithm where the following steps are applied successively: image subsampling, computation of features related to colour and texture, spatially constrained classification, and region merging. These two last steps are performed iteratively at each level of the hierarchy, thus allowing to isolate the main connected components (and to aggregate remaining pixels to these components), and then to apply once again the *kmeans* algorithm independently on each connected component. This principle of successive classifications is also presented in [Cheng and Sun, 2000] where clusters are built only from pixels with homogeneous neighbourhood. Finally, the *kmeans* algorithm can be applied on a sliding window within the image, where class centres are propagated during image scan [Leydier et al., 2004].

To summarize, various solutions for spatialising classification methods have been proposed, in order to make them adequate for image segmentation: preprocessing, postprocessing, new attributes, multiscale approach, classification correction of one pixel depending on its neighbours, etc. However, to the best of the author's knowledge, applying a classification directly in the spatial domain (i.e. pixel space) has not been studied yet, despite the fact that it seems a very intuitive solution to consider the spatial behaviour of the segmentation. So we propose in next section such a solution, which can be seen as a spatial classification using the *kmeans* algorithm.

3 Another usage of classification in image segmentation

To be applied successfully on the problem of image segmentation, a classification method has to rely on spatial information. Instead of using additional steps in the classification process, here we rather reformulate the classification in a spatial context. Thus we propose to spatialise the classification method in order to make it usable to perform image segmentation.

3.1 Spatial classification

We have pointed out in Sec. 2.1 that the only difference between a segmentation and a classification was the lack of spatial connectivity constraint in the result of the latter. Modifying the behaviour of a classification algorithm in order to ensure spatial connectivity of the produced classes (or regions in this context) is a way to elaborate a

classification method intrinsically able to solve the problem of image segmentation. Since various classification algorithms exist, we have decided here to focus on one of its most representative, the *k*means algorithm, and to study how the spatial connectivity property can be provided to it. In other words, our contribution is related to the spatialisation of the *k*means algorithm.

Spatial information is directly present in images through notions of connectivity and adjacency between neighbouring pixels. Thus we propose to apply the classification in the pixel space $\mathcal{E}$ and not anymore in a feature space (e.g. the pixel value space $\mathcal{T}$). The different operations (distance d between points and class centres during assignment step, update of class centres with mean function *avg* during update step) are also performed in the space $\mathcal{E}$, considering p rather than $f(p)$ in the different formula. Similarly, class centres c_k will be defined in $\mathcal{E}$ (i.e. they are image pixels) and not anymore in $\mathcal{T}$. Thus the criterion to be optimized in the clustering process is slightly modified to become

$$J = \sum_{k=1}^K \sum_{p \in C_k} d(p, c_k) \quad (6)$$

while the original criterion given in Eq. (5) was using $f(p)$.

However, applying directly the *k*means algorithm in the space $\mathcal{E}$ will not have the expected result for segmentation: since image pixels are regularly located on a square grid of finite size, ignoring pixel values during pixel-to-class assignment step will lead to a result independent of the image content (Fig. 1(e))! An intermediary solution has to be elaborated in order to take into account both pixels p and their values $f(p)$ during the computation of distance $d(p, c_k)$ between pixels and class centres. This will be studied in next section.

Another problem has also to be solved to ensure an adequate behaviour of the *k*means algorithm in the context of segmentation. It is related to the inherent inability of the algorithm to deal with non convex sets. During a segmentation step, it is not feasible to assume that all regions would be convex. In the case of a concave region, updating the class (or region) centre based on the computation of the gravity centre can result in locating the centre outside the class. As we will see in the next section, we rely on the assumption that the centre belongs to its region or class to ensure the spatial connectivity of produced classes. Thus we need to replace the gravity centre by another measure and we propose to use here a kind of median (thus leading to a *k*median-like algorithm rather than a *k*means-like one). Moreover, we propose to keep class centres in $\mathcal{E}$ (i.e. choosing the class centres among the pixels present in the image), like a *k*prototypes approach [Han et al., 2001] (but we still keep the overall process of the *k*means algorithm). In order to obtain such a median measure in $\mathcal{E}$, we propose to define the class centre as its central or inner most pixel, that can be obtained by means of techniques from mathematical morphology [Soille, 2003]. Implementation details will be given in Sec. 4.

Let us notice that we have chosen to consider here independently the pixel space $\mathcal{E}$ and the pixel value space $\mathcal{T}$. We could have formulate directly a solution in the space $\mathcal{E} \times \mathcal{T}$, e.g., by setting some constraints on the classes to be obtained or

by defining some appropriate representation and allocation functions. But applying a classification in such a space $\mathcal{E} \times \mathcal{T}$ would have brought a much higher computational complexity. Nevertheless, since our algorithm shares some common properties with dynamic clouds [Diday, 1971] or *kmedoids* [Kaufman and Rousseeuw, 1990] for instance, the link with such methods should be studied more deeply.

3.2 Topographic distance and connexity property

When assigning a pixel to a given class, we have to take into account both pixel coordinates and value. Thus we propose to combine p and $f(p)$ through a topographical distance for which we give some definitions below. Let us notice that updating class centres can also rely on this distance and on values p and $f(p)$, as we will see further.

Let us represent an image as a topographical surface, where each pixel p is associated to its elevation, i.e. its value $f(p)$. In this context, the cost of a path P_{pq} can be reformulated by taking into account its topography or the observed difference in height between p and q , using both the difference between values $f(p)$ and $f(q)$ and the (spatial) distance $d(p, q)$ between pixels p and q (or more precisely, between their respective locations, involving a distance measure among those already defined in Eqs. (1)-(3)). Thus, several definitions for the topographical cost of a path have been proposed [Prêteux, 1992, Meyer, 1994, Philipp-Foliguet, 2000]. The cost term introduced by Prêteux [Prêteux, 1992] can be simply formulated in the case of a finite discrete image as the highest pixel on the path:

$$\omega(P_{pq}) = \max_{p_i \in P_{pq}} f(p_i) \quad (7)$$

or in other words the path ridge, with the convention $\omega(P_{pp}) = -\infty$.

Meyer [Meyer, 1994] considers the steepest slope defined for a pixel p as

$$\delta(p) = \max_{\substack{q \in \mathcal{N}(p) \\ f(q) < f(p)}} \left(\frac{f(p) - f(q)}{d(p, q)} \right) \quad (8)$$

which can be simply computed with a morphological erosion. The topographical cost between two neighbouring pixels is defined as

$$q \in \mathcal{N}(p), \quad \omega(p, q) = \begin{cases} \delta(p) & \text{if } f(p) > f(q) \\ (\delta(p) + \delta(q))/2 & \text{if } f(p) = f(q) \\ \delta(q) & \text{if } f(p) < f(q) \end{cases} \quad (9)$$

The topographical cost of a whole path is then computed as

$$\omega(P_{pq}) = \sum_{i=1}^m (d(p_{i-1}, p_i) \cdot \omega(p_{i-1}, p_i)), \quad P_{pq} = \{p_i\}_{0 \leq i \leq m} \quad (10)$$

Philipp-Foliguet [Philipp-Foliguet, 2000] proposes a simpler definition which avoids to seek for ridges or steepest slopes, using a weighting coefficient κ (usually $\kappa = 1$) :

$$\omega(P_{pq}) = \sum_{i=1}^m (\kappa \cdot d(p_{i-1}, p_i) + |f(p_{i-1}) - f(p_i)|), \quad P_{pq} = \{p_i\}_{0 \leq i \leq m} \quad (11)$$

Finally, whatever the definition of the topographical cost in use, the topographical distance $d_T(p, q)$ is defined as the minimum cost of a path from p to q :

$$d_T(p, q) = \min_{P_{pq}} (\omega(P_{pq})) \quad (12)$$

Let us notice that d_T is a *true* distance function only with the definition of Philipp-Foliguet. In the other cases, we have $\forall p, q, d_T(p, q) = 0 \nleftrightarrow p = q$. Prêteux proposes to use $\exp(d_T(p, q))$ while Meyer presents different techniques for modifying the topographical surface in order to deal with the problematic case of plateaus (where we have $d_T(p, q) = 0$ with $p \neq q$). Let us state that another pseudo-distance measure (i.e. a distance function that does not satisfy the identity of indiscernibles) could be obtained by simply defining the cost as

$$\omega(P_{pq}) = \sum_{i=1}^m |f(p_i) - f(p_{i-1})|, \quad P_{pq} = \{p_i\}_{0 \leq i \leq m} \quad (13)$$

In the following, we will use a simplified version of the Philipp-Foliguet measure:

$$\omega(P_{pq}) = \sum_{i=1}^m \omega(p_{i-1}, p_i), \quad P_{pq} = \{p_i\}_{0 \leq i \leq m} \quad (14)$$

and consider a 8-distance where the topographical cost $\omega(p, q)$ between two neighbouring pixels p and q is given by

$$\omega(p, q) = |f(p) - f(q)| + \varepsilon, \quad q \in \mathcal{N}_8(p) \quad (15)$$

with the convention $\omega(p, p) = 0$, and $\varepsilon \ll \nabla f$ being a very small term (i.e. lowest than the minimal absolute difference between values of two neighbouring pixels) ensuring the identity of indiscernibles. An illustration of this distance is given in Fig. 2 through the computation of a distance transform from a given point in a grey level image.

We recall that our goal is to ensure the spatial connexity of classes C_k produced by the classification algorithm. Let us check that this property is verified by the pixel-to-class assignment step. Using the definition introduced in Sec. 2.1, a class C_k is connected if $\forall p, q \in C_k, \exists P_{pq}$ such that $\forall p_i \in P_{pq}, p_i \in C_k$. To simplify, we will use in this case the notation $P_{pq} \in C_k$. The assignment of a pixel p to C_k is performed



Fig. 2 Illustration of the topographical distance computed from the top-left pixel (from left to right): input image, topographical distance transform, look-up table representing the distance scale.

if $d_T(p, c_k) < d_T(p, c_j)$, $\forall j \neq k$. Obviously the centre c_k of the class C_k will stay assigned to C_k since $d_T(c_k, c_k) = 0$, $\forall j \neq k$. The connexity notion can then be written as $\forall p, q \in C_k$, $\exists (P_{pc_k}, P_{qc_k}) \in C_k$ or simply by $\forall p \in C_k$, $\exists P_{c_k p} \in C_k$. Let us denote by q the pixel preceding p in this path $P_{c_k p}$ of minimal cost, we have then $q = \arg \min_{p_i \in \mathcal{N}(p)} d_T(c_k, p_i)$ and $d_T(c_k, p) = d_T(c_k, q) + d_T(q, p)$. We can prove the connexity of C_k by showing that the assumption $q \notin C_k$ is not valid. Let us assume $q \in C_j$, $j \neq k$, we have then $d_T(c_k, q) > d_T(c_j, q)$. Combining the previous formula, we obtain $d_T(c_k, p) > d_T(c_j, q) + d_T(q, p) \geq d_T(c_j, p)$. This inequality is not valid since it would mean the assignment of p to C_j . So the proposed scheme for pixel-to-class assignment ensures to produce connected classes during the classification process.

4 Implementation issues

The k means algorithm¹ that we proposed here differs from its original version only on the few following aspects, as underlined in Algorithm 2: (1) the space where pixels are represented and centres are selected ($\mathcal{E}$ instead of $\mathcal{T}$); (2) the distance measure $d_T(p, c_k)$ which is not Euclidean anymore but rather topographical; (3) the class centre computation method which does not rely on the mean but rather on the median. These changes ensure our algorithm an adequate behaviour for image segmentation, since the resulting classification is also intrinsically a segmentation.

¹ Les ut observe that the term *kmeans* is not perfectly adequate here, since our algorithm is more of *kmeans-like* type.

Moreover, they provide a link with the field of mathematical morphology, which brings some efficient implementations.

Algorithm 2: Proposed *k*means algorithm for image segmentation.

Input: Image $f : \mathcal{E} \rightarrow \mathcal{T} : p \mapsto f(p)$
Input: Number K of classes (i.e. regions)
Output: Set of classes $\{C\}_K$ or classification (i.e. segmentation) map $\pi : \mathcal{E} \rightarrow \mathcal{C} : p \mapsto \pi(p)$

```

1 /* initialisation of classification map and class centres */
2 foreach pixel  $p$  do  $\pi^0(p) \leftarrow \emptyset$ 
3 foreach class centre  $c_k$  do  $c_k^0 \leftarrow \text{RANDOM}(\mathcal{E})$ 
4 /* iterative assignment-update process */
5 repeat
6    $l \leftarrow l + 1$ 
7   foreach pixel  $p$  do /* pixel-to-class assignment */
8     computation of topographical distances to the different centres, i.e.  $d_T(p, c_k^{l-1})$ 
9     assignment to the class with closest centre, i.e.  $\pi^l(p) = \arg \min_k d_T(p, c_k^{l-1})$ 
10  end
11  foreach class centre  $c_k$  do /* update of class centres */
12     $c_k^l = \text{med}_T\{p_i \mid \pi^l(p_i) = k\}$  with  $\text{med}_T$  the topographical median fonction
13  end
14 until  $\pi^l = \pi^{l-1}$  /* stability partition as convergence criterion */

```

Indeed, the topographical distance has been used in the context of mathematical morphology to define segmentation methods such as watershed or skeleton by influence zone [Prêteux, 1992, Meyer, 1994]. The pixel-to-class assignment step of our *k*means algorithm can then be seen as a morphological segmentation with markers corresponding to class centres. Thus it is possible to benefit from algorithmic developments in mathematical morphology, and especially hierarchical queue-based algorithms for marker-based segmentation for which we can find implementation details in the book of Soille [Soille, 2003].

As already stated, updating class centres can also be done using morphological methods. We propose to define the centre c_k of class C_k as the topographical median med_T of C_k . More precisely, we follow some definitions of multidimensional medians given in the literature [Small, 1990], and consider the convex hull peeling paradigm to determine the inner most element among data (or cluster elements). Adapted to the topographical context, the topographical median med_T of a class C_k is identified as the furthest pixel from the cluster borders. It can be obtained using the maximum of the distance transform between pixels of C_k to the background defined by $\mathcal{E} \setminus C_k$ (the background may be cleverly limited to exterior borders of C_k). This centre is defined by

$$c_k = \arg \max_{p \in C_k} d(p, \mathcal{E} \setminus C_k) \quad (16)$$

with

$$d(p, X) = \min\{d(p, x) \mid x \in X\} \quad (17)$$

the distance of p to the set X . The distance $d(p, x)$ can be the Euclidean distance d_E or preferably the topographical distance d_T , thus ensuring a better stability to the algorithm since it is the same distance measure which is used in the assignment step. Coherence between these two steps reinforces algorithm convergence through the optimisation of the quality criterion given in Eq. (6) and considering topographical distances. However this convergence cannot be completely ensured since the proposed centre computation scheme can result in outliers (i.e. the class centre maximizes the distance to other classes but may not minimize the distance within its class) and subsequently in oscillations within the iterative process. Thus, in order to guarantee the algorithm convergence, another criterion may be involved: maximal number of iterations, minimal displacement of class centres between two iterations, convergence of the global cost function, etc. Let us remark that we do not observe any lack of convergence with results given in Sec. 5. Nevertheless, this problem of convergence will be the topic of further studies.

A naive implementation of the proposed algorithm consists in seeking for the shortest path in a graph to measure the distances $d_T(p, c_k)$ between pixels and class centres. This requires $O(P^2)$ operations per pixel, or $O(P)$ if optimal data structures are used, with P the number of pixels. Considering that this computation has to be performed for all pixels and all classes, this results in a global complexity of $O(KP^3)$ (or $O(KP^2)$ with appropriate data structures) per iteration, with K the number of classes.

Since our method can be assimilated to an iterative morphological segmentation, using efficient morphological algorithms based on hierarchical queues helps to limit the algorithm computational complexity. Indeed this complexity can be independent of the number of classes K , roughly $O(P)$ by iteration. More precisely, during the assignment step, we do not compute exhaustively the K distance transforms since it is sufficient to know in each pixel which is the closest centre. By propagating distances from class centres, we can process and classify iteratively the pixels of increasing distances inserted in the queue (where only the first occurrence of each pixel is considered). Updating class centres can be obtained using the same principle and requires to apply K incomplete distance transforms (each one being limited to its related class), which is equivalent to process only once each pixel. Complexities of these two steps are thus linear in $O(P)$. Let us also observe that since the final distances between pixels and classes centres are only sums of local topographical distances between neighbouring pixels, these local distances may be precomputed in order to further limit computation time (this requires a buffer of size 4 or 8 times the image size, depending on the adjacency in use, i.e. 4- or 8-adjacency). To compare, we recall that k means algorithm has a cost equal to $O(KP)$ per iteration, K being possibly high (but still lower than P) since it corresponds here to the number of objects present in the image and not anymore to the number of classes.

Algorithm 3 describes the proposed efficient implementation of the topographical k means algorithm for image segmentation. It relies on the $\text{TOPO}(f, \{s_k, S_k\})$ function defined in Algorithm 4 which returns both a classification map π and a distance

Algorithm 3: Efficient proposed kmeans algorithm for image segmentation.

Input: Image $f : \mathcal{E} \rightarrow \mathcal{T} : p \mapsto f(p)$
Input: Number K of class (i.e. regions)
Output: Set of classes $\{C\}_K$ or classification (i.e. segmentation) map $\pi : \mathcal{E} \rightarrow \mathcal{C} : p \mapsto \pi(p)$

```

1 /* initialisation of classification map and class centres */
2 foreach pixel  $p$  do  $\pi^0(p) \leftarrow \emptyset$ 
3 foreach class centre  $c_k$  do  $c_k^0 \leftarrow \text{RANDOM}(\mathcal{E})$ 
4 /* iterative assignment-update process */
5 repeat
6    $l \leftarrow l + 1$ 
7   /* pixel-to-class assignement */
8    $(\pi^l, \phi^l) \leftarrow \text{TOPO}(f, \{c_k^{l-1}, \pi(c_k^{l-1})\})$ 
9   /* update of class centres */
10   $\Omega^l \leftarrow \{p_i \mid \exists q \in \mathcal{N}(p_i), \pi^l(q) \neq \pi^l(p_i)\}$  /* identify border pixels */
11   $(\tilde{\pi}^l, \tilde{\phi}^l) \leftarrow \text{TOPO}(f, \{\Omega^l, \pi^l(\Omega^l)\})$  /* distance transform from borders */
12  foreach class  $C_k$  do
13     $c_k^l = \max\{\tilde{\phi}^l(p_i) \mid \tilde{\pi}^l(p_i) = k\}$  /* use the furthest points from the borders */
14  end
15 until  $\pi^l = \pi^{l-1}$  /* stability partition as convergence criterion */

```

Function TOPO($f, \{s_k, S_k\}$).

Input: Image $f : \mathcal{E} \rightarrow \mathcal{T} : p \mapsto f(p)$
Input: Set $\{s_k, S_k\}$ of initial pixels s_k with related classes S_k
Output: Classification (i.e. segmentation) map $\pi : \mathcal{E} \rightarrow \mathcal{C} : p \mapsto \pi(p)$
Output: Topographic distance transform $\phi : \mathcal{E} \rightarrow \mathbb{R} : p \mapsto \phi(p)$

```

1 /* initialisation of  $\pi$ ,  $\phi$ , and the hierarchical queue */
2 foreach  $p \in \mathcal{E}$  do
3    $\pi(p) \leftarrow \emptyset$ 
4    $\phi(p) \leftarrow 0$ 
5 end
6 foreach  $s_k$  do INSERT( $s_k, S_k, 0$ )
7 /* scan all pixels with increasing  $d_T$  */
8 while NOTEMPTY do
9    $(p, C_p, d_p) \leftarrow \text{REMOVE}$ 
10  if  $\pi(p) \neq \emptyset$  then continue /* ignore already labeled pixels */
11   $\pi(p) \leftarrow C_p$  /* assign the pixel to the class with closest centre */
12   $\phi(p) \leftarrow d_p$  /* set the distance between the pixel and the closest class centre */
13  foreach  $q \in \mathcal{N}(p), \pi(q) = \emptyset$  do /* process all unlabeled neighbouring pixels */
14     $d_q \leftarrow d_p + d_T(p, q)$  /* propagate topographical distance */
15    INSERT( $q, C_p, d_q$ ) /* add the pixel to the queue */
16  end
17 end
18 return  $\pi, \phi$ 

```

transform function ϕ from a set of initial points s_k which classes S_k are assumed to be known. As we have already noticed, this function is used twice in order to compute both distances to the class centres in the assignment step and distances to the class borders in the updating step. Thus we distinguish both results and note them respectively (π, ϕ) and $(\tilde{\pi}, \tilde{\phi})$. As for the `TOPO` function, it assumes the following operations being available to manipulate the hierarchical queue: `NOTEMPTY` checks if the queue contains at least one element, `INSERT`(a, b, c) insert the element a of class b at priority c , and `REMOVE` returns and removes the current triplet (a, b, c) from the queue. A hierarchical queue may be implemented as an array of FIFO queues (each queue being indexed by its priority), thus this last operation seeks for the non-empty queue of lowest priority and returns and removes the oldest element within this queue.

5 Experiments

In order to underline the potential interest of our proposal, we have performed some experiments both on artificial and natural images. Fig. 3 shows the behaviour of our algorithm compared to a standard *kmeans* applied on grey levels, facing an increasing level of noise. We consider 3 classes (i.e. regions) in our algorithm while 2 classes (expected to be black and white) are used for the standard *kmeans*. As we can observe in this figure, noise robustness achieved by the proposed algorithm seems satisfactory, contrary to the original *kmeans* method.

The iterative process of the *kmeans* algorithm is illustrated in Fig. 4. Even if the centres are randomly located on the image, the proposed algorithm manages to correctly identify the objects present in the image as clusters in the classification paradigm. Borders between classes (i.e. regions) are displayed in white while the colours represent distances from the closest class centre, using the same colour scale as in Fig. 2, i.e. increasing from magenta to red.

Similarly to the standard use of the *kmeans* procedure in unsupervised classification, the number k of classes should be chosen very carefully. As observed in Fig. 5, an appropriate number of classes (i.e. equal to the number of objects to be segmented) will help the algorithm to produce the expected result. On the contrary a too low or too high value of k may result in undersegmentation and oversegmentation, respectively.

Finally, let us reuse the introductory image of Fig. 1. Here we have compared our algorithm with a standard *kmeans* applied on the pixel colours (represented as RGB-valued vectors) to produce 6 colour clusters, followed by a connected component labeling to identify all connected groups of pixels sharing the same colour class. The 1800 regions produced by this approach are shown in Fig. 6 (centre). On the contrary, we can define the expected number of regions as classes in the *kmeans* algorithm with our approach. Thus we have segmented the marbles using 60 classes (i.e. 60 regions), as illustrated by Fig. 6 (right). The result, despite not being perfect, is rather satisfactory and shows the relevance of our proposal.

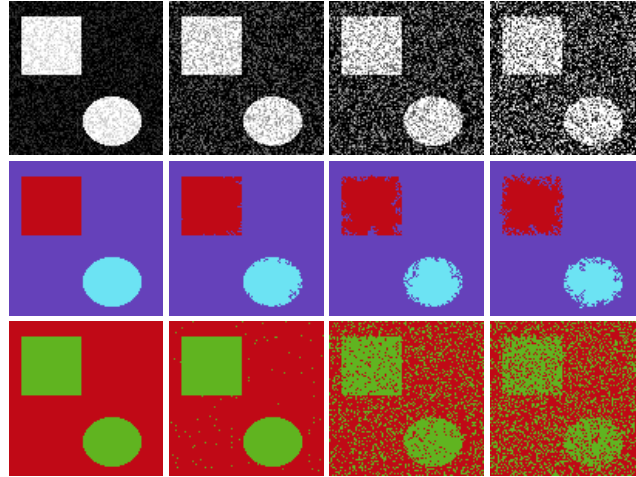


Fig. 3 Classification results on an artificial image with an increasing level of Gaussian noise (from left to right): noisy input image (top), proposed classification with 3 regions (middle), k means performed on intensities with 2 classes (bottom).

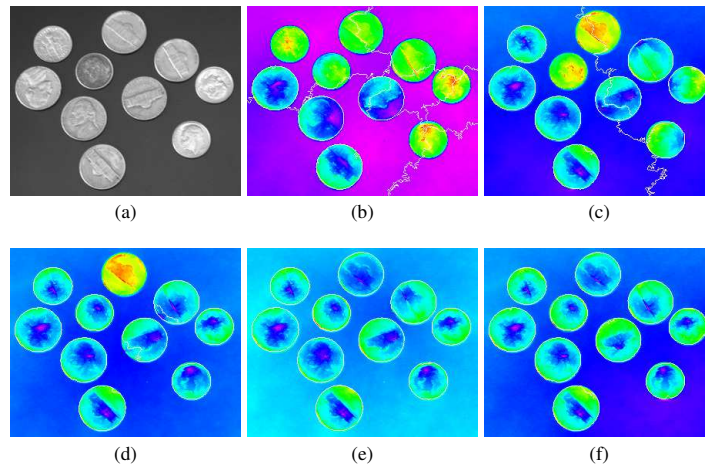


Fig. 4 Illustration of the iterative topographical k means process: (a) input image, topographical distances from centres ϕ^l and resulting classification borders Ω^l after iteration $l = 1$ (b), $l = 2$ (c), $l = 3$ (d), $l = 4$ (e), and $l = 5$ (f) where convergence has been observed.

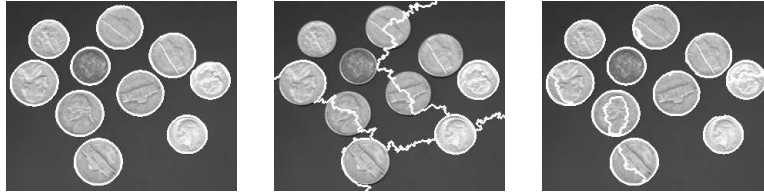


Fig. 5 Influence of the number of classes: (left) correct segmentation with 11 classes, (centre) undersegmentation with 6 classes, (right) oversegmentation with 16 classes.

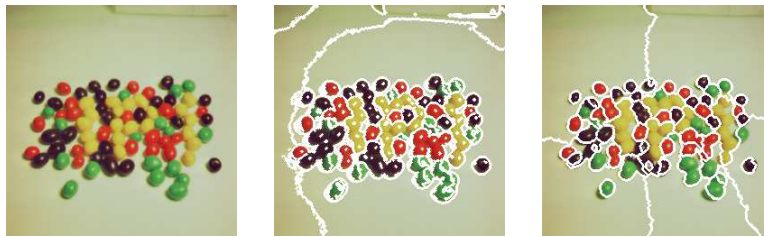


Fig. 6 Relevance of the proposed classification scheme: colour input image (left), spectral k-means followed by connected component labeling (centre), topographical k-means (right).

6 Conclusion

Classification is one of the most common solutions to solve the problem of image segmentation, i.e. transforming an image from the space of pixels to a set of objects which can be subsequently analysed in order to understand the image content. However, classification methods are often applied on the pixel value space, thus ignoring spatial information despite its primordial importance in the context of segmentation. In order to solve this problem and consider the spatial context of each pixel within a classification procedure, the numerous solutions proposed in the literature share the need for additional steps in the classification-based image segmentation process.

In this chapter, we propose a completely different approach which consists in applying the classification in the spatial domain (of the pixels) rather than in the spectral domain (of the pixel values). Thus we can keep the classification method as is (with some adaptations related to the spatial application of the method) and avoid to rely on additional ad-hoc steps. We illustrate our proposal with the spatialisation of the k means algorithm, by replacing the Euclidean distance by a topographical distance and by modifying the class centre computation scheme. Besides, a link with the field of mathematical morphology provides efficient algorithms for segmentation, with a complexity in $O(P)$ (i.e. linear and not depending on the number of classes k).

This new approach for unsupervised classification in image segmentation brings several perspectives. Among the major future works, we can mention the need

to overcome the problems related to the *k*means algorithm (e.g. choice of initial class centres, prior knowledge of the number of classes). Another research direction is to consider a fuzzy paradigm since this has already shown some interest over hard classification (see for instance the Fuzzy C-Means algorithm), and to perform a fuzzy assignment of pixels to classes [Philipp-Foliguet, 2000]. Moreover, it could be relevant to involve spatial constraints within the classification [Han et al., 2001] which in the context of segmentation may be adequate to deal with high gradient areas observed on region borders. A study on robustness and efficiency of the algorithm and experiments on larger datasets are worth being made to understand the pros and cons of our proposal. In particular, we will further study the problem of algorithm convergence. Finally, we also consider to compare our proposal with other recent methods for classification and segmentation, such as mean shift [Comaniciu and Meer, 2002], path-based clustering [Fischer and Buhmann, 2003] or its extension called robust path-based spectral clustering [Chang and Yeung, 2008], since we believe our contribution is not limited to the *k*means algorithm.

Acknowledgements The author wishes to thank his colleague Dr. Alexandre Blansch  from LSIT – CNRS / University of Strasbourg for pointing out the possible lack of convergence of the proposed algorithm, which will be the topic of further studies.

References

- [Chang and Yeung, 2008] Chang, H. and Yeung, D. (2008). Robust path-based spectral clustering. *Pattern Recognition*, 41:191–203.
- [Chen and Zhang, 2004] Chen, S. and Zhang, D. (2004). Robust image segmentation using fcm with spatial constraints based on new kernel-induced distance measure. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 34(4):1907–1916.
- [Chen and Lu, 2002] Chen, T. and Lu, Y. (2002). Color image segmentation – an innovative approach. *Pattern Recognition*, 35(2):395–405.
- [Cheng and Sun, 2000] Cheng, H. and Sun, Y. (2000). A hierarchical approach to color image segmentation using homogeneity. *IEEE Transactions on Image Processing*, 9(12):2071–2082.
- [Comaniciu and Meer, 2002] Comaniciu, D. and Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.
- [Diday, 1971] Diday, E. (1971). Une nouvelle m thode en classification automatique et reconnaissance des formes: la m thode des nu es dynamiques. *Revue de Statistique Appliqu e*, 19(2):19–33.
- [Eum et al., 1996] Eum, K., Lee, J., and Venetsanopoulos, A. (1996). Color image segmentation using a possibilistic approach. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 1150–1155.
- [Fischer and Buhmann, 2003] Fischer, B. and Buhmann, J. (2003). Path-based clustering for grouping of smooth curves and texture segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4):513–518.
- [Han et al., 2001] Han, J., Kamber, M., and Tung, A. (2001). Spatial clustering methods in data mining: a survey. In Miller, H. and Han, J., editors, *Geographic Data Mining and Knowledge Discovery*. Taylor & Francis.

- [Haralick and Shapiro, 1985] Haralick, R. and Shapiro, L. (1985). Image segmentation techniques. *Computer Vision, Graphics and Image Processing*, 29:100–132.
- [Ilea and Whelan, 2008] Ilea, D. and Whelan, P. (2008). Ctex – an adaptive unsupervised segmentation algorithm based on color-texture coherence. *IEEE Transactions on Image Processing*, 17(10):1926–1939.
- [Kaufman and Rousseeuw, 1990] Kaufman, L. and Rousseeuw, P. (1990). *Finding groups in data: an introduction to cluster analysis*. Wiley, New York.
- [Krishnapuram and Freg, 1992] Krishnapuram, R. and Freg, C. (1992). Fitting an unknown number of lines and planes to image data through compatible cluster merging. *Pattern Recognition*, (25):385–400.
- [Leydier et al., 2004] Leydier, Y., Bourgeois, F. L., and Emptoz, H. (2004). Sérialisation du k-means pour la segmentation des images en couleur: Application aux images de documents et autres. In *Colloque International Francophone sur l'Ecrit et le Document (CIFED)*.
- [Lezoray and Cardot, 2002] Lezoray, O. and Cardot, H. (2002). Cooperation of color pixel classification schemes and color watershed : a study for microscopical images. *IEEE Transactions on Image Processing*, 11(7):783–789.
- [Liew et al., 2000] Liew, A., Leung, S., and Lau, W. (2000). Fuzzy image clustering incorporating spatial continuity. *IEE Proceedings on Vision, Image and Signal Processing*, 147(2):185–192.
- [Luo et al., 2003] Luo, M., Ma, Y., and Zhang, H. (2003). A spatial constrained k-means approach to image segmentation. In *Pacific Rim Conference on Multimedia*, pages 738–742.
- [Meyer, 1994] Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing*, 38:113–125.
- [Noordam et al., 2000] Noordam, J., van den Broek, W., and Buydens, L. (2000). Geometrically guided fuzzy c-means clustering for multivariate image segmentation. In *IAPR International Conference on Pattern Recognition (ICPR)*, pages 462–465.
- [Pappas, 1992] Pappas, T. (1992). An adaptive clustering algorithm for image segmentation. *IEEE Transactions on Signal Processing*, 40:901–914.
- [Pham, 2001] Pham, T. (2001). Image segmentation using probabilistic fuzzy c-means clustering. In *IEEE International Conference on Image Processing (ICIP)*, pages 722–725.
- [Philipp-Foliguet, 2000] Philipp-Foliguet, S. (2000). Segmentation d'images en régions floues. In *Rencontres Francophones sur la Logique Floue et ses Applications*, pages 186–196, La Rochelle, France.
- [Prêteux, 1992] Prêteux, F. (1992). Watershed and skeleton by influence zones: a distance-based approach. *Journal of Mathematical Imaging and Vision*, 1:239–255.
- [Small, 1990] Small, C. (1990). A survey of multidimensional medians. *International Statistical Review / Revue Internationale de Statistique*, 58(3):263–277.
- [Soille, 2003] Soille, P. (2003). *Morphological Image Analysis : Principles and Applications*. Springer-Verlag, Berlin.
- [Tolias and Panas, 1998] Tolias, Y. and Panas, S. (1998). Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 28(3):359–369.
- [Turi, 2001] Turi, R. (2001). *Clustering-Based Colour Image Segmentation*. PhD dissertation, Monash University. 373 pages.
- [Xia et al., 2007] Xia, Y., Feng, D., Wang, T., Zhao, R., and Zhang, Y. (2007). Image segmentation by clustering of spatial patterns. *Pattern Recognition Letters*, 28:1548–1555.

Morphological Description of Color Images for Content-Based Image Retrieval

Erchan Aptoula and Sébastien Lefèvre

Abstract—Placed within the context of content-based image retrieval, we study in this paper the potential of morphological operators as far as color description is concerned, a booming field to which the morphological framework, however, has only recently started to be applied. More precisely, we present three morphology-based approaches, one making use of granulometries independently computed for each subquantized color and two employing the principle of multiresolution histograms for describing color, using respectively morphological levelings and watersheds. These new morphological color descriptors are subsequently compared against known alternatives in a series of experiments, the results of which assert the practical interest of the proposed methods.

Index Terms—Color description, content-based image retrieval (CBIR), granulometry, mathematical morphology (MM), multiresolution histogram.

I. INTRODUCTION

THE first foundations of mathematical morphology (MM) were laid down during the sixties, within the context of stereology related projects aiming to describe iron ore properties [1] and porous media [2]. Hence, visual content description can be considered as one of the initial goals, that guided the research efforts leading eventually to the development of the entire morphological framework as it is known today [3].

Since that first period the potential of MM for feature extraction and content description has been demonstrated in various fields, ranging from biomedical applications [4] to remote sensing [5], [6]. The topic of general purpose content-based image retrieval (CBIR), however, with the exception of highly application specific cases (e.g., hematological cell classification [7]), has been left relatively unexplored. At the end of the inception years of CBIR, this has started to change. Rather direct applications of classical morphological operators, such as pattern spectra [8], multiscale segmentations [9], combinations with already known descriptors (Gabor filters and watershed based segmentation [10]) as well as novel morphological approaches [11] are becoming more frequent. Nevertheless, to the best of the authors' knowledge no extensive study has yet been carried out.

Manuscript received September 25, 2008; revised June 14, 2009. First published July 10, 2009; current version published October 16, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Laurent Younes.

The authors are with the LSIIT, UMR 7005 CNRS/University Louis Pasteur—Strasbourg, Pôle API, Bd Brant, BP 10413, 67412 Illkirch cedex, France (e-mail: aptoula@lsiit.u-strasbg.fr; lefevre@lsiit.u-strasbg.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2009.2027363

Besides the relative popularity of CBIR in the image processing community, the motivation for considering morphological operators in order to tackle this problem, is mainly due to the inherent capacity of this framework to exploit spatial pixel relationships. Based on this concept, in this paper we focus on color description, and study the potential of morphological operators for content description within this context. Why color? Because, although it is certainly not the most important visual quality of image data when it comes to semantic retrieval, it still is one of the basic low level visual features, and as such it constitutes a good launching point for future work consisting of the development of more sophisticated morphological content descriptors, e.g., morphological color interest point detectors or other multiscale morphological approaches. More precisely, our contribution consists of three novel morphology based color descriptors, based on granulometries and multiresolution histograms in combination with color levelings and the watershed transform. All three proposed approaches are compared against known alternatives in a series of tests, that assert the practical interest of the morphological framework in this field.

We first provide in Section II a short overview of the application of MM to color images and its related issues as it is far from being straightforward. Then we focus on the development of morphology based color descriptors making use of granulometries and multiresolution histograms (Section III). Next we present the results of a series of comparative tests illustrating the interest of the proposed methods (Section IV), followed by concluding remarks (Section V).

II. COLOR MORPHOLOGY

The extension of morphological operators to color and more generally to multivariate images is an open problem. Specifically, since the morphological framework is based on complete lattice theory [12], it is theoretically possible to define morphological operators on any type of image data, as long as a complete lattice structure can be introduced on the image intensity range. In other words, given the vector nature of color pixels, a vector ordering becomes necessary. Several approaches have been proposed with this purpose (e.g., marginal orderings, reduced orderings, conditional orderings, etc.), a comprehensive survey of which can be found in [13]. Prior to choosing a vector ordering, however, it is imperative to select a color space to work with.

A. Color Space

Here, we choose to follow the trend of the last years in the domain of color morphology, and employ a polar color space based on the notions of hue ($h \in [0, 2\pi]$), saturation ($s \in [0, 1]$) and luminance ($l \in [0, 1]$). More precisely, although most

polar color spaces are essentially a more intuitive description of the RGB color cube, several implementations exist e.g., HSV, HLS, HSI, etc. According to Hanbury and Serra [14], [15], the cylindrical versions of these spaces have serious inconsistencies and are inappropriate for quantitative color image processing. Hence, we make our color space choice in favor of the *improved HLS* space, which employs the original biconic version of HLS. This color space has been introduced in [16], [17] under the name LSH, and can be computed from the RGB space as

$$\begin{aligned} L_{\text{LSH}} &= \frac{1}{3}(\max + \text{med} + \min) \\ S_{\text{LSH}} &= \begin{cases} \frac{3}{2}(\max - L_{\text{LSH}}) & \text{if } L_{\text{LSH}} \geq \text{med} \\ \frac{3}{2}(L_{\text{LSH}} - \min) & \text{if } L_{\text{LSH}} \leq \text{med} \end{cases} \\ H_{\text{LSH}} &= k \left[\lambda + \frac{1}{2} - (-1)^\lambda \left(\frac{\max + \min - 2\text{med}}{2S_{\text{LSH}}} \right) \right] \end{aligned} \quad (1)$$

where $\max$, med , and $\min$ denote respectively the maximum, median and minimum value of the transformed RGB triplet. k is the angle unit ($\pi/3$ for radians and 42 for 256 levels), whereas λ is set as 0 if $r > g \geq b$, 1 if $g \geq r > b$, 2 if $g > b \geq r$, 3 if $b \geq g > r$, 4 if $b > r \geq g$, and 5 if $r \geq b > g$.

One of the most important drawbacks of the cylindrical HLS space is the unintuitive definition of saturation. Specifically, it is possible to have maximized saturation values for zero luminance. This inconvenience, as well as the dependence of saturation on luminance are remedied with the LSH space, where the maximal allowed value for saturation is limited in relation to luminance. Therefore, in order to benefit from the advantages of polar spaces in the context of multivariate morphology, the ordering of LSH color vectors is required.

B. Color Ordering

As to the issue of color ordering, it is nontrivial since it is tied both to the problem of multivariate ordering as well as to the subjective notion of color. In order to realize our choice of color ordering among the available rich variety, we focused on two criteria, first its theoretical stability and second its suitability to the color space under consideration. For these reasons, we have chosen to use primarily the lexicographical ordering (i.e., a total ordering leading to theoretically valid morphological operators) and more specifically its quantization based variant [18], which renders it not only more flexible but takes into account the particular relations among the dimensions of LSH as well (e.g., the redundancy of hue when a pixel is not saturated “enough”).

More precisely, the quantization based lexicographical ordering is based on the standard lexicographical cascade, which is defined by

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbb{R}^n, \mathbf{v} <_L \mathbf{v}' \Leftrightarrow \exists i \in \{1, \dots, n\}, (\forall j < i, v_j = v'_j) \wedge (v_i < v'_i) \quad (2)$$

It is applied on the color vectors of LSH, where the importance of saturation is maximized for medium levels of luminance, whereas hue is of practical importance only for relatively high levels of saturation. Thus, we conform here to previous works [13] and use $L \rightarrow S \rightarrow H$ in the lexicographical cascade, where the first two components are subquantized nonuniformly

according to functions modeling the interchannel relations of the color space

$$\begin{aligned} \forall \mathbf{v}_1 = (l_1, s_1, h_1), \mathbf{v}_2 = (l_2, s_2, h_2) \in [0, 255]^3 \\ \mathbf{v}_1 < \mathbf{v}_2 \Leftrightarrow [l'_1, s'_1, h_1]^T <_L [l'_2, s'_2, h_2]^T \end{aligned} \quad (3)$$

In particular, we have used the exact configuration presented in [18] concerning luminance (l) and saturation (s) to produce quantized luminance (l') and saturation (s') (note that any eventual equivalences among distinct color vectors may be avoided by using the lexicographical cascade once more with the original pixel values). As to the hue, however, since it is a 2π -periodical angular value, it has been decided to use a reference based hue ordering [19]

$$\forall h, h' \in [0, 2\pi], h \leq h' \Leftrightarrow h' \div h_0 \leq h \div h_0 \quad (4)$$

where hues are ordered with respect to their distance to a reference value h_0 (to be defined depending on the application), with the closer ones being considered greater. As to the distance measure, it is defined as [19]

$$\begin{aligned} \forall h, h_0 \in [0, 2\pi] \\ h \div h_0 = \begin{cases} |h - h_0| & \text{if } |h - h_0| < \pi \\ 2\pi - |h - h_0| & \text{if } |h - h_0| \geq \pi \end{cases} \end{aligned} \quad (5)$$

Therefore, equipped with the ordering of (3) one can compute the vector extrema by means of $\inf_v$ and $\sup_v$, necessary for the definition of the basic morphological operators of vector erosion (ϵ_b) and dilation (δ_b)

$$\epsilon_b(\mathbf{f})(\mathbf{x}) = \inf_{\mathbf{s} \in b} \{\mathbf{f}(\mathbf{x} + \mathbf{s})\} \quad (6)$$

$$\delta_b(\mathbf{f})(\mathbf{x}) = \sup_{\mathbf{s} \in b} \{\mathbf{f}(\mathbf{x} - \mathbf{s})\} \quad (7)$$

from the combinations of which more sophisticated operators may be obtained. Now let us proceed to the developments of morphological operators for color images aiming to describe color content.

III. MORPHOLOGICAL DESCRIPTION OF COLOR IMAGES

Color is widely regarded as one of the most expressive visual features [20], and as such it has been extensively studied in the context of CBIR, thus leading to a rich variety of descriptors [21]. Similarly to texture and shape, it is relatively hard to establish the exact description of color content. One has to ask, given a color image, which notions describe the concept of color within it. Considering the existing amount of work on this topic, and the multitude of color descriptors employed in commercial and research systems, one can narrow down the type of information extracted from them to the following.

- The presence (or absence) of a particular color (e.g., red and blue).
- The relative amount in which a particular color is present (e.g., 80% red and 20% blue).
- The size distribution of each color (e.g., one large patch of red and multiple small patches of blue).
- The relative spatial position of each color with respect to others (e.g., red at left and blue at right).

The last of the aforementioned four notions is of use only in cases where the relative position of color components is of semantic importance. Such an assumption makes possible to use (multiscale) segmentation-based approaches where the image is represented as a set of regions through a graph (or tree) structure. In this very particular context, mathematical morphology is a possible solution [9]. An example of images where this notion is relevant is the following: in the case of beaches, the sand (brown and/or gray) is next to the sea (blue and green), above which is the sky (blue and/or gray). Of course there are also numerous other objects for which relative positions cannot be *a priori* determined, for instance accordions, books, etc. Consequently, one can understand why the vast majority of existing work on this topic concentrates on the first three notions in order to describe color content.

Any attempt on color description, however, has to address certain issues first. These are the choice of color space and quantization levels, as well as the variations due to different illumination conditions. The choice of color space is of vital importance, as it has a direct effect on channelwise correlation, and consequently on the effectiveness of the resulting feature vectors. Although perceptual spaces (e.g., $L^*a^*b^*$, $L^*u^*v^*$) are considered to better model the human perception of color, comparative studies have shown phenomenal color spaces to outperform them in combination with several descriptors [22].

As the standard number of colors in everyday digital images is usually in the order of millions, color specific approaches naturally lead to excessively large feature vectors; hence, color subquantization (i.e., quantization to fewer levels) is a common practice. The exact number of levels along with their uniformity, however, depend strongly on the chosen color space. For instance while 4 colors per channel are preferred for RGB (i.e., 4-4-4, 64 colors), 16-4-4 is usually used with HSV [20], [23]. The question of choosing the optimal quantization model and number of levels can be considered as a search for the optimal compromise between feature vector size and effectiveness.

Another significant obstacle in color description is the one caused by color variations due to different illumination conditions. Even with exactly the same object and viewing angle, changing only the illumination source may lead to totally different color distributions. One can counter this effect, either by means of an appropriate transformation, like histogram equalization (Fig. 1) [24] or by subquantizing strongly the luminance component in order to reduce its influence [25].

Despite the absence of previous work on the morphological characterization of color distributions, we believe that morphological options possess the potential to contribute to all three stages of it, specifically preprocessing (e.g., image simplification), processing (e.g., feature extraction) and postprocessing (e.g., descriptor refinement). In the light of the aforementioned color description principles, we present three global/holistic approaches for the description of color based on morphological operators, that are subsequently compared against known alternatives.

A. Color Specific Granulometry (CSG)

Since the first two requirements for describing color content can be trivially obtained through an histogram, the main

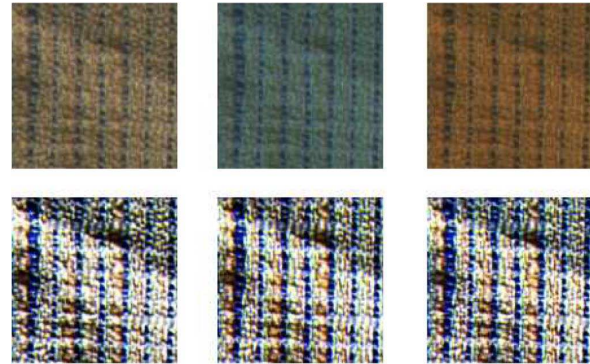


Fig. 1. In the first row, a texture example from the Outex14 texture database under three different illumination sources (left to right, 2856 K incandescent CIE A, 2300 K horizon sunlight and 4000 K fluorescent TL84) and in the second row, the same images after the application of a channelwise histogram equalization in RGB.

problem to address is finding a means to describe the spatial/size distribution of colors. To this end, an important part of the existing methods follows a color specific approach, where each color is described independently, either by means of an autocorrelation or some other spatial histogram extension, each color bin is associated additionally with a spatial description. That is why it has been decided here not to employ a vector processing strategy, and instead focus on each quantized color separately using granulometry as our description tool. Besides, CBIR oriented color granulometries have already been reported, such as the one by Angulo and Serra [26], where they are computed marginally in the RGB space and the one by Louverdis *et al.* [27], defined in HSV using a lexicographical ordering, which, however, ignores the hue component's periodicity.

Intuitively, we choose to associate each color of the input with its corresponding granulometric curve, since it is considered as the morphological tool by default for the description of size distributions [28]. In particular, this operator consists in studying the amount of image detail removed by applying morphological openings γ_λ and/or closings ϕ_λ of increasing size λ . The volumes (i.e., sum of graylevel pixel values) of the opened (or closed) images are then plotted against λ , or more usually their discrete derivative $\text{Vol}(\gamma_\lambda - \gamma_{\lambda+1})$, i.e., *pattern spectrum*.

In the present case, given a color image $\mathbf{f}$ with possible colors in $\{c_i\}_{1 \leq i \leq m}$, a graylevel image f_{c_i} for each c_i is computed, where every pixel $f_{c_i}(x, y)$ denotes the distance of $\mathbf{f}(x, y)$ to c_i

$$\forall(x, y), f_{c_i}(x, y) = d(\mathbf{f}(x, y), c_i) \quad (8)$$

where d represents a color metric suitable for the color space under consideration, e.g., the Euclidean distance for perceptually uniform spaces. Next we compute the granulometric curve of f_{c_i} using closings by reconstruction. To explain, closing by reconstruction is a high level morphological operator which can be expressed in the digital case in terms of geodesic erosions. Geodesic erosion $\varepsilon_g^{(1)}(f)$ and dually geodesic dilation $\delta_g^{(1)}(f)$ are operations involving two images, a marker image f and a mask g , where the first is processed conditionally to the second

$$\varepsilon_g^{(1)}(f) = \varepsilon(f) \vee g \quad (9)$$

$$\delta_g^{(1)}(f) = \delta(f) \wedge g \quad (10)$$

where ε and δ are the erosion and dilation by the neighborhood of the origin. Both operators can be further applied successively as follows:

$$\psi_g^{(n+1)}(f) = \psi_g^{(1)}\left(\psi_g^{(n)}(f)\right) \quad (11)$$

hence, by repeating them until stability (i times, $\psi_g^{(i)}(f) = \psi_g^{(i+1)}(f)$) one can realize respectively a reconstruction by erosion and a reconstruction by dilation

$$R_g^\varepsilon(f) = \varepsilon_g^{(\infty)}(f) \quad (12)$$

$$R_g^\delta(f) = \delta_g^{(\infty)}(f) \quad (13)$$

which reach stability after a finite number of steps. Consequently, one can define closing by reconstruction with a structuring element (SE) B

$$\phi_{R,B}(f) = R_g^\varepsilon[\delta_B(f)] \quad (14)$$

which we use in combination with SE of various sizes, in order to compute the granulometric curve of f_{c_i} that describes color c_i

$$h_f^k(c_i) = \text{Vol}(\phi_{R,B_k}(f_{c_i}))/\text{Vol}(f_{c_i}) \quad (15)$$

$$\forall c_i, h_f(c_i) = \{h_f^k(c_i) \mid k \in \{1, \dots, n\}\} \quad (16)$$

where B_k is the SE of size k to be used in the closing by reconstruction, and Vol denotes the image volume (i.e., sum of graylevel pixel values of each f_{c_i}). Consequently, the final feature vector is formed by the concatenation of all $h_f(c_i)$, leading to a descriptor of size $n \times m$. The reason for using a closing instead of an opening is because the regions representing colors similar to the one under consideration will be at minimal distance, hence, dark, whereas the choice of a reconstruction based operator is justified by the reports concerning the improvement of performance that it brings in the context of content description [29].

Moreover, granulometries, similarly to correlograms, are invariant with respect to spatial position [30], and provide a description of the size distribution of their content. Applying them directly on the input image results in a fairly effective shape descriptor [8]. By applying them “marginally” for each possible color, one can distinguish for instance a connected “large” patch of color c_i , from a number of smaller components of the same color. As to its similarity comparison, we find it pertinent to use an approach similar to the relative L_1 distance based measure of correlograms [31]

$$d_{\text{CSG}}(\mathbf{f}, \mathbf{f}') = \sum_{1 \leq i \leq m, 1 \leq k \leq n} \frac{|h_f^k(c_i) - h_{f'}^k(c_i)|}{1 + h_f^k(c_i) + h_{f'}^k(c_i)} \quad (17)$$

An illustrative example of CSG is given in Fig. 2, where the curves obtained for three colors of different size distribution are shown. Possible extensions to the definition of (16) could include using multivariate granulometries (e.g., size-shape) [32], as well as higher order statistical measures instead of the volume [33].

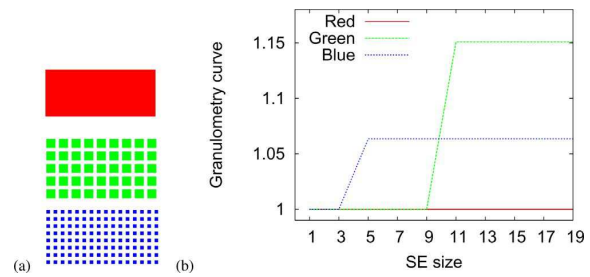


Fig. 2. Input image (a) and the granulometric curves (b) corresponding to red, green and blue, obtained by processing the distance image of (a) computed using an Euclidean distance in RGB.

B. Multiresolution Histograms in the Leveling Scale-Space (MHL)

Color histograms remain simple yet effective tools since the early days of CBIR [34], with positive properties such as invariance to geometric transformations. As they are capable of providing the first two elements required for the description of color, their foremost drawback is the lack of any spatial information. To this end, several extensions have been proposed [35], [21]; among which one particular approach is using multiple resolutions, that has proven itself to be able to capture effectively color content [36], [37]. Multiresolution decomposition has been achieved by means of Gaussian filters [36], Gabor filters [38] as well as wavelets [39]. For an in-depth study on this topic the reader can consult [36]. Morphological scale spaces, however, have not yet been used to this end, despite being used for other multiscale image representations (e.g., related to segmentations [9]).

Scale-spaces based on morphological levelings have been studied in particular by Meyer and Maragos [40], where the useful properties of levelings, such as contour preservation and invariance to translation and rotation among others have been mentioned. This formulation later on has been used for optimal scale selection by exploiting the property of maxima propagation within levelings [41], and for multiscale image description through segmentation tree [9]. Here, however, we propose to employ this scale space for the description of color distribution by means of multiresolution histograms. Before proceeding any further, let us recall the relative definitions. Levelings are powerful morphological operators representing a subclass of connected operators, that work on a reference image $\mathbf{f}$ and a marker image $\mathbf{g}$. The marker image is modified in such a way that it becomes the leveling $\mathbf{g}' = \lambda(\mathbf{f}, \mathbf{g})$ of $\mathbf{f}$. From an implementational point of view, the expression

$$\mathbf{g}' = [\mathbf{f} \wedge \delta(\mathbf{g})] \vee \varepsilon(\mathbf{g}) \quad (18)$$

is iterated until idempotence, although more efficient algorithms also exist [42].

At this point two issues need to be addressed, first the extension of levelings to color images, as well as the formulation of a scale-space. The first has been studied by Gomila and Meyer [43] and by Angulo and Serra in [44], where different orderings and color spaces have been tested. It consists in replacing the scalar morphological as well as binary operators of (18) by their

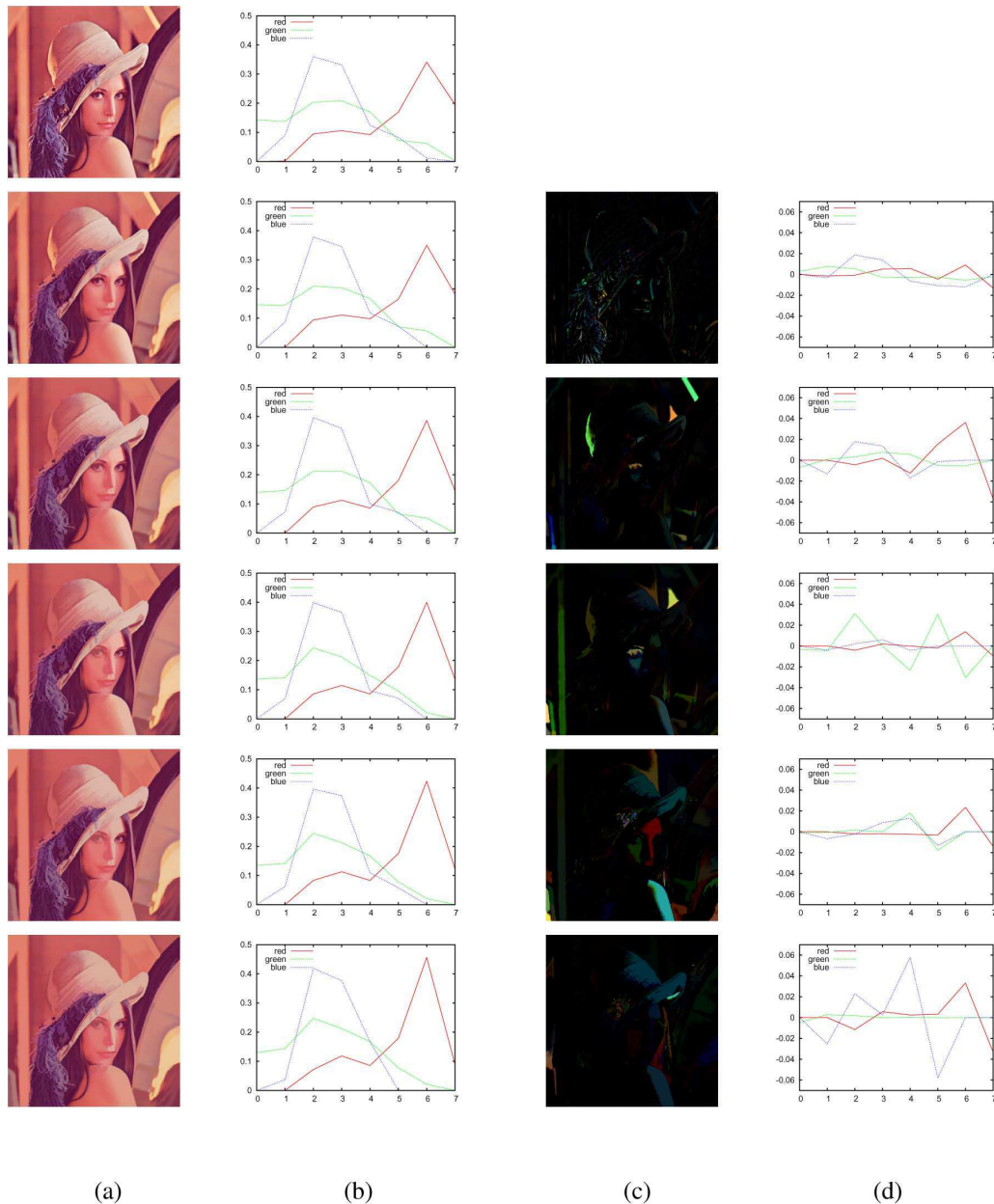


Fig. 3. From top to bottom: Original Lenna image and its vectorly leveled versions using (18) and (19), with marker images obtained using alternating sequential filters (ASF) of increasing size 3, 5, 7, 9, and 11. Each line contains (from left to right): the (leveled) color image (a) and its corresponding 8-bin RGB color normalized histograms (b), the differences with the finer image (i.e., from previous line or resolution) emphasized with contrast stretch (c) and considering its color normalized histograms (d). (This figure is best viewed in color).

vector versions. As to the second, multiple methods are possible such as using a family of extensive and anti-extensive operators within (18) or a family of marker images $\{g_i\}$ that lead to progressively simplified images

$$\text{For resolution } i \in \{1, \dots, n\} \\ g'_0 = f \quad \text{and} \quad g'_i = \lambda(f, g_i). \quad (19)$$

Of course, a variety of filters may be used in order to obtain the marker images, including alternating sequential filters (ASF),

Gaussian or even anisotropic diffusion filters, that according to the study in [45] provide the best quantitative performance in terms of noise reduction capacity. An example of successive color levelings applied on the Lenna image (256×256 pixels) is given in Fig. 3(a), where the multiple marker images are obtained through vectorial ASF (with opening first) of various sizes i expressed as

$$\text{ASF}_i(f) = \phi_i \gamma_i \dots \phi_2 \gamma_2 \phi_1 \gamma_1(f) \quad (20)$$

where ϕ_i and γ_i respectively represent the closing and opening of size i . Using a smoother marker leads to larger flat regions and results in a removal of more details as shown in Fig. 3(c), where contrast has been stretched for visualization purpose. Moreover, the leveling process modifies the corresponding color histograms as illustrated in Fig. 3(b), which contains normalized histograms. Differences between color histograms at successive resolutions (i.e., when computed from successive leveled images) are emphasized in Fig. 3(d). These differences can reach up to 4000 pixels for a given quantized color and correspond to the large areas shown in Fig. 3(c).

Given n resolutions, one can then compute the color histogram of size m of each g_i^j and concatenate them in order to form the feature vector of size $n \times m$. Moreover, the histogram differences may also be used. As to the similarity measure, while the histogram intersection is of course the standard choice, a weighting factor may also be added, following the principle of a *pyramid match* [46]. In detail, since highly simplified images are bound to be more alike than the lower resolutions where more image detail is preserved, it makes sense to weight the corresponding histogram intersections. For instance, given $H(\mathbf{f}) = \{H_i(\mathbf{f})\}_{0 \leq i \leq n}$ and $H(\mathbf{h}) = \{H_i(\mathbf{h})\}_{0 \leq i \leq n}$, the multiresolution histogram families for two images $\mathbf{f}$ and $\mathbf{h}$, respectively, their weighted distance d_w would become

$$d_w(H(\mathbf{f}), H(\mathbf{h})) = \sum_{i=0}^n \frac{1}{2^i} \sum_{j=1}^m \min\{H_{ij}(\mathbf{f}), H_{ij}(\mathbf{h})\} \quad (21)$$

where H_{ij} denotes the j th histogram bin of resolution i . Further refinement may be achieved by modifying the bin width at each resolution.

C. Multiresolution Histograms Based on Watersheds (MHW)

Another morphological tool that may be used in order to boost the spatial sensitivity of the standard color histogram is the watershed transform. This powerful operator represents the foremost morphological approach to the problem of image segmentation. Although fully automatic, it is relatively sensitive to image noise, hence leading often to oversegmented results unless some form of smoothing is applied first. Nevertheless, the regions that are formed represent uniform areas in terms of spectral content or “flat” in topological terms. Here we propose to use this operator in order to produce the different image resolutions that are to be used in order to construct a multiresolution histogram. To do so we rely on multiple smoothed images, e.g., the leveled images presented in previous section. Let us notice, however, that it would have also been possible to involve multiscale watershed-based techniques (e.g., waterfalls [47], using contour saliencies [48], etc.) which produce directly segmentations at multiple resolutions, thus avoiding the need for multiple smoothed images.

In particular, given a color image $\mathbf{f}$, it is first smoothed relatively to resolution i and then segmented by means of the watershed transform, which leads usually to several regions. Then, the pixels of each region are associated with the representative

color of that region. Several techniques may be involved to compute this representative color, among which we can cite: the average color (using a standard average function computed marginally on each component of the color space under consideration); the most dominant color (seeking for the main peak of the color histogram); the most homogeneous color (considering the color of the valley in the topographic surface, i.e., the pixel with minimum gradient value). Consequently, one can at this point compute the standard histogram taking into account the newly formed regions. Besides, using multiple resolutions means that the initial image will be further smoothed, hence leading to the formation of even larger flat regions, and, thus, we obtain a multiresolution representation of the input the histograms of which capture indirectly the spatial organization of its content. This procedure is further illustrated in Fig. 5. Consequently, the end feature vector is formed by the concatenation of the histograms of each resolution, thus leading to a size of $n \times m$, where n is the number of resolutions and m the number of bins of each histogram. As a similarity measure, we consider the pyramid match of (21) as adequate.

Although it appears as a straightforward approach, its implementation is hindered, however, by certain elements. First, there is the issue of a morphological color image segmentation [17]. The watershed transform is most often applied on a graylevel input representing the topographic relief of the image under consideration, usually its gradient, that is why an effective means of computing the color variations is necessary. To this end, we choose to combine the color channels by means of a channel-wise maximum of marginal gradients

$$\forall (l, s, h) \in [0, 1]^3 \rho_{\text{LSH}}(l, s, h) = \max\{\rho(l), \rho(s), \rho_H(h, s)\} \quad (22)$$

where $\rho = f - \varepsilon(f)$ is the standard internal morphological gradient, chosen instead of its alternatives based on experimental observations. Although the components of the polar color spaces are highly intuitive, their combination is relatively problematic. In particular, hue is of no importance if saturation is “low,” while the biconic shape of the color space assures that no high saturation levels exist, if luminance is not “high enough.” Hence, the hue gradient needs to be weighted with a coefficient that has a strong output only when both compared saturation values are “sufficiently high”. Besides, on the contrary of the other two components, it has been observed that the max – min form provides visually superior results with the hue, hence leading us to use the following expression:

$$\rho_H(h, s) = \max_{i \in B} \{j(s, s_i) \times h \div h_i\} - \min_{i \in B} \{j(s, s_i) \times h \div h_i\} \quad (23)$$

where B is the local 8-neighborhood, $\div$ the hue distance defined in (5) and $j(\cdot, \cdot)$ a double sigmoid controlling the transition from “low” to “high” saturation levels

$$j(s_1, s_2) = \frac{1}{(1 + \exp(\alpha \times (s_1 - \beta))) \times (1 + \exp(\alpha \times (s_2 - \beta)))} \quad (24)$$

where the slope $\alpha = -10$ and the offset $\beta = 0.5$.

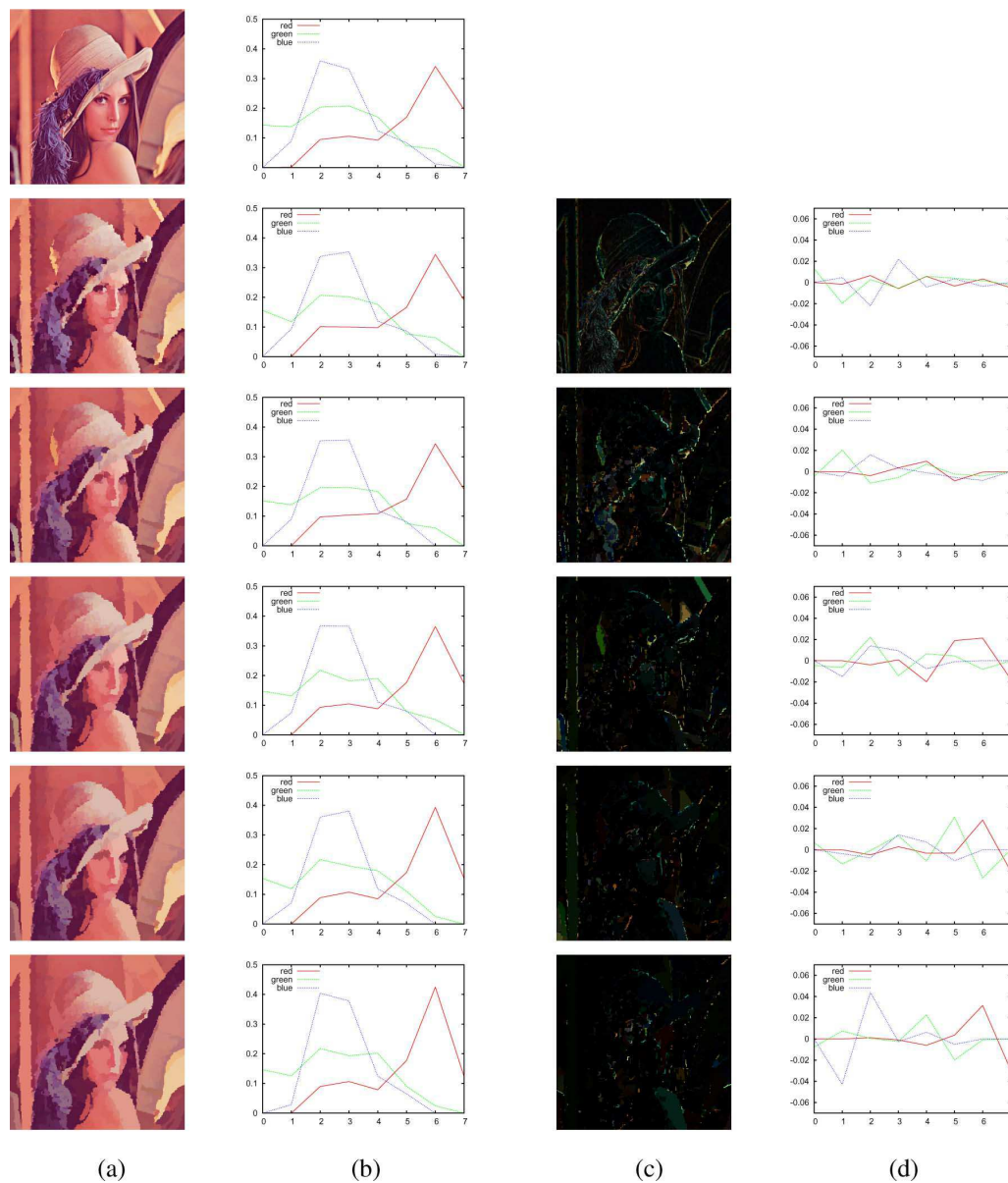


Fig. 4. From top to bottom: Original Lenna image and the results obtained after the watershed computation on the levelings of Fig. 3(a), where pixel values have been replaced by the valley color of each basin. Each line contains (from left to right): the resulting color image (a) and its corresponding 8-bin RGB color normalized histograms (b), the differences of each image (i.e., from previous line or resolution) emphasized with contrast stretch (c) and considering its color normalized histograms (d). (This figure is best viewed in color).

Moreover, one also has to choose a means to compute the progressively smoothed versions of the input image. Here, we consider color levelings with this purpose, the marker images of which may be obtained with any of the options mentioned in Section III-B. Following the flattening realized by levelings, the flat regions will be easily detected by the watershed transform, that will also detect as regions all basin shaped areas that have not been flattened by levelings. Whether this nuance has a positive or negative effect on the operator's color content description capacity is a question to be answered experimentally in Section IV but we still provide here a comparison on the standard

Lenna image. The application of the watershed transform on the computed gradient leads to the result shown in Fig. 4(a), whereas Fig. 4(b) shows their color normalized histograms. Similarly to the leveled images, interresolution differences are present though hard to observe. Thus, we also provide contrast-stretched visual and normalized histogram differences in Fig. 4(c) and (d), respectively. Here we consider the valley color as the representative color of each region, but let us notice that some other tests with average color provide very similar results.

To illustrate the difference between the two multiscale approaches (using levelings only or levelings followed by wa-

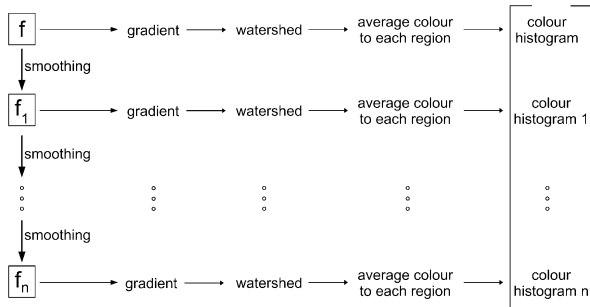


Fig. 5. Illustration of the steps required for forming the watershed transform based multiscale histograms.

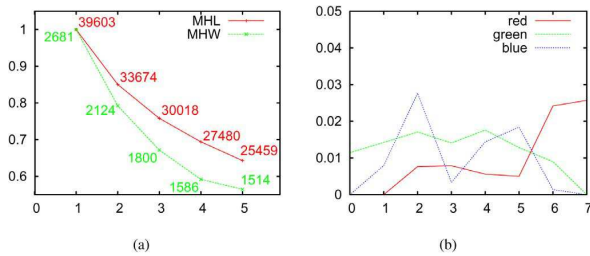


Fig. 6. Difference between multiscale histograms using levelings only (MHL) and levelings followed by watersheds (MHW): (a) evolution of the color number at different image resolutions relative to the first resolution, (b) average difference (among the different resolutions) between the three 8-bin RGB color normalized histograms obtained from MHL and MHW.

tersheds), we measure in Fig. 6(a) the color number at each image resolution (the initial Lenna image contains 52319 different colors, i.e., RGB triplets). We can observe that the watershed segmentation greatly reduces (by a factor of 15) the color number. However, since a quantization is usually applied on the input image as indicated at the beginning of this section, the difference appears not on the color number but rather on the quantized color distributions, as illustrated in Fig. 6(b). As we can notice in left column of Fig. 7, the watershed-based additional step introduces changes in particular on the object borders. Moreover, these changes appear more on a region-basis than on a pixel-basis, thus asserting the effect of the segmentation step. These interresolution differences lead to different evolutions of color distributions, see right column of Fig. 7.

IV. EXPERIMENTAL EVALUATION

A. Setup

The retrieval tests that have been carried out have served multiple purposes. First they have been used in order to verify the practical interest of the presented morphological color descriptors, as well as to determine the optimal color space and quantization couple for this task. To this end, a total of 3000 images from three different sources have been used [49]–[51]. This set contained 16 semantic categories of 100 images each, along with 1400 bulk images, in order to simulate real-life retrieval from an unorganized database. Examples from each category are given in Fig. 8. Furthermore, in order to obtain more reliable estimates, each image of the 16 categories served as the query subject in

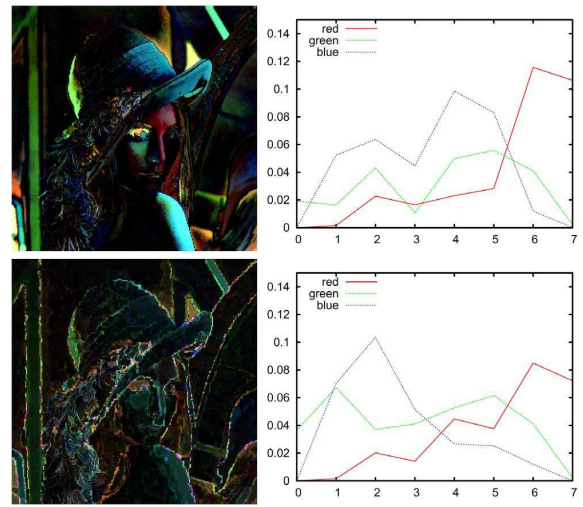


Fig. 7. Amount of differences accumulated over the five different resolutions using levelings only (top) and levelings followed by watersheds (bottom). The visual differences (left) lead to differences in the color distributions (right).

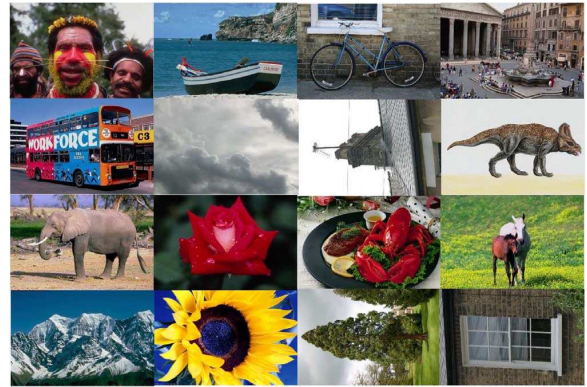


Fig. 8. Representative category examples of the image database used for color descriptor comparisons. From top-left to bottom-right: Africa, beach, bicycles, building, bus, clouds, chimneys, dinosaur, elephant, flower, food, horse, mountain, sunflower, tree, and window.

turn, making a total of 1600 queries. From the results of which the precision versus recall charts have been computed, as well as ANMRR values that provides an overall rating of performance.

ANMRR stands for *average normalized modified retrieval rank* and is the MPEG-7 retrieval effectiveness measure commonly accepted in the CBIR community [20]. More precisely, this measure has been introduced in order to overcome problems such as those related to queries with varying ground truth set sizes. In particular, given a query q , assume that the k th ground truth image is retrieved at $\text{Rank}(k)$. Then a number $K(q)$ is defined, which denotes the ranks that are considered as feasible in terms of retrieval evaluation. $K(q)$ is often set as twice the size of the ground truth set $NG(q)$ associated with the query under consideration. Hence, $\text{Rank}^*(k)$ is defined, which represents the penalty attributed to a retrieved item

$$\text{Rank}^*(k) = \begin{cases} \text{Rank}(k), & \text{if } \text{Rank}(k) \leq K(q) \\ 1.25K(q), & \text{if } \text{Rank}(k) > K(q) \end{cases} \quad (25)$$

Thus, one gets the average rank (AVR) for a query q

$$\text{AVR}(q) = \frac{1}{\text{NG}(q)} \sum_{k=1}^{\text{NG}(q)} \text{Rank}^*(k). \quad (26)$$

Skipping the intermediate notions, we can directly define the ANMRR as

$$\text{ANMRR} = \frac{1}{\text{NQ}} \sum_{q=1}^{\text{NQ}} \frac{\text{AVR}(q) - 0.5(1 + \text{NG}(q))}{1.25K(q) - 0.5(1 + \text{NG}(q))} \quad (27)$$

where NQ is the number of queries. Thus, ANMRR can provide values between 0 for best results (i.e., all images of the ground truth of the queries have been retrieved), and 1 for worst results (i.e., none of the ground truth images of the queries has been retrieved).

It should be additionally noted that for the charts the first retrieved image was always ignored as it represents the query image itself. In our tests, the proposed morphological color descriptors have been compared against the standard histogram, the autocorrelogram (AC) [31], the superiority of which with respect to the histogram has been shown by multiple comparative studies [22], [25], the color structure descriptor (CSD) [20] of the MPEG-7 standard and color distribution entropy (CDE) [21].

Moreover, a question that arises at this point is “why use a nonstandard image database?”. Indeed, our initial intent was in the opposite direction; however, it became soon clear that contemporary image collections, such as Caltech-256 [50] are rather unsuitable for this particular testing suite, in the sense that they require descriptors capable of taking into account multiple visual dimensions (e.g., color, shape, etc.) simultaneously, thus rendering them inappropriate for the present case where we wish to focus on color, and color only. Of course this also represents a limit for this series of tests, as the results rest valid for collections with a certain level of color separability among categories.

B. Test Results

1) *Color Space and Quantization:* According to previous comparative studies [22] and the MPEG-7 standard [20], polar color spaces in general outperform most of their alternatives in the context of content description, in combination with different quantization models. Considering, however, the design problems of these spaces (Section II-A), and the alternative formulations that have been proposed, it has been considered appropriate to test their effectiveness in order to determine the most suitable color space along with its quantization model for the image database under consideration.

To this end, we have employed color histograms along with RGB, $L^*a^*b^*$, HSV, and LSH with three different uniform quantization models, namely 8 (8-8-8), 6 (6-6-6), and 4 (4-4-4) bins per channel. The resulting ANMRR values are given in Table I. Judging from the obtained results, there are primarily two remarks to be made: first, the general superiority of LSH with respect to its counterparts, a result which confirms previous studies [25]. And additionally, one can observe the relatively light influence of histogram bin numbers. As a matter of fact, even with eight times less colors, the performance of

TABLE I
ANMRR VALUES FOR DIFFERENT QUANTIZATION MODELS (HISTOGRAM BINS) AND COLOR SPACES

Color spaces	8-8-8 (512)	6-6-6 (216)	4-4-4 (64)
RGB	0.4993	0.5042	0.5478
$L^*a^*b^*$	0.5544	0.5718	0.6127
HSV	0.4912	0.4998	0.5518
LSH	0.4759	0.4824	0.5158

TABLE II
ANMRR VALUES FOR DIFFERENT EXPRESSIONS OF BRIGHTNESS USING 256 BIN BRIGHTNESS HISTOGRAMS

Value	Intensity	Luminance
0.707	0.6843	0.7086

TABLE III
ANMRR VALUES FOR DIFFERENT EXPRESSIONS OF SATURATION USING 256 BIN SATURATION HISTOGRAMS

Luminance dependent	Independent
0.7053	0.6655

TABLE IV
ANMRR VALUES FOR SATURATION WEIGHTED AND UNWEIGHTED 256 BIN HUE HISTOGRAMS

Saturation weighted	unweighted
0.5892	0.6275

the histograms decreases only by 4% for the best color space under study and less than 6% for the others.

Based on the performance of the two polar color spaces, from this point on we focus on the optimization of their arguments. For instance, which brightness expression is most adequate for this task? Value (maximum of R, G, B), intensity (average of R, G, B), or luminance (perceptually weighted combination of R, G, and B)? Table II presents the results obtained using a 256 bin brightness histogram with the aforementioned expressions. According to the obtained results, the relatively simple expression of red green blue average, provides an overall better performance with this set.

Proceeding to the other two channels, as far as saturation is concerned, we test the brightness dependent version of HSV against the brightness independent definition of LSH, by means of a 256 bin saturation histogram. Table III shows the obtained ANMRR values. According to the results, brightness independence aids saturation's retrieval performance.

As to the hue component, we choose the implementation of LSH, and test the effect of using a saturation weighted hue histogram against a “raw” hue histogram, where the hue values of achromatic pixels take place as well. Table IV shows the obtained results. The weighting factor has indeed a positive effect.

In the light of these results, we opt for the LSH color space, equipped with the R, G, B average as brightness expression, and its hue weighted by its saturation. Furthermore, since

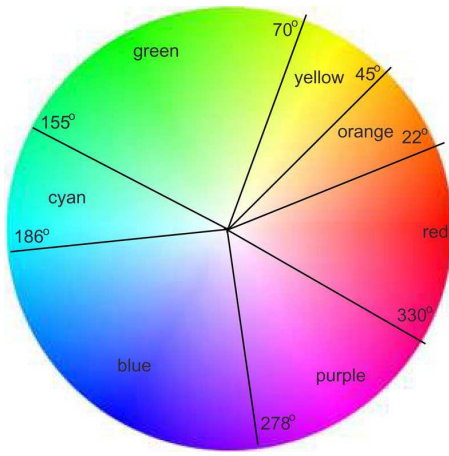


Fig. 9. Nonuniform hue circle quantization.

TABLE V
ANMRR VALUES FOR VARIOUS QUANTIZATION MODELS.
NU STANDS FOR NONUNIFORM

Quantization model	Color number	ANMRR
3-3-7 (NU)	63	0.4908
3-3-8	72	0.5045
4-4-8	128	0.4912
3-3-12	108	0.4813
3-3-16	144	0.4800
4-4-12	192	0.4729
4-4-16	256	0.4717

the hue histogram is evidently more pertinent with respect to brightness and saturation, similarly to the general tendency [20], [23], nonuniform quantization of LSH is tested next. Besides, allowing a lower number of levels for brightness, has the additional advantage of limiting the effect of illumination variations. But the question is, how exactly should one quantize this space? Common models include 4-4-16 [20], 3-3-12 [25], as well as 3-3-9 [21]. Moreover one has also the option of quantizing the hue circle nonuniformly, according to [52], which is based on the fact that human color perception has a varying capacity of hue distinction. For instance, the average person can distinguish more hues of red than blue. This principle is illustrated in Fig. 9, and the quantization option is tested as 3-3-7 (NU). Table V summarizes the results obtained from various quantization models. Although in general the exact configuration has relatively small effect on the retrieval performance, as a practical compromise between color number (i.e., feature size) and effectiveness, the nonuniform 3-3-7 stands out. Hence, it will be subsequently used as the default quantization scheme, unless stated otherwise.

2) *Descriptors*: Having chosen the LSH color space with a nonuniform subquantization of type 3-3-7 [52] (see Fig. 9), we now continue to compare the color descriptors presented in Section III against some of the classical methods of color content

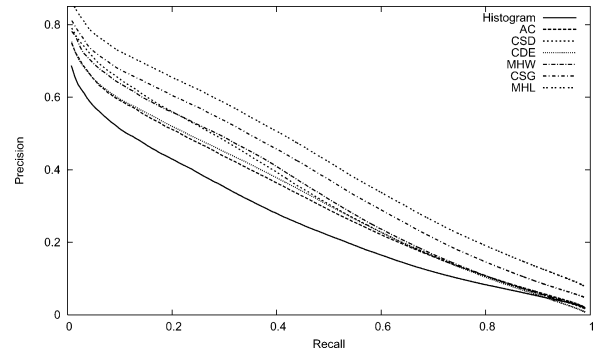


Fig. 10. Precision versus recall chart of the tested color descriptors, from worst to best descriptors: standard histogram, autocorrelation (AC), color structure descriptor (CSD), color distribution entropy (CDE), multiresolution histograms with watersheds (MHW), color specific granulometry (CSG), multiresolution histogram with levelings (MHL).

description. As far as CSG is concerned, the projection of each quantized color within the 63 color set is computed using the distance between the color under consideration and the original unquantized image. The distance expression is set to a weighted combination of hue and luminance difference

$$\begin{aligned} \forall c_1 = \{h_1, s_1, l_1\}, c_2 = \{h_2, s_2, l_2\} \\ d(c_1, c_2) = (h_1 \div h_2) \times j(s_1, s_2) \\ + |l_1 - l_2| \times (1 - j(s_1, s_2)) \end{aligned} \quad (28)$$

where all color components have been normalized to [0,1], the weight $j(\cdot, \cdot)$ is given in (24) and $\div$ represents the standard hue distance defined in (5). Moreover, the granulometries have been computed using square shaped SE with 4 sizes, 1, 5, 9, and 13 pixels wide using the efficient algorithm of [53]. Thus, with $3 \times 3 \times 7 = 63$ colors, it leads to a feature vector of length $4 \times 63 = 252$. As to MHL, based on empirical observations it has been decided to use ASF (with opening first) in order to realize the levelings, in 4 different resolutions, with the same SE options as CSG. Levelings of the same setup have been also used along with MHW, prior to applying the watersheds. Consequently, all three operators lead to identical feature vector sizes.

On the other hand, as to their rivals, the standard color histogram is computed in LSH using the subquantization model 4-4-16 with histogram size 256. The AC is calculated with the same setup as in [25], where it is defined for the HSV color space in combination with a subquantization model of 12-3-3. Namely, with four distances (1,3,5,7), hence leading to a feature vector of length $108 \times 4 = 432$. For the CSD, we use the configuration presented in [20] but with the aforementioned color space related choices instead of the HMMD space. Consequently it provides a feature vector of the same length as the standard histogram. In order to calculate the CDE we follow the definition given in [21], with the number of radius quantization levels set to $N = 4$ which results in a description of length 252.

The precision versus recall charts resulting from the 1600 queries are shown in Fig. 10, whereas the ANMRR values are given in Table VI. As expected, the total lack of spatial information places the standard color histogram last, while considerable improvements are achieved by both AC and CDE that provide

TABLE VI
ANMRR VALUES FOR THE TESTED COLOR DESCRIPTORS

Histogram	AC	CSD	CDE	CSG	MHW	MHL
0.4717	0.4341	0.4132	0.4289	0.3935	0.4156	0.3767

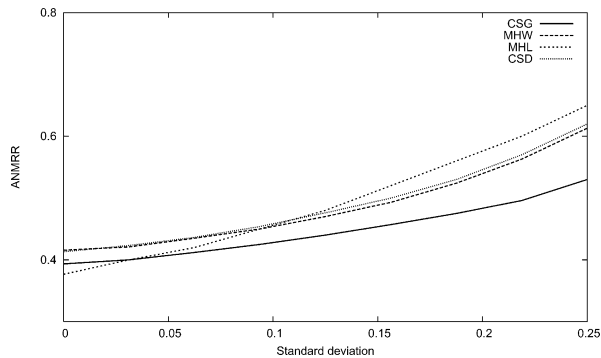


Fig. 11. Chart of ANMRR values with respect to various noise deviation levels of the tested color descriptors.

TABLE VII
RELATIVE COMPUTATION TIMES OF VARIOUS DESCRIPTORS TO THE STANDARD COLOR HISTOGRAM

Histogram	AC	CSD	CDE	CSG	MHW	MHL
1	127	89	9	548	470	443

likewise performances. Among the legacy descriptors, similarly to the study of Ojala *et al.* [23], the CSD gives the best retrieval results. As far as the proposed color description operators are concerned, with the exception of MHW, they all lead to superior performances, especially in terms of precision. In particular, we can observe that the MHW has a similar output with CSD, while both CSG and MHL outperform their counterparts, with the latter providing the best results. We now proceed to test their robustness against noise.

To test the noise robustness of the proposed descriptors, we have corrupted the images of our database with various levels of zero-mean additive Gaussian noise, and measured their retrieval performances. In particular, the images have been contaminated with noise in the RGB color space. The ANMRR values that have been obtained are shown in Fig. 11. This time they are compared only against CSD, as its noise robustness is known to be superior to its alternatives [20]. Judging from the obtained values, we can immediately remark the sensitivity of the previously outperforming MHL to noise levels, while MHW and CSD once more provide very similar outputs. CSG, however, exhibits a somewhat higher robustness. Further steps to improve their performance with corrupted image data could include the use of rank based morphological operators.

As to their computation times, they are given in Table VII with respect to that of a standard color histogram. Generally speaking all three morphology based approaches are significantly slower than their counterparts, nevertheless they still possess a certain margin for further efficiency optimization. Hence, one tends to conclude that they are unsuitable, at least in this

form, for online retrieval systems. Nevertheless, we have not yet used all optimized versions of morphological operators during the elaboration of the proposed descriptors. A deep execution analysis of our descriptors gives the following results. The main bottleneck within CSG is the closing by reconstruction step involved in the granulometry. Since we are using the efficient reconstruction algorithm from Vincent [53], most of the time is computed in the morphological closing operator itself. Possible ways of decreasing computation time are to involve efficient low-level (e.g., erosion/dilation) or high-level (e.g., granulometry) operators [54], perform SE decomposition to reduce the number of pixel comparisons, etc. As far as the multiresolution histograms are concerned, the high computation time is due to the ASF and levelings operations. Nevertheless, it is possible to rely on other multiscale watershed segmentations to produce the scale-space from which is computed the histogram. Such efficient multiscale segmentations have focusing the attention of the mathematical morphology community for years [47], [48] and is still very topical [9].

V. CONCLUSION

This paper has focused on the application of the morphological framework to the problem of color description. In particular, the main properties of color from a descriptive point of view have been determined and a state-of-the-art ordering approach has been implemented for the extension of mathematical morphology to color images. Subsequently, three global descriptors have been proposed, focusing mainly on the integration of spatial information with spectral color distribution through morphological means. Specifically, the proposed approaches include the color specific granulometry, which provides a size distribution independently computed for each subquantized color, and two multiresolution histograms based on morphological levelings and watersheds. The tests that have followed, have provided results asserting their competitive performance with respect to known alternatives. Nevertheless, the relatively high computation time that is required prohibits their use with online retrieval systems, while also leaving a considerable margin for speed optimization. A the field of mathematical morphology offers numerous efficient algorithms, this issue is not so critical.

In conclusion, the proposed morphological descriptors possess mixed properties, and while providing a retrieval performance as well as noise robustness comparable to contemporary approaches, they do have still efficiency issues that need to be resolved if they are to be used along with online CBIR systems. Nevertheless, since the existence of efficient morphological algorithms will most probably greatly reduce the computation time of proposed descriptors, this step into the field of CBIR-oriented morphological image content description will serve as a basis for the subsequent development of more sophisticated operators, e.g., morphological color interest point detectors, multiscale watershed segmentations.

REFERENCES

- [1] J. Serra, "Patent: 'Automatic Scanning Device for Analyzing Textures'," France 23 273, Jul. 2, 1965a.
- [2] G. Matheron, *Éléments Pour Une Théorie Des Milieux Poreux*. Paris, France: Masson, 1967.
- [3] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.
- [4] B. Naegel, N. Passat, and C. Ronse, "Grey-level hit-or-miss transforms—Part II: Applications to angiographic image processing," *Pattern Recognit.*, vol. 40, no. 2, pp. 648–658, Feb. 2007.
- [5] S. Lefèvre, J. Weber, and D. Sheeren, "Automatic building extraction in VHR images using advanced morphological operators," in *Proc. IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion Over Urban Areas*, Paris, France, Apr. 2007.
- [6] J. Weber and S. Lefèvre, "A multivariate hit-or-miss transform for joint spatial and spectral template matching," in *Proc. IEEE Int. Conf. Image and Signal Processing*, Cherbourg, France, Jul. 2008, vol. 5099, pp. 226–235, ser. Lecture notes in computer science.
- [7] J. Angulo, "Morphologie Mathématique et Indexation d'Images Couleur. Application à la Microscopie en Biomédecine," Ph.D. dissertation, Ecole des Mines, Paris, France, 2003.
- [8] M. H. F. Wilkinson and F. Tushabe, "Content-based image retrieval using shape-size pattern spectra," in *CLEF Workshop*, A. Nardi and C. Peters, Eds., Budapest, Hungary, 2007.
- [9] R. Lerallut, "Modélisation et Interprétation d'Images à l'Aide de Graphes," Ph.D. dissertation, Ecole des Mines, Paris, France, 2006.
- [10] I. Pratikakis, I. Vanhamel, H. Sahli, B. Gatos, and S. Perantonis, "Watershed-driven region-based image retrieval," in *Mathematical Morphology: 40 Years on*, ser. Computational Imaging and Vision, C. Ronse, L. Najman, and E. Decencière, Eds. Dordrecht, The Netherlands: Springer-Verlag, 2005, vol. 30, pp. 207–216.
- [11] F. A. Andaló, P. A. V. Miranda, R. D. S. Torres, and A. X. Falcão, "A new shape descriptor based on tensor scale," in *Proc. 8th Int. Symp. Mathematical Morphology*, U. M. Braga-Neto, G. J. F. Bannon, and J. Barrera, Eds., Rio de Janeiro, Brazil, Oct. 2007, vol. 1, pp. 141–152, INPE.
- [12] H. J. A. M. Heijmans, *Morphological Image Operators*, ser. Advances in Electronics and Electron Physics Series. Boston, MA: Academic, 1994.
- [13] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognit.*, vol. 40, no. 11, pp. 2914–2929, Nov. 2007.
- [14] A. Hanbury and J. Serra, "Colour image analysis in 3d-polar coordinates," in *Proc. DAGM*, Magdeburg, Germany, Sep. 2003, vol. 2781, pp. 124–131, ser. Lecture notes in computer science, Springer-Verlag.
- [15] A. Hanbury, "Constructing cylindrical coordinate colour models," *Pattern Recognit. Lett.*, vol. 4, no. 29, pp. 494–500, Mar. 2008.
- [16] J. Angulo, "Unified morphological color processing framework in a lum/sat/hue representation," in *Mathematical Morphology: 40 Years On*, ser. Computational Imaging and Vision, C. Ronse, L. Najman, and E. Decencière, Eds. Dordrecht, The Netherlands: Springer-Verlag, 2005, vol. 30, pp. 387–396.
- [17] J. Angulo and J. Serra, "Modelling and segmentation of colour images in polar representations," *Image Vis. Comput.*, vol. 25, no. 4, pp. 475–495, Apr. 2007.
- [18] E. Aptoula and S. Lefèvre, "On lexicographical ordering in multivariate mathematical morphology," *Pattern Recognit. Lett.*, vol. 29, no. 2, pp. 109–118, Jan. 2008.
- [19] A. Hanbury and J. Serra, "Morphological operators on the unit circle," *IEEE Trans. Image Process.*, vol. 10, no. 12, pp. 1842–1850, Dec. 2001.
- [20] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 703–715, Jun. 2001.
- [21] J. Sun, X. Zhang, J. Cui, and L. Zhou, "Image retrieval based on colour distribution entropy," *Pattern Recognit. Lett.*, vol. 27, no. 10, pp. 1122–1126, Oct. 2006.
- [22] W. Y. Ma and H. J. Zhang, "Benchmarking of image features for content-based image retrieval," in *Proc. 32nd Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 1998, vol. 1, pp. 253–257.
- [23] T. Ojala, T. Maenpää, M. Pietikainen, J. Viertola, J. Kyllönen, and S. Huovinen, "Outex: New framework for empirical evaluation of texture analysis algorithms," in *Proc. 16th ICPR*, Quebec City, Canada, Aug. 2002, vol. 1, pp. 701–706.
- [24] G. D. Finlayson, S. Hordley, G. Schaefer, and G. Y. Tian, "Illuminant and device invariant colour using histogram equalisation," *Pattern Recognit.*, vol. 38, no. 2, pp. 179–190, Feb. 2005.
- [25] T. Ojala, M. Rautiainen, E. Matinmikko, and M. Aittola, "Semantic image retrieval with HSV correlograms," in *Proc. 12th Scandinavian Conf. Image Analysis*, Bergen, Norway, 2001, pp. 621–627.
- [26] J. Angulo and J. Serra, "Morphological color size distributions for image classification and retrieval," in *Proc. Int. Conf. Advanced Concepts for Intelligent Vision Systems*, Ghent, Belgium, Sep. 2002, pp. 46–53.
- [27] G. Louverdis, I. Andreadis, and P. Tsalides, "Morphological granulometries for color images," in *Proc. 2nd Hellenic Conf. Artificial Intelligence*, Thessaloniki, Greece, Apr. 2002, pp. 333–342.
- [28] G. Matheron, *Random Sets and Integral Geometry*. New York: Wiley, 1975.
- [29] E. R. Urbach, J. B. T. M. Roerdink, and M. H. F. Wilkinson, "Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 272–285, Feb. 2007.
- [30] M. H. F. Wilkinson, "Generalized pattern spectra sensitive to spatial information," in *Proc. 16th ICPR*, Quebec City, Canada, Aug. 2002, vol. 1, pp. 21–24.
- [31] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlogram," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, San Juan, PR, Jun. 1997, pp. 762–768.
- [32] S. Lefèvre, "Beyond morphological size-distribution," *J. Electron. Imag.*, vol. 18, no. 1, 2009, article 013010.
- [33] F. Sand and E. Dougherty, "Robustness of granulometric moments," *Pattern Recognit.*, vol. 32, pp. 1657–1665, 1999.
- [34] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11–32, Nov. 1991.
- [35] A. Rao, R. K. Srihari, and Z. Zhang, "Spatial color histograms for content-based image retrieval," in *Proc. 11th IEEE Int. Conf. Tools With Artificial Intelligence*, Chicago, IL, Nov. 1999, pp. 183–186.
- [36] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Multiresolution histograms and their use for recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 831–847, Jul. 2004.
- [37] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. Int. Conf. Computer Vision and Pattern Recognition*, New York, Jun. 2006, vol. 2, pp. 2169–2178.
- [38] A. K. Jain and F. Farrokhi, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognit.*, vol. 24, no. 12, pp. 1167–1186, 1991.
- [39] M. Popovic, "Texture analysis using 2d wavelet transform: Theory and applications," in *Proc. 4th Int. Conf. Telecommunications in Modern Satellite, Cable and Broadcasting Services*, Nis, Serbia, Oct. 1999, vol. 1, pp. 149–158.
- [40] F. Meyer and P. Maragos, "Nonlinear scale space representation with morphological levelings," *J. Vis. Commun. Image Represent.*, vol. 11, no. 2, pp. 245–265, June 2000.
- [41] C. Vachier, "Morphological scale-space analysis and feature extraction," in *Proc. Int. Conf. Image Processing*, Thessaloniki, Greece, Oct. 2001, vol. 3, pp. 676–679.
- [42] F. Meyer, "Levelings, image simplification filters for segmentation," *J. Math. Imag. Vis.*, vol. 20, pp. 59–72, 2004.
- [43] C. Gomila and F. Meyer, "Levelings in vector spaces," in *Proc. IEEE Conf. Image Processing*, Kobe, Japan, Oct. 1999.
- [44] J. Angulo and J. Serra, "Morphological coding of color images by vector connected filters," in *IEEE Proc. 7th Int. Symp. Signal Processing and its Applications (ISSPA'2003)*, Paris, France, Jul. 2003, vol. 1, pp. 69–72.
- [45] K. Karantzalos, D. Argialas, and N. Paragios, "Comparing morphological levelings constrained by different markers," in *Proc. 8th Int. Symp. Mathematical Morphology*, U. M. Braga-Neto, G. J. F. Bannon, and J. Barrera, Eds., Rio de Janeiro, Brazil, Oct. 2007, vol. 1, pp. 113–124, INPE.
- [46] K. Grauman and T. Darrell, "Benchmarking of image features for content-based image retrieval," in *Proc. Int. Conf. Computer Vision*, Beijing, China, Oct. 2005, vol. 2, pp. 1458–1465.
- [47] S. Beucher, "Watershed, hierarchical segmentation and waterfall algorithms," in *Mathematical Morphology and its Applications to Image Processing*, J. Serra and P. Soille, Eds. Boston, MA: Kluwer, 1994, pp. 69–76.
- [48] L. Najman and M. Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 2, pp. 1163–1173, Feb. 1996.
- [49] Corel Corporation, Corel Gallery Images [Online]. Available: <http://www.corel.com>
- [50] G. Griffin, A. Holub, and P. Perona, Caltech-256 Object Category Dataset California Inst. Technol., Tech. Rep. 7694, 2007 [Online]. Available: <http://authors.library.caltech.edu/7694>
- [51] Google Search Engine, [Online]. Available: <http://www.google.com>
- [52] H. Y. Lee, H. K. Lee, and Y. H. Ha, "Spatial color descriptor for image retrieval and video segmentation," *IEEE Trans. Multimedia*, vol. 5, no. 3, pp. 358–367, May 2003.

- [53] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Trans. Image Process.*, vol. 2, no. 2, pp. 176–201, Apr. 1993.
- [54] L. Vincent, "Granulometries and opening trees," *Fundam. Inf.*, vol. 41, no. 1–2, pp. 57–90, 2000.



Erchan Aptoula received the B.Sc. degree in computer engineering from Galatasaray University, Istanbul, Turkey, in 2004, the M.Sc. and Ph.D. degrees in computer science from Louis Pasteur University, Strasbourg, France, in 2005 and 2008. His doctoral work focused on color mathematical morphology applied to content-based image retrieval.

Apart from morphological analysis, his interests include multispectral image processing, pattern recognition, and machine learning.



Sébastien Lefèvre received the M.Sc. and Eng. degrees in computer engineering from the University of Technology of Compiègne, France, in 1999, and the Ph.D. degree in computer science from the University of Tours, France, in 2002.

He is currently an Assistant Professor in the Department of Computer Science and the LSIT, University Louis Pasteur, Strasbourg. His research interests are in image/video processing using mainly mathematical morphology with applications in remote sensing, astronomical and medical imaging, and content-based image/video retrieval.

Supervised Segmentation Using Machine Learning and Evolutionary Computation

S. Derivaux, G. Forestier, C. Wemmert, S. Lefèvre*

*Image Sciences, Computer Sciences and Remote Sensing Laboratory, LSIT UMR 7005
CNRS–University of Strasbourg, Pôle API, Blvd Sébastien Brant, PO Box 10413, 67412
Illkirch Cedex, France*

Abstract

Automatic image interpretation could be achieved by first performing a segmentation of the image, i.e. aggregating similar pixels to form regions, then use a supervised region-based classification. In such a process, the quality of the segmentation step is of great importance. Nevertheless, whereas the classification step takes advantage from some prior knowledge such as learning sample pixels, the segmentation step rarely does. In this paper, we propose to involve machine learning to improve the segmentation process using the watershed transform. More precisely, we apply a fuzzy supervised classification and a genetic algorithm in order to respectively generate the elevation map used in the watershed transform and tune segmentation parameters. The results from our evolutionary supervised watershed algorithm confirm the relevance of machine learning as a mean to introduce knowledge within the watershed segmentation process.

Key words: supervised segmentation, watershed, genetic algorithms

1. Introduction

The goal of image understanding is to identify meaningful objects (from an user point of view) within an image. This process usually relies on two distinct steps: segmentation and classification. The segmentation clusters pixels into regions (i.e. assigns to each pixel a region label) whereas classification clusters regions into classes (i.e. assigns to each region a class label). A region is a set of connected pixels from which region features can be extracted (e.g. shape, textural indexes, ...). These extended features, which cannot be extracted at pixel level, should ease the classification step. Nowadays this kind of approach is widely used, in particular in the remote sensing field (refs).

To build an accurate classification, the segmentation should return a set of regions corresponding to the semantic objects contained in the image. How-

*Corresponding author. Tel: +33 (0)3 90 24 45 70; fax: +33(0)3 90 24 44 55
Email address: lefevre@lsiit.u-strasbg.fr (S. Lefèvre)

ever, this is hardly possible and most often two problems arise in the segmentation step: undersegmentation and oversegmentation. Undersegmentation occurs when a given region spans over objects of different classes. Whatever the classifier is, some parts of the region will be misclassified. Undersegmentation leads to errors that can not be recovered in the classification step. On the other side, oversegmentation occurs when a semantic object is covered by many regions. In this case, extracted attributes, especially shape and topological properties, are less representative of the object class. The classification, using such noisy attribute values will produce a lower quality result.

A segmentation algorithm is usually designed to cluster connected pixels according to an homogeneity criterion. For instance, in a gray level image, it seems a good idea to cluster connected pixels sharing a similar value. However, the final objective of the segmentation is that each resulting region perfectly matches one semantic object. For complex images (like very high resolution remotely sensed images) the assumption that a semantic object is composed of homogeneous pixels is no longer true. Indeed, a semantic object like a house could be quite heterogeneous, due to the presence of windows on the roof, or a different illumination on each side of the roof. Because of this wrong assumption, existing segmentation methods can not give satisfying results in such contexts.

In this paper, we propose a new supervised segmentation method relying in two ways on ground truth data. Firstly, ground truth is used to learn the better set of parameters for the segmentation using a genetic algorithm. Secondly, ground truth information are used to learn how to project the source image in a more relevant data space, where the homogeneity assumption is true and so where a well-known segmentation method can be applied.

First section of this paper presents a review of the related work. Then, we describe the proposed supervised segmentation methods: segmentation parameters optimization, space transformation and the hybrid method using the two approaches. Then, an evaluation of the algorithms is carried out on a very high remote sensing image. Finally, conclusion and some research directions are drawn.

2. Related work

The most frequent use of ground truth is to perform an optimization approach to find segmentation parameters (Martin et al., 2006; Pignatelli et al., 2003; Bhanu et al., 1995; Song and Ciesielski, 2003; Feitosa et al., 2006). This kind of methods involves a usual segmentation algorithm which can be tuned by a set of parameters. The genetic algorithm finds a set of parameters which optimize a fitness function. Different fitness functions were proposed using different segmentation criteria based on ground truth.

In the watershed algorithm, knowledge is often introduced using markers (Meyer and Beucher, 1990) and/or applying the watershed on a modified image. Many methods have been proposed for the choice of markers using knowledge, we review the major contributions hereafter.

Haker et al. (2000) use manually segmented images to extract a priori membership probabilities for each object to each class of interest. Then, they are combined using Bayes rule. Other kinds of knowledge on the data can be included in the process, for example spatial relation information on the objects of interest. This approach is comparable to a supervised classification and has the same problem of undersegmentation. Nevertheless, it produces better results if the user can approximatively determine the position of the objects in the scene.

Levner and Zhang (2007) proposed a method working with probability maps. They use a first classification, based on an eroded ground truth to find some seeds. Another classification is applied using original ground truth and the resulting inverted probability map is used as an elevation. This approach is currently only applied on binary classification. Also, this method assumes the detection of all seeds. If a seed is missed then the underlying object is not segmented.

Grau et al. (2004) use a probability map for each class of interest. Markers are generated using an atlas. Each marker has an associated class. A region growing approach is used to simulate flooding. The elevation between two pixels relies on the original marker class as it uses the difference on probability of these pixels in the probability map for the marker class (i.e. it is a markovian process). This approach needs the knowledge of markers locations.

Other forms of knowledge introduction in segmentation have been proposed. Hamarneh and Li (2007) perform a watershed segmentation with the classical oversegmentation problem. They use a modified k-means in order to cluster segments by intensity and position. Using appearance knowledge they select the appropriate cluster and iteratively align a shape histogram over the result to remove irrelevant remaining segments. This approach relies heavily on the assumption that objects have homogeneous intensity values, assumption which cannot be made in the paper context.

Chen et al. (2003) extract a shape and intensity model of the object of interest from a set of reference segmentations. After the learning step, they use an active contour model in order to segment the objects in respect with the shape and intensity model previously defined. This method works only for single object detection and approximative location need to be known.

From this brief review of related work, we can notice that involving knowledge into the segmentation process is a relevant idea which leads to several approaches proposed recently. In order to highlight our contribution and the goal of this paper, we point out the main features which differs our work from other existing approaches :

- ability to deal with many classes;
- no knowledge about the position of objects is needed;
- ability to deal with spectrally unseparable classes i.e. where marker creation using classification is not possible.

3. Segmentation parameters optimisation

A genetic algorithm (GA) is an optimization method (Gersho and Gray, 1992), based on a function to maximize, called the *fitness function*. The definition of this fitness function is a critical point of these methods. Indeed, the fitness has to evaluate the solutions proposed by the GA, in order to drive it to the *best* solutions. In the next section, we present and compare different kinds of segmentation evaluation criteria that could be used as fitness functions.

3.1. Segmentation evaluation

In the literature, many criteria for segmentation quality evaluation have been proposed. The reader can refer to (Zhang, 1996, 2001) for some surveys of this topic. In this paper, we do not consider all existing criteria, but rather focus on criteria based on discrepancy, i.e. comparing a resulting segmentation with some reference regions. This is particularly relevant since we are interested here in evaluation of GA methods in the context of optimal segmentation parameters learning. Criteria which are not based on learning samples are useless when investigating machine learning capabilities of the GA solutions.

Let us define reference samples as a set of connected components $R = \{R_i\}_{i \in [1; \Omega(R)]}$ where each connected component R_i is labeled with a class $C_k = c(R_i)$ from the set $C = \{C_k\}_{k \in [1; \Omega(C)]}$, with Ω the cardinality operator and c the class assignment function. For instance, we could define $C = \{\text{house, road, vegetation}\}$ in the remote sensing context. If no class are meaningful, we assign a new class to each reference sample, thus $c(R_i) = C_i$ and $\Omega(R) = \Omega(C)$. We also note R^{C_k} the set of reference samples, sharing the same class label, i.e. $R^{C_k} = \{R_i : c(R_i) = C_k\}$.

We can define three types of discrepancy criteria: classification errors criteria, matching criteria and generalization criteria. In our study, we illustrate these categories by a few representative criteria which will now be described.

3.1.1. Classification errors criteria

These criteria are based on the classification error principle. An image segmentation can be seen as an image classification process, and then, the percentage of misclassified pixels can be used. Since labels are assigned to both produced and reference regions, the number of pixels with different labels between the segmentation and the reference image can be computed.

The criterion we use here to represent this category is derived from the E criterion from Carleer et al. (2005). In the original paper, each reference region has a unique label. In our case, we assign to each reference region a class label. This way, reference regions which share the same semantics, have the same label. To each segmented region is then assigned the label of the most overlapping reference region (i.e. the region sharing the greatest number of pixels). Here we define the TMA criterion (Theoretical Maximum Accuracy), which uses class labels instead of a label for each region. If a segmented region spans over two reference regions of the same class, the TMA criterion does not track an error, whereas the E criterion does, as each reference region has a

different label. For each class, error is measured and weighted by the inverse number of reference pixels in order to give the same importance to each class. Then, a per-pixel confusion matrix K is computed. For each evaluation pixel of a class C_i , assigned to a label C_j by the matching, the value of the cell K_{ij} is incremented by $(\Omega(C_i))^{-1}$ where $\Omega(C_i)$ is the number of reference pixels for class C_i . Thus, the evaluation function TMA is the classifier precision (the overall accuracy):

$$TMA = \frac{1}{\Omega(C)} \sum_{i=1}^{\Omega(C)} K_{ii} \quad (1)$$

The TMA criterion gives the best available accuracy of a subsequent classification step of the resulting segments.

3.1.2. Matching criteria

Matching criteria measure spatial differences between segmented and reference regions. They rely on a matching function $m(R_i, S_j)$ which computes a matching score between a reference region R_i and a segmented region S_j , where $S = \{S_j\}_{j \in [1; \Omega(S)]}$ is the set of segmented regions. Let us additionally define R_{S_j} the set of reference regions overlapping S_j , and inversely S_{R_i} the set of segmented regions overlapping R_i . To apply these criteria on a complete segmentation, the average matching value μ_m of the best matching score for each reference region is computed:

$$\mu_m = \frac{1}{\Omega(R)} \sum_{i=1}^{\Omega(R)} \text{best}_{1 \leq j \leq \Omega(S)} (m(R_i, S_j)) \quad (2)$$

where the best function is the optimum function, i.e. minimum or maximum function depending on the matching criterion.

The first criterion used here is taken from Feitosa et al. (2006) and defined by:

$$F(R_i, S_j) = \frac{\Omega(R_i \setminus (R_i \cap S_j)) + \Omega(S_j \setminus (R_i \cap S_j))}{\Omega(R_i)} \quad (3)$$

We observe that the F criterion favors oversegmentation over undersegmentation and should be minimized to obtain the best segmentation.

The second criterion is taken from Janssen and Molenaar (1995). It is quite similar to F but does not have the bias to avoid oversegmentation. It considers reference and segmented regions in the same way and should be maximized.

$$J(R_i, S_j) = \sqrt{\frac{\Omega(R_i \cap S_j)^2}{\Omega(R_i) \times \Omega(S_j)}} \quad (4)$$

In this formulation, if a segmented region S_j spans over two reference regions R_i and $R_{i'}$ of the same class C_k , both matching scores $J(R_i, S_j)$ and $J(R_{i'}, S_j)$

will be low. Nevertheless, as R_i and $R_{i'}$ belongs to R^{C_k} , they could be merged, thus resulting in a high matching score $J(R_i \cup R_{i'}, S_j)$.

This principle leads to a new criterion JC which relies on class labels. For a given couple (R_i, S_j) , we consider the subset of $R^{c(R_i)} = \{R_{i'} : c(R_{i'}) = c(R_i)\}$ (i.e. the union of all reference regions $R_{i'}$ sharing the label assigned to R_i) overlapping S_j , or $R_{S_j}^{c(R_i)} = R^{c(R_i)} \cap S_j$. The modified criterion is then:

$$JC(R_i, S_j) = \sqrt{\frac{\Omega(R^{c(R_i)} \cap S_j)^2}{\Omega(R_i) \times \Omega(S_j)}} \quad (5)$$

A similar evaluation criterion is the Jaccard index (Jaccard, 1912) which should also be maximized. It is defined as the ratio between the cardinalities of the intersection and the union of the two sets:

$$J'(R_i, S_j) = \frac{\Omega(R_i \cap S_j)}{\Omega(R_i \cup S_j)} \quad (6)$$

Here, we also extend this criterion to handle class labels:

$$JC'(R_i, S_j) = \frac{\Omega(R^{c(R_i)} \cap S_j)}{\Omega(R_i \cup S_j)} \quad (7)$$

We can also mention the ultimate measurement accuracy criterion (Zhang and Gerbrands, 1992), which measures the difference between features extracted from R_i and S_j . Since it strongly depends on the regional features extracted, and thus, is hardly compatible with a generic solution for parameter tuning, we do not consider this criterion in our study.

3.1.3. Generalization criteria

Generalization criteria measure the coarseness of the segmentation.

The *Gen* criterion (Carleer et al., 2005) measures oversegmentation through a simple ratio between the number of segmented and reference regions, i.e. $Gen = \Omega(S)/\Omega(R)$.

Here we consider only segmented regions spanning over a reference one, in order to deal with an incomplete reference segmentation. Moreover, we take into account class information and compute the average oversegmentation for all classes. Thus the proposed criterion *OV* is defined as:

$$OV = \frac{1}{\Omega(C)} \sum_{k=1}^{\Omega(C)} \frac{\Omega(S_{R^{C_k}})}{\Omega(R^{C_k})} \quad (8)$$

where $S_{R^{C_k}}$ denotes the set of segmented regions overlapping at least one of the reference region assigned to the class C_k while R^{C_k} is the set of reference regions assigned to the class C_k .

Another criterion belonging to this category is the average region size (noted p/r), i.e. $\Omega(I)/\Omega(S)$ where $\Omega(I)$ and $\Omega(S)$ represent respectively the number of pixels in the image and the number of regions produced by the segmentation.

It is rather simplistic and does not involve any sample. Nevertheless, it allows to compare two segmentations to determine the coarsest one.

3.1.4. Hybrid criteria

Among the previous criteria, some criteria measure mainly oversegmentation (e.g. *OV* and *p/r*) while others measure mainly undersegmentation (e.g. *TMA*). So it is relevant to combine these single criteria and build some aggregated criteria. Combination is one solution for resolving multiobjective optimization. Another solution is to use the Pareto front (Fonseca and Fleming, 1996). The Pareto front returns a set of results representing different trade-offs between all the considered criteria. Thus, handling a set of results needs more user interaction, which is out of the scope of this paper.

We propose here two multiobjective criteria, combining *TMA* and *OV*.

The first one *TMA/OV*, avoids mainly undersegmentation (using *TMA*) and secondarily oversegmentation (using *OV*). It is simply defined by weighting *OV* with a small coefficient (ε):

$$TMA/OV = TMA + \varepsilon \frac{1}{OV} \quad (9)$$

The second criterion is $TMA \oplus OV(\alpha)$. It also primarily rely on undersegmentation (using *TMA*), but limits its effect with the α parameter:

$$TMA \oplus OV(\alpha) = \min(TMA, \alpha) + \varepsilon \frac{1}{OV} \quad (10)$$

Of course the α parameter is dependent of the application. It represents the amount of errors (measured by the *TMA* criterion) tolerated by the user or system. For instance, if the *TMA* quality should be at least 95%, the user sets $\alpha = 0.95$.

3.2. Segmentation optimisation

We choose to use the watershed segmentation as base algorithm for our experiments. It is a well-known segmentation method which considers the image to be processed as a topographic surface. In the immersion paradigm from Vincent and Soille (1991), this surface is flooded from its minima, thus generating different growing catchment basins. Dams are built to avoid merging water from two different catchment basins. The segmentation result is defined by the locations of the dams (i.e. the watershed lines) when all the image have been flooded, as illustrated in Figure 1.

In this approach, the topographic surface is most often built from an image gradient, since object edges (i.e. watershed lines) are most probably located at pixels with high gradient values. Different techniques can be involved to compute the image gradient. Since it does not affect our study, we consider here as an illustrative example, the morphological gradient (Soille, 2003) computed independently for each band and combined through an Euclidean norm.

In its markerfree version, the watershed segmentation is proven to easily generate an oversegmentation (i.e. a segmentation where the number of regions

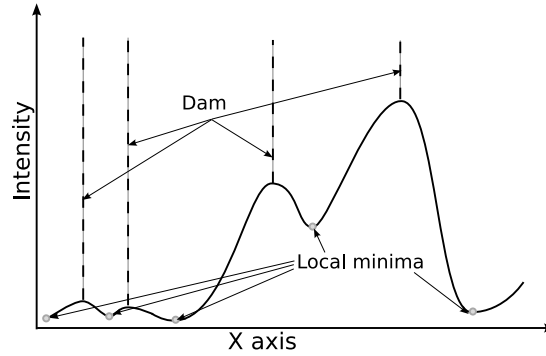


Figure 1: Illustration of the watershed segmentation principle.

created is far larger than the number of actual regions in the image). A smoothing filter is often applied on the input image to overcome this problem. Here we have decided to process all channels of the image with a median filter (with a 3×3 structuring element which is adequate for our task) which is known to preserve edges.

To further reduce oversegmentation, we use several other methods. First, the gradient thresholding method (Haris et al., 1998) is used. On the final gradient image, each pixel with a value below a threshold is set to zero. This step removes small heterogeneity effects. On Figure 2, this step is represented by the *hmin* line, all values under this line are set to null, and thus, two watersheds are removed.

The concept of dynamics (Najman and Schmitt, 1996) is also involved. Catchment basins with a dynamic under a threshold are filled. On Figure 2 this step is represented by the catchment basin which starts from A. If its dynamic *d* is below the threshold, this catchment basin is filled and the left watershed is removed.

The last method used is region merging (Haris et al., 1998). We characterize each produced region by its mean on each original source channel. If the Euclidian distance between the feature vector of two neighboring regions is below a threshold *M*, these two region are merged.

To summarize, the *segmentation* method used here needs three parameters.

Let us emphasize that this base segmentation method (and its parameters) is just a simple example to illustrate our contribution which consists in a general evolutionary framework for optimizing segmentation parameters. Another segmentation algorithm could have been used instead.

Optimization algorithm. As it has been underlined previously, the base segmentation algorithm requires four parameters to be set. We explain here how the genetic algorithm proceeds to tune these parameters.

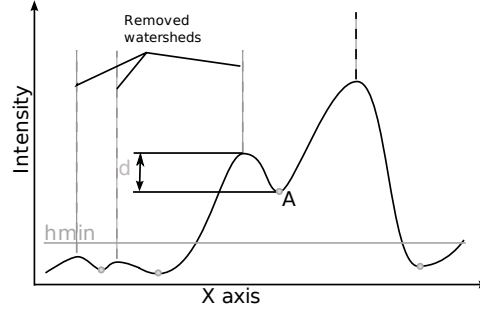


Figure 2: Illustration of oversegmentation reduction methods for the watershed transform.

Given an evaluation function $f(G)$ where G (the genotype in the genetic framework) is taken in a space $\mathbb{G}$, the GA searches the optimal value of G , i.e. $\arg \max_{G \in \mathbb{G}} f(G)$. GA are known to be effective even if $f(G)$ contains many local minima. This optimization can be considered as a learning process, if and only if it is performed on a learning set but can be generalized to other (unlearned) datasets.

The genotype G is defined as an array containing the parameters to be automatically tuned in the watershed segmentation process, i.e. $G = [\omega_1, \dots, \omega_{b+1}, t]$ with all parameters normalized into $[0; 1]$.

A GA requires an initial population defined as a set of genotypes, to perform the evolutionary process. In this process, the population evolves to obtain better and better genotypes, i.e. solutions of the optimization problem under consideration. In order to build the initial population, each genotype is randomly chosen in the space $\mathbb{G}$.

Once the initial population has been defined, the algorithm relies on the following steps, which represent the transition between two generations:

1. assessment of genotypes in the population: genotypes are sorted by their relevance;
2. selection of genotypes for crossover weighted by their rank;
3. crossover: two genotypes (G_1 and G_2) breed by combining their parameters (or genes in the genetic framework) to give a child E . The resulting child is E with $E[i] = G_{p_i}[i] + \alpha_i \times |G_1[i] - G_2[i]|$ where α_i and p_i are randomly selected in $[-1; 1]$ and $\{1, 2\}$ respectively. We apply an elitist procedure and keep the best solution of the current generation in the next generation;
4. mutation: each parameter may be replaced by a random value with a probability P_m . Thus, we avoid the GA to be trapped in a local minimum. As indicated previously, the best genotype of a generation is kept unchanged.

In our study, we use the following parameters for the GA: a population size

of 15 genotypes, a mutation probability P_m of 1%, and an evolution number $N = 30$ generations (experiments shown that no significant improvement is obtained with more generations). The results are presented in the Section 6.

4. Supervised segmentation by space transformation

In the previous section, we proposed an optimization method, that automatically finds the best parameters for a segmentation algorithm. It enables to produce better results, because it uses the information contained in the available examples given by the user. However, as the segmentation algorithm is not modified, it is not possible to extract non-heterogeneous objects from the image. Indeed, the segmentation algorithm assumes that the objects are composed of homogeneous pixels in the image. This assumption is not correct when dealing with very high resolution remote sensing images. Thus, we propose here another approach that modifies the data space in which the segmentation is applied.

The main idea is to use the examples given by the user to define a new homogeneity between the pixels. For this, we project the pixels in a new data space in which the sample regions are composed of homogeneous pixels. Then, classical segmentation algorithms can be applied and should give better results (according to the samples given by the user).

A graphical representation of the supervised segmentation process is presented in Figure 3. The proposed method breaks down into two parts:

- fuzzy classification: based on the samples given by the user;
- watershed segmentation: the segmentation is applied on the membership map given by the fuzzy classification (not on the image).

Let S_i be the input space defined by Eq. 11.

$$\begin{aligned} S_i : E &\rightarrow \mathbb{R}^i \\ x &\mapsto S_i(x) \quad \text{with } S_i(x) \text{ the spectral} \\ &\quad \text{signature of the pixel } x \end{aligned} \quad (11)$$

The problem is that no function exists which takes a pixel of S_i , assigns a class to this pixel and makes no errors. As only approximation functions exist, let S_m be the memberships space defined by Eq. 12. In the memberships space, each class of objects (let $|C|$ be the number of classes) contained in the image is a dimension. The value in each dimension denotes the membership of the pixel to the corresponding class of objects.

$$\begin{aligned} S_m : E &\rightarrow [0; 1]^{|C|} \\ x &\mapsto S_m(x) \quad \text{with } S_m(x) \text{ the memberships} \\ &\quad \text{vector of the pixel } x \end{aligned} \quad (12)$$

For the transformation from S_i to S_m , we propose to use the available ground truth to learn the transformation using data mining tools.

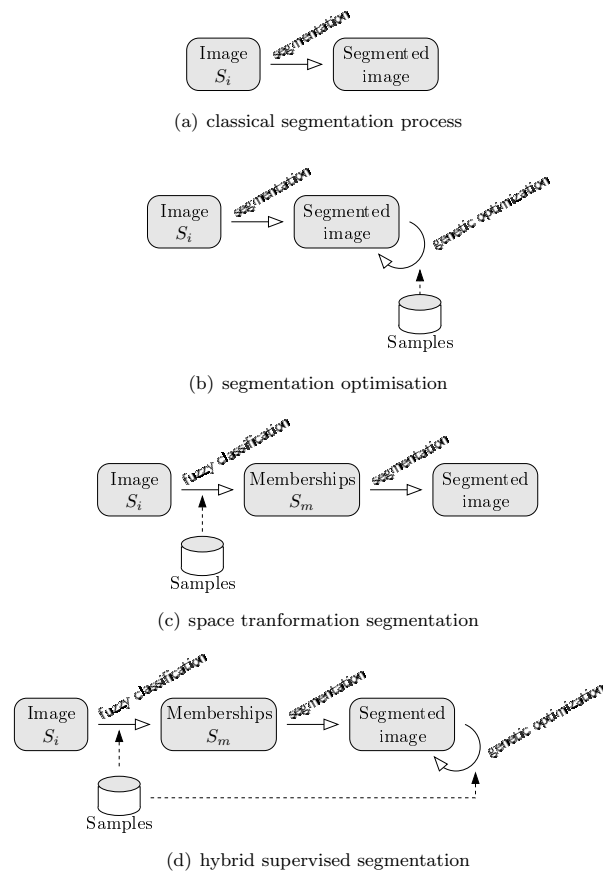


Figure 3: Supervised segmentation process

In our case, the *fuzzy classification* step induces a N nearest neighbors classifier (Aha et al., 1991). For each input pixel p , the N (with $N = 5$) nearest labeled pixels in the S_i space are selected. Each neighboring pixel p_n will increase the membership degree of the class it has been labeled with, weighted by the inverse of the distance $d(p, p_n)$ in the feature space. The memberships $m_{p,k}$ are then obtained by:

$$m_{p,k} = \left(\sum_{n=1}^N \sum_{l=1}^K w_{n,l} \right)^{-1} \sum_{n=1}^N w_{n,k} \quad (13)$$

$$\text{where } w_{n,k} = \begin{cases} d(p, p_n)^{-1} & \text{if } p_n \text{ is labeled with class } k \\ 0 & \text{otherwise} \end{cases}$$

5. Hybrid approach

In this section, we describe a hybrid method, integrating the two previous ideas presented above. In an offline phase, the method learns how to segment an image using a learning set (composed of images and masks corresponding to objects of interest). The learning process occurs in two steps: a space transformation step and a core segmentation step. Once the learning is finished, a segmentation algorithm (i.e. the space transformation step and the core segmentation step) is produced and can be used to segment images. No learning set is needed in this application phase. The proposed method does not need input parameters in neither phases. A flow chart is shown on Figure 3.

The learning set is composed of learning images and corresponding learning masks. A learning mask is a semantic interpretation of a learning image made by a human. For each object, the corresponding pixels in the image are labeled with a class C_k where $k \in [1 \dots K]$ and K is the number of classes. Some pixels could be left unlabeled, denoting the inability to label them.

5.1. Segmentation supervision by genetic algorithm

Here we propose a genetic algorithm in order to handle the parameters from the *segmentation* step. As already stated in Section 3.2 The watershed algorithm needs three parameters to be set: $hmin$ to reduce oversegmentation, d for the dynamics and M as the threshold for the region merging. In the space transform segmentation algorithm, another parameter is added, which is the same as the M threshold, but applied with the mean of memberships maps and a threshold M_m . Thus, we have four parameters to optimize.

5.2. Evaluation function

As already discussed in the Section 3.1, a critical point of the genetic algorithm optimization method is the way the quality of the potential solutions (i.e. genotypes) is estimated. Here, as we are interested in evaluation of segmentation results, we focus on empirical discrepancy evaluation methods following the work from Carleer et al. (2005). Nevertheless, our criteria are adapted to

both mixed and user-meaningless pixels which do not appear in such a manual reference segmentation. They are compatible with partially segmented images defined as (incomplete) sets of labeled pixels. We use the term region for a reference region given by the user and the term segment for a region of a produced segmentation.

From the evaluation criteria introduced in Section 3.1, we can define the evaluation function. We can choose to optimize one of the two criteria or a combination of them. Here, we chose to optimize a criterion which represents oversegmentation and undersegmentation using:

$$\mathbb{F}(g) = \frac{1}{OV(g)} \times \max(0, TMA(g) - 0.98) \quad (14)$$

In the proposed function, $\mathbb{F}(g)$ increases as $OV(g)$ is reaching 1 (no oversegmentation) and decreases when $TMA(g)$ decreases. The function is null if $TMA(g)$ is under 98%, i.e. the maximum accuracy is 98% well classified pixels. This threshold was set to give more importance to avoid undersegmentation. It could be modified by the user depending on the image noise and complexity. 98% seems a good compromise. If $TMA(g)$ falls below this threshold the resulting segmentation will be useless.

6. Evaluation

The evaluation of the proposed algorithm follow the evaluation scheme proposed by Zhang (1996) using an analytical evaluation and a empirical discrepancy evaluation. The empirical goodness evaluation is not performed as it is not relevant. An empirical goodness evaluation usually relies on the assumption that segment are spectrally homogeneous.

6.1. Analytical evaluation

The first part of the evaluation is an analytical review of the proposed algorithm. Such a review is helpful to know if the algorithm is suitable to an image or not. The proposed algorithm requires some knowledge from the user to be able to segment an image:

- Class knowledge: the user needs to know the classes of objects which are seeked in the image.
- Samples for each class: some samples of each class are needed for the learning step. The *fuzzy classification* step can work with isolated samples, but the *genetic optimization* step requires labeling of image parts.

There are also some limits which should be noted in the proposed algorithm:

- Connected objects of the same class: if two objects of the same class are spatially connected and have similar memberships to classes, they will be merged together (i.e. undersegmentation). The same problem arises in usual segmentation methods when two objects have similar spectral values.

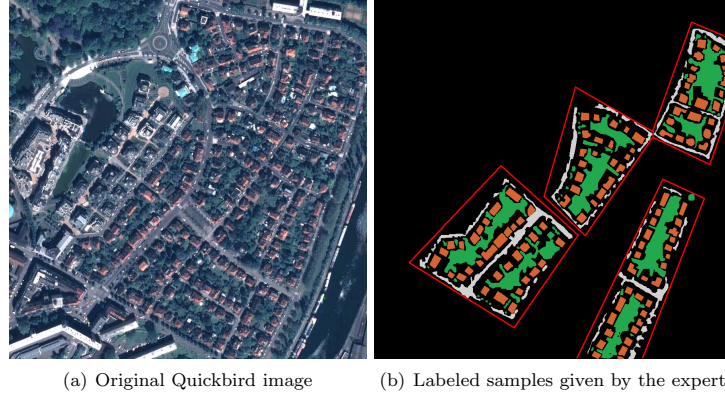


Figure 4: Remotely sensed image of a part of Strasbourg (France)

- Objects having heterogeneous spectral values and memberships values: in such a case, the algorithm produces an oversegmentation.

Nevertheless, these limits are weaker than those of classical segmentation algorithms. If an object has heterogeneous spectral and memberships values, it will be oversegmented by classical segmentation methods. The case where two spatially connected objects have similar memberships values and dissimilar spectral values and each object has homogeneous spectral values seems less frequent than objects with heterogeneous spectral values. It is a tradeoff that should be analyzed depending on the application.

6.2. Application to a real urban image

This case study is a typical example of VHSR image interpretation in remote sensing, where a segmentation is first performed before applying a supervised region-based classification. The input data is a pan-sharpened Quickbird¹ image of the city of Strasbourg (France) with 4 spectral bands (Figure 6.2). The spatial resolution is 0.7 meter per pixel. In four areas of the studied zone, some regions have been labeled by the expert in three classes: road, vegetation and house (Figure 6.2).

Choice of the fitness function. The aim of the first set of experiments carried out on this data was to evaluate the influence of the choice of the fitness function. Indeed, we presented in Section 3.1 many criteria that could be used as fitness function to optimize the parameters of the segmentation methods. The

¹image provided by the LIVE laboratory from University of Strasbourg

	<i>TMA</i>	<i>OV</i>	<i>F</i>	<i>JC</i>	<i>JC'</i>
<i>TMA/OV</i>	98.036	48.016	0.226	0.529	0.442
<i>TMA</i> $\oplus$ <i>OV</i>	99.125	95.100	0.185	0.434	0.364
<i>F</i>	98.569	61.286	0.243	0.534	0.448
<i>JC</i>	96.747	34.833	0.214	0.567	0.484
<i>JC'</i>	96.910	41.124	0.212	0.555	0.488

Table 1: Watershed parameters optimisation

	<i>TMA</i>	<i>OV</i>	<i>F</i>	<i>JC</i>	<i>JC'</i>
<i>TMA/OV</i>	98.054	4.51	0.121	0.661	0.527
<i>TMA</i> $\oplus$ <i>OV</i>	99.504	28.79	0.312	0.646	0.575
<i>F</i>	99.408	23.59	0.317	0.653	0.576
<i>JC</i>	98.275	7.58	0.190	0.689	0.572
<i>JC'</i>	99.175	12.88	0.279	0.676	0.592

Table 2: Probashed parameters optimization

question is *which criteria shall we optimize to obtain the best result ?* We performed a genetic optimization on two segmentation algorithms proposed before: classical watershed and *probashed* (which corresponds to the space transformation method given in Section 4). For the watershed algorithm, three parameters have to be tuned as stated in Section 3.2: *hmin*, *d*, and *M*. For the *probashed* algorithm, four parameters are used (Section 5) : *hmin*, *d*, *M* and *M_m*.

In our experiments, we consider the following parameters for the genetic algorithm: a population size of 15 genotypes, a mutation probability $\mathbb{P}_m$ of 1% and an evolution number equals to 90 generations. Experiments show that stability and convergence is achieved at this step.

We present in Tab.1 and Tab.2 the results obtained by optimizing the parameters of the segmentation method. The first column shows the criterion that has been used as fitness function. Then, each column corresponds to the value obtained by the final result for each evaluation criterion.

The first remark concerns the three last lines of the two tables. It is obvious that optimizing one criterion will produce the best result for this criterion. This is verified on these results for the three criteria *F*, *JC* and *JC'*.

Concerning the hybrid criteria, *TMA* $\oplus$ *OV* seems to be a better compromise as *TMA/OV* because it optimizes well the *TMA* criterion, without having bad results with the other ones.

Comparison of the different approaches proposed. The second experiment tries to compare the different approaches proposed in this paper. To have a more thorough study, we also included two results given by two commercial remote sensing segmentation software: eCognitionTM from Definiens² and ENVI FX

²<http://earth.definiens.com/>

	<i>TMA</i>	<i>OV</i>	<i>F</i>	<i>JC</i>	<i>JC'</i>
watershed	99.180	99.04	0.171	0.415	0.300
watershed+genetic	98.569	61.286	0.243	0.534	0.448
probashed	99.520	24.33	0.317	0.655	0.483
probashed+genetic	99.408	23.59	0.317	0.653	0.576
eCognition	91.420	35.26	0.129	0.483	0.512
ENVI FX	84.950	2.75	0.013	0.473	0.598

Table 3: Comparison of the different approaches proposed and two commercial segmentation software

from ITT Visual Information Solutions ³. These results were manually computed by a geographer expert.

Again, we present in Tab.3 the evaluations calculated from the different evaluation criteria on the results given by the different proposed methods. For the optimised methods, we only give the result when optimizing the *F* criterion for better readability. We choose *F* because it has good results with quite all the evaluation criteria.

As it is difficult to grasp the influence of a small change on a criterion, we show in Figure 5 the segmentations produced by the different methods. Thus, it is possible to have a visual appreciation of the quality of the results. It is clear that the watershed, even in its optimized version, produces results that could not be used directly in the classification step. For example, the vegetation zones in the blocks are really oversegmented as well as the houses. It is then very difficult to use geometrical attributes in the classification, as the shape of the regions does not necessarily correspond to the expected one.

When comparing the probashed method and its optimized version, the values for the evaluation criteria are comparable or better for the optimized version. But the main differences are visible on the segmentation results (Figure 5). It is obvious that the river (East of the image) is better delimited as the houses in the blocks.

7. Conclusion

In this article, we presented and compared different criteria to optimize segmentation parameters, when examples are available. We also exposed another way to take advantage of ground truth, in changing the data space before applying the segmentation algorithm. The space transformation is performed by a fuzzy classification based on the examples given by the expert. It has been shown that using this knowledge to guide the segmentation enables to produce better results, even better than manually produced segmentations by an expert.

In future work, we would like to focus on the study of the integration of other kinds of knowledge (not only examples) in the segmentation process. For

³<http://www.ittvis.com/>

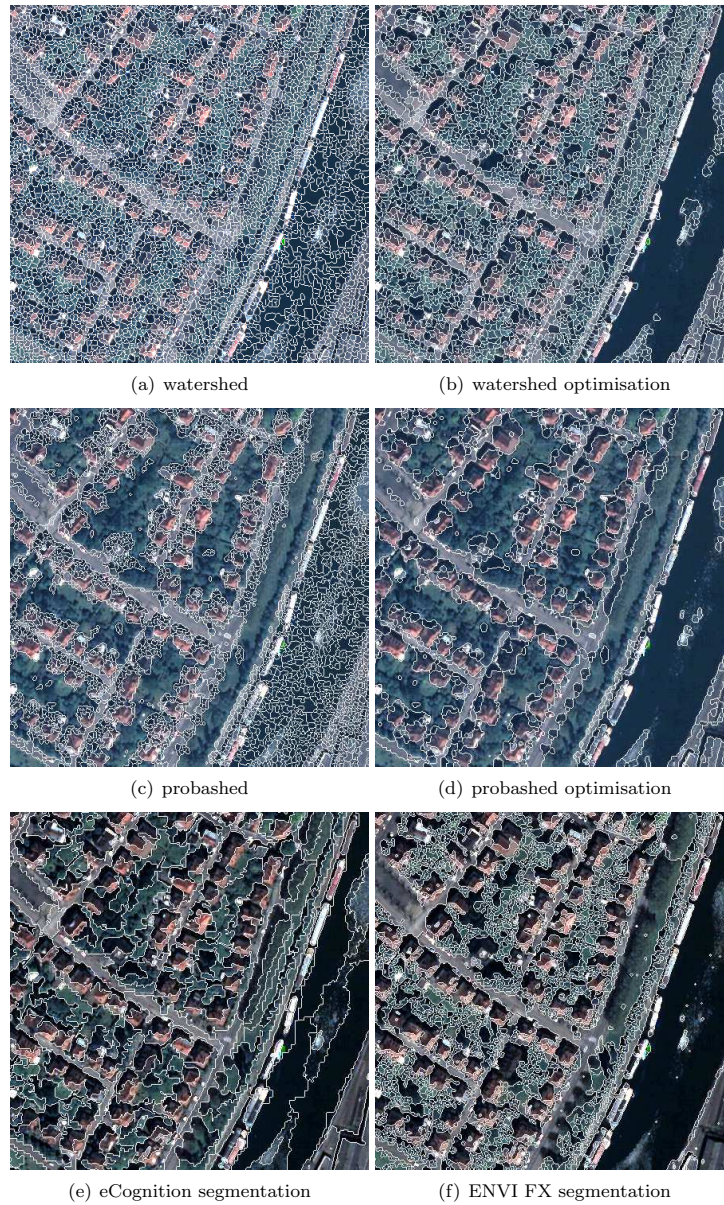


Figure 5: Segmentation results obtained by the different approaches proposed (extract from the studied image)

example, a hierarchy of concepts describing the objects of interest could help to better identify which regions are well segmented. We also plan to use several segmentation algorithms and make them collaborate to find a better segmentation.

References

- Aha, D. W., Kibler, D. F., Albert, M. K., 1991. Instance-based learning algorithms. *Machine Learning* 6, 37–66.
- Bhanu, B., Lee, S., Das, S., 1995. Adaptive image segmentation using genetic and hybrid search methods. *IEEE Transactions on Aerospace and Electronic Systems* 31 (4), 1268–1291.
- Carleer, A. P., Debeir, O., Wolff, E., 2005. Assesment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering and Remote Sensing* 71 (11), 1285–1294.
- Chen, Y., Huang, F., Tagare, H., Rao, M., Wilson, D., Geiser, E., 2003. Using prior shape and intensity profile in medical image segmentation. In: *Proceedings of the Ninth IEEE International Conference on Computer Vision*. pp. 1117–1124.
- Feitosa, R. Q., Costa, G. A., Cazes, T. B., B., F., 2006. A genetic approach for the automatic adaptation of segmentation parameters. In: *International Conference on Object-based Image Analysis*.
- Fonseca, C. M., Fleming, P. J., 1996. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary Computation* 1 (3), 1–16.
- Gersho, A., Gray, R. M., 1992. *Vector quantization and signal compression*. Kluwer Academic Publishers.
- Grau, V., Mewes, A., Alcaniz, M., Kikinis, R., Warfield, S., 2004. Improved watershed transform for medical image segmentation using prior information. *IEEE Transactions on Medical Imaging* 23 (4), 447–458.
- Haker, S., Sapiro, G., Tannenbaum, A., 2000. Knowledge-based segmentation of SAR data with learned priors. *IEEE Transactions on Image Processing* 9 (2), 299–301.
- Hamarneh, G., Li, X., 2007. Watershed segmentation using prior shape and appearance knowledge. *Image and Vision Computing*. doi:10.1016/j.imavis.2006.10.009.
- Haris, K., Efstradiadis, S. N., Maglaveras, N., Katsaggelos, A. K., 1998. Hybrid image segmentation using watersheds and fast region merging. *IEEE Transaction On Image Processing* 7 (12), 1684–1699.
- Jaccard, P., 1912. The distribution of flora in the alpine zone. *The New Phytologist* 11 (2), 37–50.
- Janssen, L., Molenaar, M., 1995. Terrain objects, their dynamics and their monitoring by the integration of gis and remote sensing 33, 749–758.

- Levner, I., Zhang, H., 2007. Classification-driven watershed segmentation. *IEEE Transaction on Image Processing* 16 (5), 1437–1445.
- Martin, V., Maillot, N., M., T., 2006. A learning approach for adaptive image segmentation. In: *IEEE International Conference on Computer Vision Systems*. pp. 40–48.
- Meyer, F., Beucher, S., 1990. Morphological segmentation. *Journal of Visual Communication and Image Representation* 1 (1), 21–46.
- Najman, L., Schmitt, M., 1996. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (12), 1163–1173.
- Pignalberi, G., Cucchiara, R., Cinque, L., Levialdi, S., 2003. Tuning range image segmentation by genetic algorithm. *EURASIP Journal on Applied Signal Processing* 2003 (8), 780–790.
- Soille, P., 2003. *Morphological Image Analysis*, 2nd Edition. Springer-Verlag.
- Song, A., Ciesielski, V., 2003. Fast texture segmentation using genetic programming. In: *IEEE Congress on Evolutionary Computation*. Vol. 3. pp. 2126–2133.
- Vincent, L., Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Pattern Analysis and Machine Intelligence* 13 (6), 583–598.
- Zhang, Y. J., 1996. A survey on evaluation methods for image segmentation. *Pattern Recognition* 29 (8), 1335–1346.
- Zhang, Y. J., 2001. A review of recent evaluation methods for image segmentation. In: *International Symposium on Signal Processing and its Applications*. Vol. 1. pp. 148–151.
- Zhang, Y. J., Gerbrands, J. J., 1992. Segmentation evaluation using ultimate measurement accuracy. In: *Proc. SPIE, Image Processing Algorithms and Techniques III*. Vol. 1657. pp. 449–460.



Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

A robust hit-or-miss transform for template matching applied to very noisy astronomical images

B. Perret*, S. Lefèvre, Ch. Collet

Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT), Université de Strasbourg, UMR CNRS 7005, France

ARTICLE INFO

Article history:

Received 24 July 2008

Received in revised form 23 February 2009

Accepted 28 February 2009

Keywords:

Hit-or-miss transform

Template matching

Mathematical morphology

Noisy environment

Astronomy

Low surface brightness galaxies

ABSTRACT

The morphological hit-or-miss transform (HMT) is a powerful tool for digital image analysis. Its recent extensions to grey level images have proven its ability to solve various template matching problems. In this paper we explore the capacity of various existing approaches to work in very noisy environments and discuss the generic methods used to improve their robustness to noise. We also propose a new formulation for a fuzzy morphological HMT which has been especially designed to deal with very noisy images. Our approach is validated through a pattern matching problem in astronomical images that consists of detecting very faint objects: low surface brightness galaxies. Despite their influence on the galactic evolution model, these objects remain mostly misunderstood by the astronomers. Due to their low signal to noise ratio, there is no automatic and reliable detection method yet. In this paper we introduce such a method based on the proposed hit-or-miss operator. The complete process is described starting from the building of a set of patterns until the reconstruction of a suitable map of detected objects. Implementation, running cost and optimisations are discussed. Outcomes have been examined by astronomers and compared to previous works. We have observed promising results in this difficult context for which mathematical morphology provides an original solution.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Template matching is a subject of interest in many computer vision processes and more generally in digital image analysis. In a general formulation it consists of searching for occurrences of a pre-defined object in an image. The objects of interest can have several different visual appearances which lead to several patterns sought in the images.

Among the methodologies to be used for pattern matching in digital images, mathematical morphology is particularly relevant since it performs spatial based image analysis relying on a pattern called structuring element (SE). The hit-or-miss transform (HMT) is a well-known morphological operator dedicated to template matching. Initially defined for binary images, it has been extended recently to greyscale images [1].

The use of the morphological HMT to perform template matching has grown during the recent years. Different formulations have been proposed and their application to various domains (medical

imaging [2,3], remote sensing [4,5], character recognition [6], and face localisation [7]) have proven their ability to provide a generic pattern matching method. However, it lacks robustness to noise and thus cannot be yet considered as a valid option for object detection in noisy images, even if some promising results have already been obtained in other challenging tasks such as galaxy segmentation and classification [8].

In this paper we propose a new formulation of the morphological HMT for grey level images. This formulation leads to a fuzzy result and can be applied even on very noisy images with a signal to noise ratio (SNR) below 0 dB. An application to detect faint objects: low surface brightness (LSB) galaxies [9] in astronomical images is also proposed. Results obtained with our new method are compared to those obtained with a traditional correlation algorithm. Detecting LSB galaxies is a very topical problem since astronomers lack information about them. It also well fits our interest of adapting the HMT to deal with very noisy environment for several reasons. The shape and the luminance profile of LSB galaxies are known, thus we can easily define a set of templates representing the various visual appearances they can take. However, their detection is a hard task as astronomical images have very heterogeneous content. The pixel dynamic can be very large with a brightness ratio between faintest and brightest objects coming up to 1:1000. The size variability of objects present in astronomical images can also be huge with objects

* Corresponding author.

E-mail addresses: perret@lsiit.u-strasbg.fr (B. Perret), lefevre@lsiit.u-strasbg.fr (S. Lefèvre), collet@lsiit.u-strasbg.fr (Ch. Collet).

covering only a few pixels to the whole image. Moreover, due to projective effects, different objects may overlap. Finally, to observe such faint objects as LSB galaxies, astronomers have to use long exposition time to collect as many photons as possible, thus leading to also collect a lot of noise.

This paper is organised as follows. First we briefly recall the basics of grey level morphology. Then we discuss the existing approaches for grey level HMT and their relative robustness to noise. We introduce a new formulation of the fuzzy HMT and give a methodology to involve this formulation in template matching applied to the detection of LSB galaxies in astronomical images. Finally results of this application are discussed and a comparison with existing astronomical results is proposed.

2. Review of existing HMT

We first recall the basics of grey level morphology, using the notations given by Naegel et al. [10]. Let E be a digital space (e.g. $E = \mathbb{Z}^n$). Let T be a set of grey levels. We require T to be a complete lattice with respect to the order " $\leq$ " [11,12]. We will next assume that $T = \mathbb{R} \cup \{+\infty, -\infty\}$ or $T = \mathbb{Z} \cup \{+\infty, -\infty\}$. We denote $\perp$ and $\top$, respectively, the lowest and the greatest elements of T .

Let I and G be two functions of T^E , the set of functions going from a subspace of E to T . We call the support of I ($\text{supp}(I)$) and dual support ($\text{supp}^*(I)$) the set of points where I is strictly above $\perp$ and, respectively, under $\top$. Dilation $\oplus$ and erosion $\ominus$ of I by G at every point $p \in E$ are then given by

$$(I \oplus G)(p) = \sup_{k \in \text{supp}(G)} (I(p - k) + G(k)) \quad (1)$$

$$(I \ominus G)(p) = \inf_{k \in \text{supp}(G)} (I(p + k) - G(k)) \quad (2)$$

You can notice that these formulas can lead to disinclination like $+\infty$ plus $-\infty$. To keep consistency $+\infty$ plus $-\infty$ must be valued as $-\infty$ in Eq. (1) and as $+\infty$ in Eq. (2). Further readings about the grey level morphology can be found in Refs. [13,14]. We also define the dual G^* of the function $G \in T^E$ by

$$G^* : E \rightarrow T$$

$$p \rightarrow -G(-p) \quad (3)$$

We now recall the principles of the HMT. It consists of searching for all pixels in an image which have a desired neighbourhood. It requires two SEs: the first one called *foreground* SE describes the shape and the second one called *background* SE is used to describe the spatial neighbourhood of this shape. In the binary case the interpretation of background and foreground is straightforward: the foreground is a set of pixels which has to be present in the image and the background is a set of pixels which has to be present in the complement of the image (see e.g. Fig. 1).

The extension of HMT to complete lattice requires to define *foreground* and *background* in such a general case. Thus several approaches have been proposed [15,10,3,16]. We will now present these approaches and discuss their resistance to noise in images with very low SNR ($\text{SNR} \leq 0$ db), the SNR being defined as signal energy divided by noise energy and is expressed in logarithmic scale

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_n^2} \right) \text{ db}$$

with σ_s and σ_n being, respectively, the standard deviation of signal and noise.

2.1. Existing formulations of grey level HMT

We now introduce the various existing approaches to grey level HMT. For each operator, we discuss its ability to deal with poor SNR which is of primary importance for template matching in noisy images.

Let $I, F, B \in T^E$, F and B are the foreground and background SEs such that $F \leq B$ (i.e. $\forall p \in E, F(p) \leq B(p)$), I stands for the image.

2.1.1. Ronse's and Soille's HMT

The definition of Ronse's HMT given in Ref. [10] is

$$\text{RHMT}_{[F,B]}(I)(p) = \begin{cases} (I \ominus F)(p) & \text{if } (I \ominus F)(p) \geq (I \oplus B^*)(p) \neq +\infty \\ -\infty & \text{otherwise} \end{cases} \quad (4)$$

It assigns the maximum grey value where it is possible to fit both SEs.

Soille's definition given in Ref. [10] is an extension to functional SE of the initial version proposed in Ref. [1]

$$\text{SHMT}_{[F,B]}(I)(p) = \max\{(I \ominus F)(p) - (I \oplus B^*)(p), 0\} \quad (5)$$

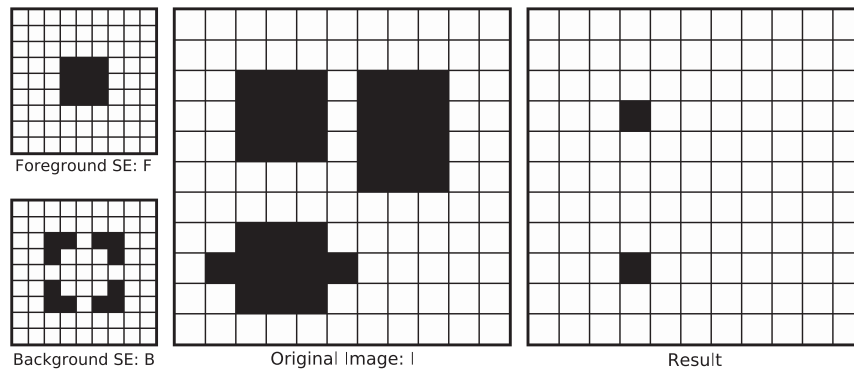


Fig. 1. Example of application of binary HMT to detect 3×3 squares with possible extension along edges. The top left image represents the foreground SE: F , origin of the image is at the centre of the square. The bottom left image is the background SE B , with same origin as F . The middle image is the original image I . And the right image is the result of the HMT applied on I with SEs F and B .

We can observe that Soille's and Ronse's approaches are very similar. Soille's definition returns the difference between the highest and the lowest values where it is possible to fit the SEs. A unified formulation of these two approaches has been proposed in Ref. [10]. These approaches are adapted to perform pattern matching in quite good SNR environments, as we will discuss in the next section.

2.1.2. Barat et al.'s HMT

Barat et al. introduced in Ref. [3] an operation called *morphological probing* that is defined by

$$BHMT_{[F,B]}(I)(p) = (I \oplus B^*)(p) - (I \ominus F)(p) \quad (6)$$

It measures the distance between the best inner fit of F and the best outer fit of B . The lower is this value the best is the fit. This method presents no special noise related issue. One can note that there is a direct link between Barat's and Soille's HMT:

$$SHMT_{[F,B]}(I)(p) = -\min\{BHMT_{[F,B]}(I)(p), 0\} \quad (7)$$

thus Soille's HMT can be seen as a constrained version of Barat's HMT.

2.1.3. Khosravi-Schaefer's HMT

Khosravi and Schaefer introduced in Ref. [16] another version of HMT which uses a single SE and is equivalent to Barat et al.'s [3] HMT. It is defined by

$$\begin{aligned} KHMT_{[F]}(I)(p) &= (I \ominus F)(p) - (-I \ominus (-F))(p) \\ &= (I \ominus F)(p) - (I \oplus F^*)(p) \\ &= -BHMT_{[F,F]}(I)(p) \end{aligned} \quad (8)$$

$KHMT$ is always less or equal to 0 which is reached for a perfect match.

2.2. Generic solution to improve noise robustness

We place this study in the context of very noisy images, considering that images are generally affected by impulsive (or speckle), Gaussian or Poisson noise, we will first review the different generic solutions that have been proposed to deal with these kinds of noise. Another aspect of noise is the inherent slight variations or uncertainties of a real signal to the ideal model. These variations can be of different types: deformations (local or global), overlapping with other objects.

A simple solution to improve the tolerance to impulsive noise of all methods is to use rank operators instead of traditional erosion and dilation [16,17]. However, this solution implies to determine an appropriate rank for both dilation and erosion operations. Khosravi and Schaefer [16] have shown that a lower and upper bounds for the rank are given, respectively, by impulsive and Gaussian noise, but they did not provide a general formula to obtain these bounds. They also concluded that the effect of rank operator on Gaussian or Poisson noise is limited. Moreover impulsive noise can be strongly reduced by applying a rank filter (e.g. a median filter) before the pattern matching operator with negligible loss of quality [18].

Bloomberg and Maragos [17] proposed several solutions to improve robustness to noise and shape uncertainties. First idea is to perform a subsampling of both image and SEs. This solution is attractive (by considering less pixels and a Gaussian white noise, the probability of having a very noisy pixel under the SE is decreased),

it reduces small shape variations and saves computation time. Nevertheless it requires the resolution to be good enough so the subsampling will preserve major features of the shape. Another possibility is to perform a spatial decimation of the support of structural functions following a regular grid. The advantages of this method are the same as subsampling's ones, but it can be applied even at low resolution. Despite their capacities, these two methods are not sufficient to deal with very noisy images.

It is also possible to improve noise resistance of the morphological HMT by first dilating both foreground and background before performing erosions, as proposed by Bloomberg [17] in the binary case. In this case the foreground and the background overlap, thus it becomes easier to fit the SEs. However, the extension to the greyscale case is not trivial since we have to define the functional SE to be used.

Another way to improve robustness against noise and shape uncertainties is to provide less informative SEs (i.e. increase the distance between the foreground and the background). This is nearly equivalent to previous Bloomberg's method and can be the easiest way to work with low level SNR. Nevertheless this method fails at very low SNR as the SEs become so blurred that almost every configurations can provide a good match.

3. Robust hit-or-miss

Khosravi and Schaefer [16] have widely tested the use of KHMT with general rank order operators face to impulsive and Gaussian noise at various SNR. They explain the difficulty to find a suitable rank which is highly dependent on noise level and type. They have also explored partitioning of the template which is a kind of variation of the decimation method. Their conclusion joined Bloomberg's one saying that it should be done following a regular grid to obtain best performances. Nevertheless with all these improvements, performances still fall when the SNR goes under 19 dB [16].

The intrinsic difficulty of Soille's and Ronse's HMT to deal with high noise level is due to their rigid behaviour. Because we have no intermediate level between a true answer (object is detected at point p) and a false one (object is not detected) it is really hard to find the good SEs F and B to be used. In fact, in these HMT the main feature that provides a robustness to noise is the distance taken between F and B , the larger it is the more noise they can absorb and inversely. If F and B are too precise, too close then the transform nearly always gives a negative result. On the contrary, if F and B are too simple and the distance between each other is too large, this will lead to a high number of false positives.

Barat et al.'s HMT or Khosravi-Schaefer's HMT are quite different because they evaluate a distance, thus they look more flexible. Let us consider a given pixel expected to return a positive result (ground truth is available). However, applying Barat et al.'s HMT with the SEs F and B on this pixel may return a positive value t instead of a negative one. Then to ensure a positive result with Ronse's or Soille's HMT at this point we have to increase the distance between the two SEs by t (e.g. $F' = F - t$ and $B' = B$ or $F' = F$ and $B' = B' + t$).

To sum up, Soille's and Ronse's HMT noise robustness is mainly determined by the distance between the two SEs. In Barat et al.'s HMT or Khosravi-Schaefer's HMT the distance between the two SEs is not important as it only results in a shift of the result. Thus the problem of determining a good distance is transformed in a problem of thresholding. The important point is to understand that these problems are equivalent and that all these formulations will react similarly when a lot of noise is present. None of them will deal with this case accurately. Fig. 2 illustrates the behaviour of the various HMTs when SNR is decreasing.

In the binary case Maragos [19] had the idea to introduce a fuzzy HMT that allows to keep well-defined SEs, close to each other and that works in noisy environments. It measures a ratio between the

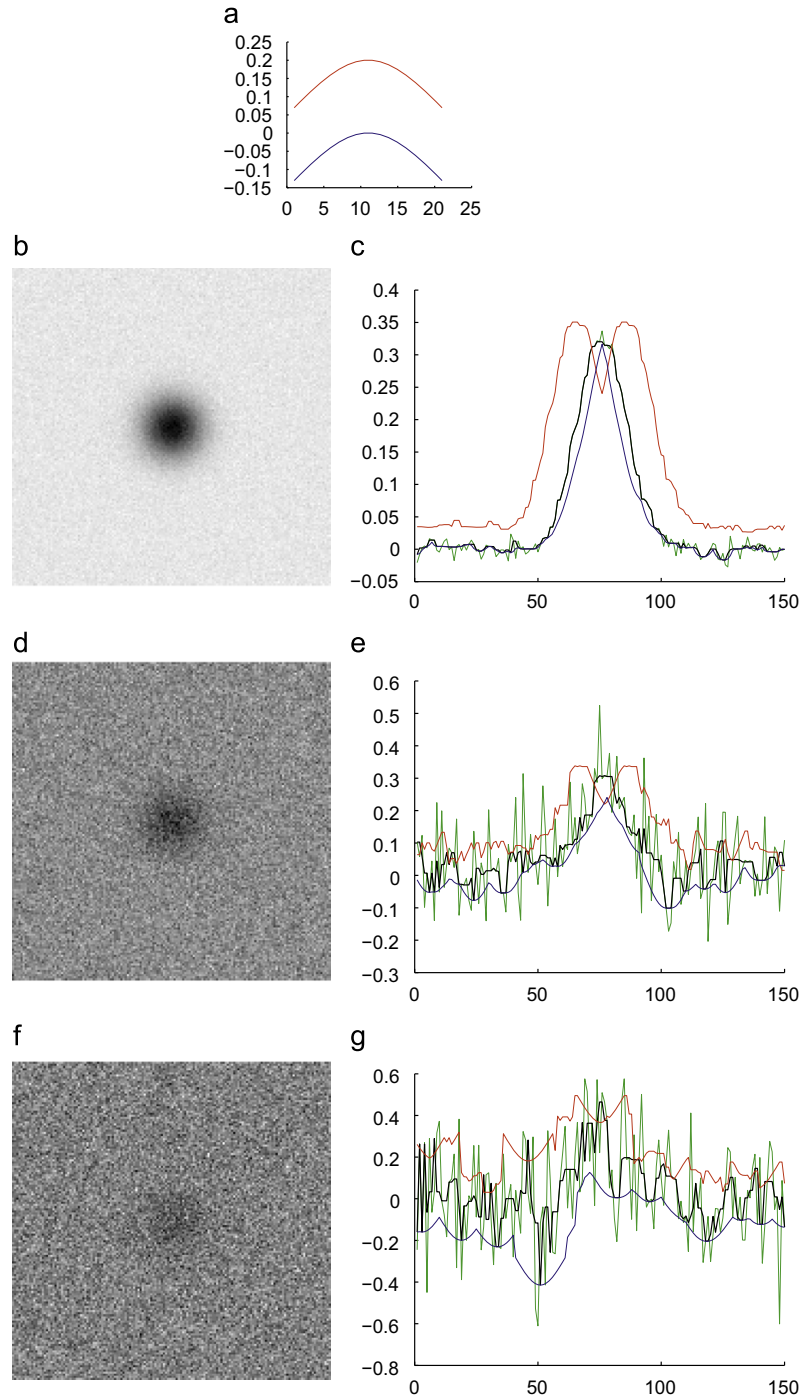


Fig. 2. This shows how the grey dilation and erosion and consequently the various formulations of HMT react when SNR falls. In this example a round shape with a Gaussian luminosity profile is hidden in noise. Image (a) is a one-dimensional cut of the foreground SE F (blue curve) and the background SE B (red curve). Then the left column shows images with decreasing SNR: ~ 20 dB for image (b), ~ 0 dB for image (d) and ~ -8 dB for image (f). Right column represents a one-dimensional cut of left images (in green), a cut of a median filtered (size 5×5) version of left image (black), the erosion of the filtered image by the foreground SE: $I \ominus F$ (in blue), and the dilation of the filtered image by the dual of the background SE $I \oplus B^*$ (in red). In plots (c) and (e), the red curve goes under the blue one, which would give a positive result for RHMT and SHMT. In the third case detection fails as the dilation and the erosion are disturbed by noise. One can note that if we increase the distance between SE F and B , for example, by defining $F' = F - k$ with k a grey value, the result would be equivalent to a vertical translation of the blue curve. And this is obviously useless as no vertical translation of the blue curve can make it cross the red curve only at the centre (i.e. if the blue line crosses the red one at another point this means that a false positive is detected). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

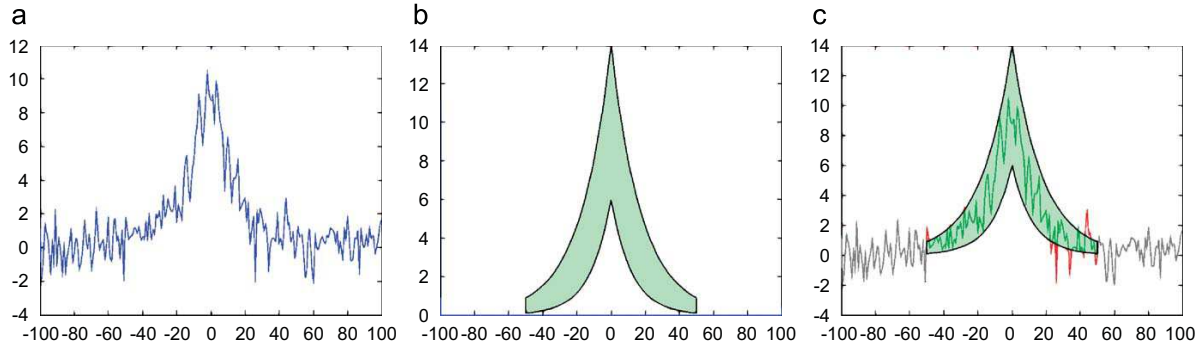


Fig. 3. Example of application of the fuzzy HMT to detect an exponential profile with Gaussian noise. Image (a) represents a 1D noisy signal (SNR $\approx$ 9 dB). Image (b) represents the uncertainty area defined by the 2 SEs F and B . Image (c) shows how well the pattern can fit the signal.

number of pixels that actually fit in the SEs and the total number of pixels contained in the SEs. Hence, instead of looking for the positions where the two SEs perfectly fit the image, we measure how well the SEs fit the image everywhere.

This can be extended to grey level images. For this we consider that the foreground and the background SEs describe, respectively, a local lower and upper bound for our image. The fitting process measures the ratio between the number of pixels in the neighbourhood between these two bounds and the total number of pixels in the area covered by the SEs. The result is a score between 0 (no points of the neighbourhood fits in the bounds) and 1 (all pixels of the image are between the lower and the upper bounds).

Whereas in Ronse's or Soille's approach the distance between SEs F and B and so their global shape has to be modified to tolerate noise, in this approach the SEs can keep a well-defined shape and tolerate fluctuations due to noise.

3.1. Definition

We now assume that T is an infinite set of grey values with a total ordering relation noted $\leq$ (e.g. $T \subset \mathbb{Z}$). Formally we have to decompose the image I and the pair of SEs (F, B) as *suprema* of impulses and then consider each triplet (I, F, B) of impulses independently. Let $i_{(p,t)}$ be an impulse at point $p \in E$ of level $t \in T$

$$\forall x \in E, \quad i_{(p,t)}(x) = \begin{cases} t & \text{if } p = x \\ -\infty & \text{otherwise} \end{cases} \quad (9)$$

For practical reasons and because the positions of the impulses have no influence on the following definitions we write i_t for $i_{(0,t)}$. Let us write $IC_{F,B}^t(q)(v)$ the proposition "the pixel value v comprised between $F+t$ and $B+t$ at point q in the SE", we have

$$IC_{F,B}^t(q)(v) \iff i_{F(q)+t} \leq v \leq i_{B(q)+t} \quad (10)$$

$$\iff i_v \oplus i_{B(q)+t} \leq t \leq i_v \ominus i_{F(q)+t} \quad (11)$$

A proof of equivalence between Eqs. (10) and (11) can be found in Ref. [10].

Let us define $S = \text{supp}(F) \cup \text{supp}^*(B)$ and card is the cardinality of a set. The proposed fuzzy HMT is then defined by

$$FHMT_{[F,B]}(I)(p) = \max_{t \in T} \frac{\text{card}\{q \in S | IC_{F,B}^t(q)(I(p+q))\}}{\text{card}(S)} \quad (12)$$

Hence it evaluates the best ratio of pixels satisfying the requirements of Proposition 11 with SEs F and B translated at all possible grey levels.

3.2. Implementation issues and parameters selection

Application of this operator is not straightforward. Indeed T is theoretically an infinite and perhaps continuous set, thus it is not directly possible to determine at which level t the score in Eq. (12) will be maximised. In fact the interval of possible values for t can be restricted quite easily by a local analysis of pixel values. For example, in Fig. 3, we can estimate that the upper part of our pattern should be between grey levels 6 and 10, and this can obviously be done by applying a simple mean filter and keep the resulting value as an estimation of t . One can then compute the score for different values near this estimation. This solution obviously depends on the possibility to extract a main feature from the pattern (for example "the peak" in Fig. 3) that can be used to perform a first and quick estimation of t .

Another issue is how to determine the ideal distance between the two SEs. Indeed it can be estimated from local noise information. For example, assuming that our image is corrupted by a Gaussian noise $\mathcal{N}(0, \sigma^2)$, and that we are looking for a pattern given by function P we can set $F = P - \sigma$ and $B = P + \sigma$. Then, according to noise statistic we can expect to have a matching score of about 68% in a correct case.

4. Application to astronomical images

Since the discovery of the various surface brightnesses of galaxies, astronomers have become more and more interested in a way to automatically detect very faint objects in astronomical images. LSB galaxies [9] have a central surface brightness higher than 22.5 mag/arcsec² (the magnitude is an inverse logarithmic brightness scale used in astronomy because of its similarities with the human eye behaviour, the arcsec is an angular distance). Traditional method of sigma clipping (i.e. adaptive thresholding method) used commonly in astronomy is not able to detect such objects as they are eliminated with background (see Section 4.2 for details on the sigma clipping method).

When searching for objects with a good SNR, astronomical images are quite easy to describe because they are mainly composed of bright objects on a dark background. Unfortunately several problems arise when approaching background signal level (i.e. looking at the early universe or at low brightness objects). The background is not flat over all the images, even in small images of about 512×512 pixels. Images are also affected by a Poisson noise (approximated by a Gaussian noise of mean 0 and deviation σ) and some salt and pepper noise. This last one may be caused by defect pixels on the CCD camera or more usually by cosmic rays. All these properties make the problem of LSB galaxies detection hard to solve (see Fig. 6).

The whole process of LSB galaxies detection is decomposed in two steps: first a segmentation map of interesting objects is built, and then the objects present in segmentation map are characterised using physical criteria. Astronomers of the Centre de Données astronomiques de Strasbourg¹ (CDS) have recently developed such a tool that performs an automatic characterisation of astronomical objects [20]. This tool has been tested to detect LSB galaxies based on segmentation map obtained using a Markovian quadtree segmentation tool [21]. The Bayesian approach is useful to involve a prior on noise but does not deal with large shape knowledge, thus we proposed a new method to compute more selective segmentation map using FHMT.

In this section we describe a method based on FHMT to automatically build a segmentation map of potential LSB galaxies. The algorithm is decomposed into several steps. First we need to build several patterns corresponding to LSB galaxies of various shapes and orientations. Because the objects we are looking for are very close to the background in terms of photometry, we also need to compute a precise map of the background (i.e. evaluate the intrinsic luminosity of the sky at all points). Next the original image is preprocessed with a median filter to reduce noise. Filtered image, background map, and pattern set are then used to calculate FHMT. The result is thresholded and the original shape of LSB galaxies is reconstructed.

4.1. Description of patterns

We now describe how LSB galaxies are modelled. They are elliptic objects with a luminosity profile roughly following an exponential law. Let us consider a two-dimensional discrete plane E . Let $(c_x, c_y) \in E$ be the centre of the galaxy, then we have for all $p \in E$ the following expression for the luminance:

$$LSB(p_x, p_y) = \mu_0 e^{-1/R_0 \sqrt{(p_x - c_x)^2 + b^2(p_y - c_y)^2}} \quad (13)$$

where μ_0 is the central surface brightness given in pixel value unit, R_0 is the scale length, $b \in]0; 1]$ defines the ellipticity of the model. Using this model we are able to generate several patterns with various brightnesses, elongations, orientations and scale lengths.

Because of the non-constant background and the huge number of possible central brightnesses we choose to generate patterns at a normalised brightness ($\mu_0 = 1$) and dynamically stretch patterns depending on local background value and noise statistics. Since there is a high correlation between the noise standard deviation σ_n and magnitude scale, brightness t will then be expressed as a factor k of the noise deviation: $t = k * \sigma_n + \mu_b$ with μ_b the background level.

We considered scale length from 1.33 to 10.33 arcsec by step of 1.0 arcsec (the equivalence between arcsec and pixel unit can be computed from image file header). We quantify 14 orientations between 0 and π . The final set of SEs was obtained by combining all possible scale lengths, elongations and orientations while avoiding identical cases induced by symmetries. It is composed of 640 templates (see e.g.² Figs. 4 and 5).

All SEs are then convolved with a Gaussian kernel of 5×5 pixels to reproduce the point spread function (PSF) of the telescope (seeing and optics deformations). This is an approximation because PSF depends on the image and the location in the image, but this approximation is negligible in comparison with shape uncertainties.

Finally each pattern is composed of two SEs of same orientation and elongation. We choose to take the foreground SE to have a scale length of 1 arcsec smaller than the background one. The distance



Fig. 4. Example of SEs obtained with a variation of orientations convolved with a Gaussian point spread function.



Fig. 5. Example of SEs obtained with a variation of elongations convolved with a Gaussian point spread function.

between the SEs is adjusted dynamically with respect to local noise statistics (see Section 4.2.).

4.2. Computation of FHMT

Once we have generated all different patterns, we estimate background and noise of the image by following a regular grid of 32×32 pixels in a window of 256×256 pixels with the sigma clipping method widely used in astronomical imaging. This method performs in two steps: in the first one it computes the average value μ (mean or median) and deviation σ of pixels in the estimation window, then in the second step it masks all pixels for which the value is greater (resp. lower) than $\mu + k\sigma$ (resp. $\mu - k\sigma$), and goes back to step one until convergence. Parameter k gives the confidence interval with respect to noise power, a value of 3 is generally used.³ Background and noise estimation cannot be performed over all pixels because of its prohibitive computation time. Nevertheless as these parameters

¹ <http://cdsweb.u-strasbg.fr/>

² All images are given in inverse grey level to save ink.

³ See SExtractor documentation for details: <http://terapix.iap.fr/>

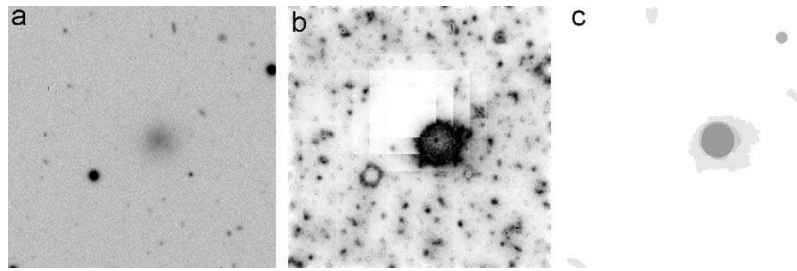


Fig. 6. Application of the FHMT to LSB galaxy detection, an easy case. (a) Original image of 512×512 pixels with a LSB galaxy at centre. One can note that the pixel dynamic has been adjusted manually so the LSB galaxy becomes visible. (b) Score map obtained after the application of the FHMT. The LSB galaxy is clearly visible. We can observe that both the centre and the wings of the object give a good score. (c) Thresholded and reconstructed map. The final map contains different classes proportional to the object brightness. This allows deblending capabilities for overlapping objects having different brightnesses.

vary slowly over the image, we can have a good estimation of them by simply interpolating values of the background grid.

Computation of scores is done according to the principles given in Section 3.2. Estimation of the best grey level t for the vertical translation of the SEs is done using a median filter (5×5). We transpose foreground (resp. background) SE to local background level and stretch it so that its highest value is $t - \sigma$ (resp. $t + \sigma$) with σ the local deviation of noise. Matching scores between each pattern and median filtered image is then computed and the best one is kept as the result (Figs. 6 and 10).

4.3. Reconstruction of score map

In the next step, the score map is thresholded. The threshold value offers a good tuning option and was determined to minimise false negative rate and, with less importance, minimise false positive rate. We found that a score of 80% offers a good compromise. The use of the same threshold for every observation is not a problem. As all parameters of the algorithm are set automatically according to observation parameters and statistics, the score gives an absolute measure which is independent of the observation.

Once we have a binary map, we need to reconstruct the final map. To do this, each pixel of the binary map is dilated by the support of the pattern that gives the maximum score in this position (Figs. 6 and 10). Scheme of the procedure is given in Fig. 7.

To improve map's precision we could have used a geodesic reconstruction instead of a simple dilation so the reconstructed areas would exactly fit the detected objects. In our case this is useless because the next step of the algorithm rather prefers to work with idealised shapes instead of precise but noisy boundaries.

4.4. Optimisation

Without any optimisation the proposed algorithm takes about 20 min to compute the segmentation map on an image of 512×512 pixels. This time includes estimation of background and generation of the 640 patterns. The code is written in Java and executed on a Pentium IV computer. In this section we propose several optimisations to improve algorithm performances:

- We perform a spatial decimation of patterns using Bloomberg's method [17]; we choose to keep one pixel every two.
- As the resolution is sufficient in our images, we perform segmentation in a sub-grid of the images at scale 1:2.
- Score computation and thresholding are done simultaneously, so the computation of a fit can be discarded as soon as threshold becomes unreachable.
- Last optimisation is heuristic and comes from the idea that we

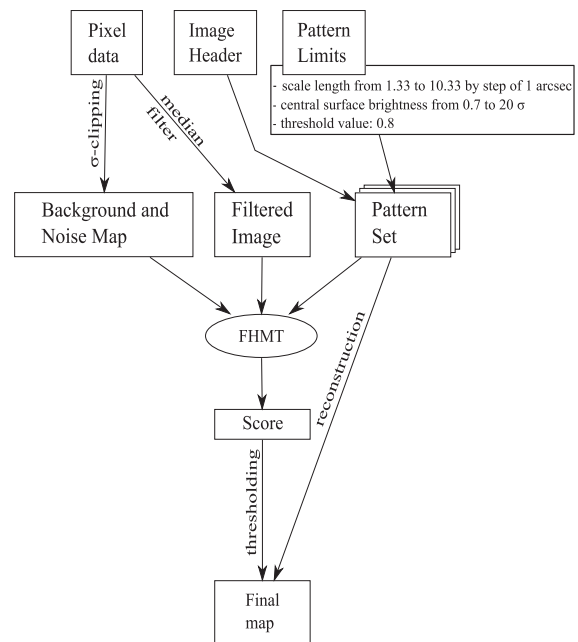


Fig. 7. Scheme of the pattern matching method. The algorithm takes three inputs: the image pixel values, the image file header, and physical limits. The whole procedure is automatic.

can perform the segmentation in two steps. The first step will make a fast selection of possibly interesting pixels, second step will perform a full computation for selected pixels. During the first step we will segment the images using only a few patterns (we take only circular patterns, orientation and elongation are not considered) and a low threshold (60%). During the second step we segment pixels selected at the first step with the full process (all patterns, threshold at 80%).

With all these optimisations we are able to compute a segmentation map of a 512×512 images in about 2 min, which is ten times faster than original algorithm. Detection maps obtained with and without optimisation are nearly the same. One can notice that last optimisation is purely heuristic and so we cannot assert that target will never be lost. However, we have compared results obtained with both the original strategy and this 2-step approach, and we have

observed only minor differences (small difference in the shape or size of the detected targets). In these experiments no target was lost.

5. Results

In this section, the method was evaluated through two different datasets. First, to compensate the lack of ground truth in real images, we created a simulated dataset. However, we also involved a set of real images which have already been analysed by experts [22] to show the relevance of the proposed approach in this context. Due to the lack of related work and the clear inability of existing HMT to deal with this problem (as shown in Section 3), evaluation can rely only on quantitative and qualitative analysis (and thus cannot involve any comparison).

5.1. Simulated dataset

It is very important for astronomers to know the detection limits of the algorithm they are using. This information is used to determine the statistical bias of their conclusions due to selection effects of the algorithm.

A batch of tests has been designed with astronomers of the CDS to determine the detection limits of our algorithm in terms of brightness and scale length. These tests apply on simulated images of LSB galaxies for which we know the ground truth. The algorithm used to generate artificial LSB galaxies is nearly the same as the one used to build the SEs in the detection algorithm. We used a Gaussian function to reproduce PSF. We also had the possibility to add other astronomical objects in our images (like stars or bright galaxies) to test robustness of our algorithm in crowded part of the sky. Noise power and pixels statistics used to create images were taken directly from a real astronomical image. Fig. 8 shows the examples of simulated images of LSB galaxies.

Our test protocol is divided into three parts (Fig. 8), and we applied the algorithm on:

- images containing only one LSB galaxy (one test per couple (brightness, radius)) to determine absolute limits;
- images containing one LSB galaxy and a bright star (one test per couple (brightness, radius)) to study reaction in a sparse environment;
- images containing one LSB galaxy and between three and seven other objects located randomly over the images (21 tests per couple (brightness, radius)) to test robustness in crowded environment.

The LSB galaxies do not have strict boundaries, thus we choose to define their limits as the ellipse where the signal has lost 95% of its central energy. The area enclosed in this radius gives us the reference truth to evaluate the map proposed by the detection algorithm. Let C_{ref} be the set of pixels present in the reference map. There may be several objects in the segmentation map provided by the method, thus we choose to keep the connected component having the largest intersection with the reference map as the final result and denote it by C .

Finally we used an F1-measure [23] to evaluate the correctness of the detection. The later combines the *precision* and the *recall* measures to obtain a score between 0 and 1. It is given by

$$F1\text{-measure} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (14)$$

with *precision* defined as

$$\text{precision} = \frac{\text{card}(C \cap C_{ref})}{\text{card}(C)} \quad (15)$$

and *recall* defined as

$$\text{recall} = \frac{\text{card}(C \cap C_{ref})}{\text{card}(C_{ref})} \quad (16)$$

Experiments have shown that an F1-measure value equal to at least 0.4 is sufficient for the following steps to perform accurately.

We have determined that the FHMT algorithm is able to detect LSB galaxy up to a scale length of 2 arcsec (this is PSF limit anyway) and a central surface brightness of $\sim 26.2 \text{ mag/arcsec}^2$ (0.7σ over sky background level which is nearly equivalent to 0 dB SNR). Moreover detection is nearly not affected by other objects except if the wing of another bright object covers the LSB galaxy (the two objects overlap and the LSB galaxy is drowned in the other object luminosity).

We have also compared the results of our method to those obtained with a traditional correlation method [18]. This was done by replacing the FHMT measure by the correlation one and adjusting the threshold value in the process to ensure the best possible results. The correlation based method performs well until 1.5σ over sky background level (Fig. 9), i.e. its robustness to noise, is lower than the FHMT one. This shows that (robust) morphological approaches can be relevant alternatives to statistical methods even in such noisy environments.

5.2. Real dataset

Next the FHMT operator was applied on a test set of 16 blue band (a filter centred at wavelength of 4500 Å corresponding to blue colour is used) images of 2048×4096 pixels coming from the INT Wide Angle Survey.⁴ The dataset covers the Virgo cluster which has the advantage to be close to the Galaxy and in a relatively sparse area of the sky. Moreover the population of LSB galaxies in this survey has already been studied by astronomers [22], which gives us a reference catalogue to evaluate our results.

The algorithm has been run over the whole dataset and segmentation maps were used as inputs for the characterisation tool of the CDS: DetectLSB. At the end of the process, this tool provides HTML files describing potential LSB galaxies found and XML catalogues in the VOTable standard [20].

A full analysis of the result for two images was performed in collaboration with an astronomer. The reference catalogue contains nine LSB galaxies in these areas. The algorithm proposed 23 candidates. It found six objects of the reference catalogue and provided eight new LSB galaxies. The spurious detections were mainly due to artefacts created by two close sources (when wings of two close objects intersect, they locally create an artificial raise of luminosity) or very distant galaxies (the brightness of the galaxy is then decreased by the absorption of intergalactic environment). This leads to a total of 17 interesting objects (nine already known and eight newly discovered) for the two images. The recall of the method is then 0.82 and the precision is 0.6. One can note that the three LSB galaxies missed from the reference catalogue were correctly segmented by FHMT but were rejected by the characterisation algorithm. Because LSB galaxies detection is a very challenging task for which no automatic solution has been proposed yet, these scores can be considered as really good.

The traditional correlation method performed on these two images, using the same methodology as in Section 5.1, proposed 54 candidates. It found 9 of the 17 LSB galaxies previously mentioned, yielding a recovery rate equal to 53%. The false positives were mainly multiple detections (two or more close objects are fused in a single one in the final map) or very distant galaxies. Fig. 10 shows a

⁴ http://www.ast.cam.ac.uk/~rgm/int_sur/

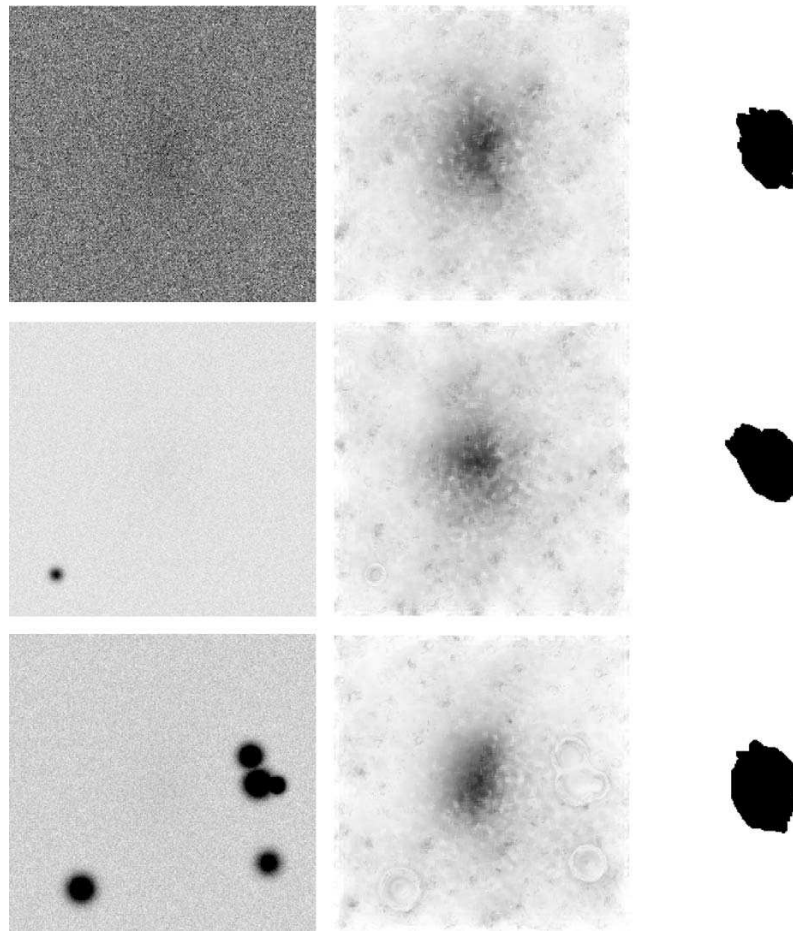


Fig. 8. Example of simulated images for the three steps of our evaluation protocol. The LSB galaxies displayed in the centre of each image have a scale length of 10 pixels and a central surface brightness of $25.7 \text{ mag/arcsec}^2$ (one σ over sky background level with σ the deviation of noise). The SNR is very low about -1 dB . The first line shows the results for a single LSB galaxy, the second line for a LSB galaxy with a star near it, and the third line for a LSB galaxy surrounded by five stars randomly placed. The first column shows the simulated image, the second one shows the result of FHMT, and the third one the final map.

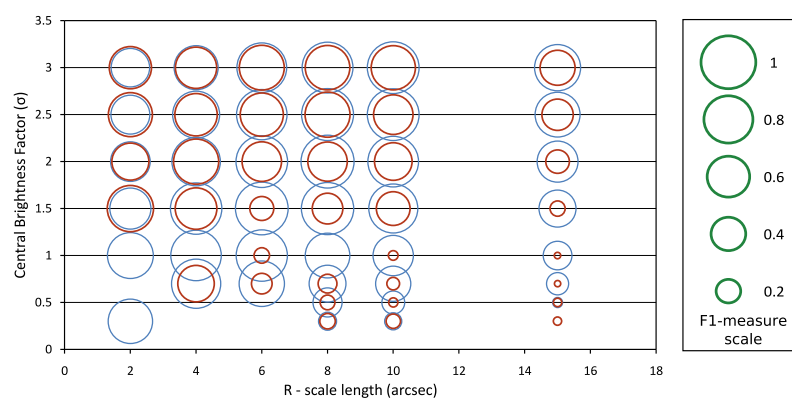


Fig. 9. Comparison of the results obtained using the FHMT (blue circles) and the correlation (red circles) on simulated LSB galaxies. For each couple (radius, brightness factor), the circle area is proportional to the F1-measure value. The green circles on the right side give a reference scale. The experiments showed that the detection can be considered to be effective when the F1-measure is above or equal to 0.4. While the performances of the correlation drops under 1.5σ , the FHMT remains effective until 0.5σ over sky background level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

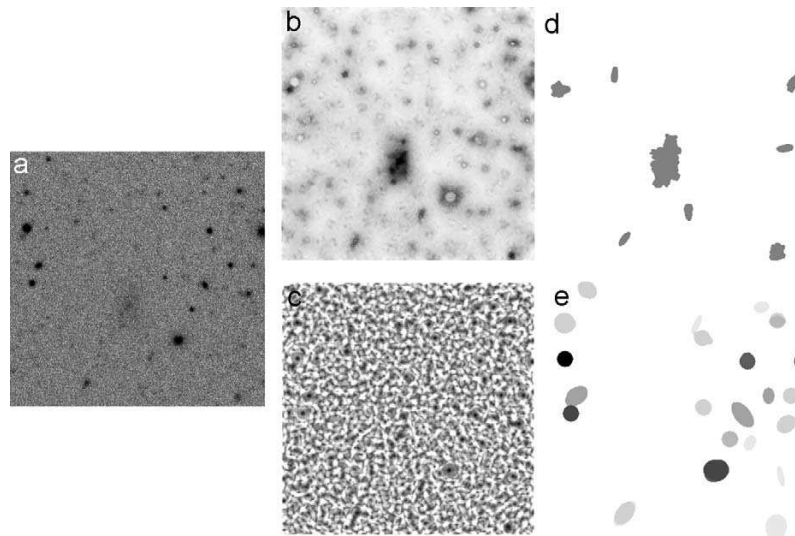


Fig. 10. Example of comparison between the FHMT and the traditional correlation for LSB galaxies detection. (a) Original image of 512×512 pixels with a LSB galaxy at the centre. One can note that the pixel dynamic has been adjusted manually: the LSB galaxy is still hard to view. (b) Score map obtained after the application of the FHMT. The LSB is clearly visible. We can observe that the object appears to be composed of several small ellipses, which is due to the irregular aspect of the galaxy. (c) Thresholded and reconstructed map for the FHMT. There is clearly one false positive on this map on the top right border (it corresponds to the wings of a brighter object). Other objects detected have to be deeply analysed to determine their physical nature. (d) Score map obtained after the application of the correlation. (e) Thresholded and reconstructed map for the correlation. In spite of the high number of false positives, the main target is lost.

comparison between results given by the FHMT and the correlation on a real small image extracted from one of the two big images.

We also applied the FHMT to the 16 images from the real dataset. We performed a cross correlation between the new catalogue obtained with the FHMT and the existing one: the FHMT was able to recover 87% of the LSB galaxies contained in the reference catalogue and several new candidates were proposed. These candidates have to be analysed now by astronomers' community before any conclusion can be drawn on them.

6. Conclusion

We have addressed the particular problem of template matching in very noisy environment using the morphological HMT. We have shown that the existing formulations of HMT are not adapted to do this, thus we proposed a formulation of a "fuzzy" HMT. It relies on two SEs: the foreground and the background one and consists of measuring the ratio of the image area that fit the SEs compared to the total area covered by the SEs.

We have described how it can be applied to a real and still unsolved case in astronomical imaging: the automatic detection of LSB galaxies. We have tested the new FHMT on simulated and real images and showed that our method performs better than a traditional correlation method. Our results have been analysed by an astronomer and confirmed that FHMT is suitable for pattern matching at very low SNR.

Thereby the field of application of the morphological HMT has been extended to very noisy images for which statistical approaches are generally preferred. The process we described can be easily adapted for pattern matching in other domains such as radar imaging. Moreover, in case of noisy multispectral data, the proposed operator should be extended to multivariate mathematical morphology [24].

Acknowledgements

The authors thank Bernd Vollmer (Observatoire de Strasbourg, Université de Strasbourg), François Bonnarel (Centre de Données astronomiques de Strasbourg, Université de Strasbourg), Mireille Louys (Laboratoire des Sciences de l'Images, de l'Informatique et de la Télédétection, Université de Strasbourg) and Wim Van Driel (Observatoire de Meudon, Institut d'Astrophysique de Paris) for their enlightening discussion about LSB galaxies and the various aspects of astronomical image analysis.

References

- [1] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer, New York, Inc., Secaucus, NJ, USA, 2003.
- [2] B. Naegel, N. Passat, C. Ronse, Grey-level hit-or-miss transforms—part ii: application to angiographic image processing, *Pattern Recognition* 40 (2) (2007) 648–658.
- [3] C. Barat, C. Ducottet, M. Jourlin, Pattern matching using morphological probing, in: *International Conference on Image Processing*, vol. 1, 2003, pp. 369–372.
- [4] S. Lefèvre, J. Weber, D. Sheeren, Automatic building extraction in VHR images using advanced morphological operators, in: *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (URBAN)*, Paris, 2007.
- [5] A. Puissant, J. Weber, S. Lefèvre, Coastline extraction in VHR imagery using mathematical morphology with spatial and spectral knowledge, in: *ISPRS International Congress*, Beijing, 2008.
- [6] D.S. Bloomberg, L.M. Vincent, Pattern matching using the blur hit-miss transform, *Journal of Electronic Imaging* 9 (2000) 140–150.
- [7] B. Raducanu, M. Graña, A grayscale hit-or-miss transform based on level sets, in: *International Conference on Image Processing*, 2000, pp. 931–933.
- [8] E. Aptoula, S. Lefèvre, C. Collet, Mathematical morphology applied to the segmentation and classification of galaxies in multispectral images, in: *European Signal Processing Conference (EUSIPCO)*, 2006.
- [9] G. Bothun, C. Impey, S. McGaugh, Low-surface-brightness galaxies: hidden galaxies revealed, *Publication of the Astronomical Society of the Pacific* 109 (1997) 745–758.
- [10] B. Naegel, N. Passat, C. Ronse, Grey-level hit-or-miss transforms—part i: unified theory, *Pattern Recognition* 40 (2) (2007) 635–647.
- [11] H.J.A.M. Heijmans, Theoretical aspects of gray-level morphology, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 13 (6) (1991) 568–582.

- [12] C. Ronse, Why mathematical morphology needs complete lattices, *Signal Processing* 21 (2) (1990) 129–154.
- [13] J. Serra, Morphological image operators, *Society for Industrial and Applied Mathematics Review* 38 (1) (1996) 178–179.
- [14] S.R. Sternberg, Grayscale morphology, *Computer Vision, Graphics, and Image Processing* 35 (3) (1986) 333–355.
- [15] C. Ronse, A lattice-theoretical morphological view on template extraction in images, *Journal of Visual Communication & Image Representation* 7 (3) (1996) 273–295.
- [16] M. Khosravi, R. Schafer, Template matching based on a grayscale hit-or-miss transform, *IEEE Transactions on Image Processing* 5 (5) (1996) 1060–1066.
- [17] D.S. Bloomberg, P. Maragos, Generalized hit-miss operations, in: *SPIE Conference 1350, Image Algebra and Morphological Image Processing*, San Diego, CA, 1990, pp. 116–128.
- [18] R. Gonzalez, R. Woods, *Digital Image Processing*, Prentice-Hall, Englewood, 2002 (ISBN: 0-20-118075-8).
- [19] P. Maragos, Optimal morphological approaches to image matching and object detection, in: *International Conference on Computer Vision*, 1988, pp. 695–699.
- [20] M. Louys, B. Perret, B. Vollmer, F. Bonnarel, S. Lefèvre, C. Collet, LSB galaxies detection using Markovian segmentation on astronomical images, in: *ASP Conference Series, Astronomical Data Analysis Software and Systems XVII*, London, UK, 2007, pp. 116–128.
- [21] C. Collet, F. Flitti, Variations on Markovian quadtree model for multiband astronomical image analysis, in: *International Symposium on Applied Stochastic Models and Data Analysis*, Brest, France, 2005.
- [22] S. Sabatini, J. Davies, R. Scaramella, R. Smith, M. Baes, S.M. Linder, S. Roberts, V. Testa, The dwarf LSB galaxy population of the Virgo cluster–I. The faint-end slope of the luminosity function, *Monthly Notices of the Royal Astronomical Society* 341 (2003) 981–992.
- [23] T.Y. Chen, F.-C. Kuo, R. Merkel, On the statistical properties of the F-measure, in: *Proceedings of the Fourth International Conference on Quality Software, QSIQ*, 2004, pp. 146–153.
- [24] E. Aptoula, S. Lefèvre, A comparative study on multivariate mathematical morphology, *Pattern Recognition* 40 (11) (2007) 2914–2929.

About the Author—BENJAMIN PERRET received the M.Sc. in Computer Science from Louis Pasteur University, Strasbourg, France, in 2007, where he is currently pursuing a Ph.D. degree in Computer Science with a thesis on galaxy classification using Bayesian modelling, Markov chains Monte Carlo methods and mathematical morphology in multispectral astronomical images.

About the Author—SÉBASTIEN LEFÈVRE received in 1999 the M.S. and Engineering degrees from the University of Technology of Compiègne in Computer Engineering, and in 2002 the Ph.D. degree from the University of Tours in Computer Science. He is currently an Assistant Professor in the Department of Computer Science and the LSIIT, University Louis Pasteur, Strasbourg. From 1999 to 2002 he was with AtosOrigin as a Research and Development Engineer. In 2003, he was with the Polytechnical School of the University of Tours as an Assistant Professor. His research interests are in multispectral image and colour video processing, multimedia analysis and indexing, and mathematical morphology.

About the Author—CHRISTOPHE JEAN-FRANÇOIS COLLET was born in 1966 in France. He graduated from the Université Paris-Sud Orsay in 1989 (Master in Signal Processing, DEA) and received a Ph.D. in Image Processing from the University of Toulon in 1992. His thesis research was carried out at the Ecole Nationale Supérieure des Télécommunications de Bretagne (ENST-Br) and at the French Naval Academy. He spent 8 years at the French Naval Academy and was the Chairman of the Laboratory GTS “Groupe de Traitement du Signal” from 1994 to 2000, where he developed hierarchical Markovian approaches for SONAR image segmentation. Since 2001 he holds a Full Professor position at Strasbourg University (LSIIT UMR CNRS 7005). His major research interests include multi-image segmentation and classification with hierarchical approaches (wavelet decomposition, multigrid optimization, multiscale modeling), Bayesian inference, Markovian approaches for pattern recognition, Bayesian networks, with a particular focus on astronomy (hyperspectral) and medical (multimodal) images.

Table des matières

Préambule	5
1 Contexte	9
1.1 Mon parcours : de l'applicatif au fondamental avec des applications	9
1.1.1 Analyse et indexation de séquences vidéo	9
1.1.2 Morphologie mathématique	10
1.2 Résumé des travaux en morphologie mathématique	12
1.2.1 Travaux fondamentaux	12
1.2.2 Applications	13
1.2.3 Synthèse	15
1.3 Curriculum vitae	16
1.4 Activités d'enseignement et responsabilités pédagogiques	17
1.4.1 Enseignements	17
1.4.2 Charges administratives et fonctions d'intérêt collectif	19
1.5 Animation de la recherche	20
1.5.1 Encadrement d'activités de recherche	20
1.5.2 Rayonnement scientifique	22
1.5.3 Activités contractuelles et projets de recherche	26
1.6 Liste de publications	28
2 Travaux fondamentaux	35
2.1 Introduction à la morphologie mathématique	35
2.1.1 Introduction	35
2.1.2 Cadre théorique et notations	36
2.1.3 Opérateurs et filtres morphologiques	38
2.1.4 Quelques exemples de traitements morphologiques	42
2.1.5 Descripteurs morphologiques	45
2.1.6 Conclusion	47
2.2 Approches morphologiques pour les images multivaluées	48
2.2.1 Introduction	48
2.2.2 Panorama des ordres vectoriels	48
2.2.3 Ordres totaux équilibrés	52
2.2.4 Décompositions en ensembles d'images	55
2.2.5 Conclusion	58
2.2.6 Perspectives	61
2.3 Intégration de connaissances en morphologie mathématique	64
2.3.1 Introduction	64
2.3.2 Formalisation par les éléments structurants	66
2.3.3 Connaissance limitée au nombre de classes d'intérêt	67
2.3.4 Utilisation de régions de référence	70
2.3.5 Conclusion	72
2.3.6 Perspectives	74

3	Applications	77
3.1	Analyse d'images couleur	77
3.1.1	Introduction	77
3.1.2	Détection d'objet et segmentation en régions	79
3.1.3	Description d'image	80
3.1.4	Annotation semi-automatique et recherche d'image couleur	84
3.1.5	Conclusion	85
3.1.6	Perspectives	87
3.2	Télédétection	88
3.2.1	Introduction	89
3.2.2	Télédétection en environnement urbain	90
3.2.3	Télédétection en environnement du littoral	96
3.2.4	Conclusion	100
3.2.5	Perspectives	101
3.3	Imagerie astronomique	103
3.3.1	Introduction	103
3.3.2	Segmentation de galaxies	105
3.3.3	Détection de galaxies	106
3.3.4	Classification de galaxies	107
3.3.5	Conclusion	108
3.3.6	Perspectives	109
3.4	PELICAN : une plate-forme pour l'analyse et le traitement des images	110
3.4.1	Introduction	111
3.4.2	Architecture et fonctionnalités	112
3.4.3	Interfaces et usages	115
3.4.4	Conclusion	116
3.4.5	Perspectives	117
4	Programme de recherche	119
4.1	Imagerie du vivant	120
4.1.1	Invariance et imprécision	120
4.1.2	Applications en imagerie biologique	120
4.1.3	Applications en imagerie médicale	121
4.2	Analyse de séquences vidéo	121
4.2.1	Segmentation spatio-temporelle	121
4.2.2	Indexation multimédia	122
4.3	Morphologie mathématique : plus loin que l'analyse d'image	123
4.3.1	Traitements des différents média	123
4.3.2	Classification et indexation de données multidimensionnelles	124
4.4	Conclusion générale	125
	Bibliographie	125
A	Publications jointes	135

Table des figures

2.1	Morphologie mathématique ensembliste et images binaires	37
2.2	Morphologie mathématique ensembliste et images à niveaux de gris	37
2.3	Arbre des coupes Max-Tree	42
2.4	Processus de segmentation par ligne de partage des eaux	44
2.5	Comparaison entre espaces-échelle Gaussien et morphologique	45
2.6	Comparaison des stratégies marginale et vectorielle	49
2.7	Apparition de fausses couleurs dans le cas d'une stratégie marginale.	50
2.8	Importance de l'ordre des composantes dans l'ordre lexicographique	52
2.9	Illustration du principe de quantification spectrale	53
2.10	Illustration de l'intérêt de l'information spatiale dans la quantification	54
2.11	Illustration du principe de quantification itérative	55
2.12	Treillis d'images binaires construit par regroupements successifs	57
2.13	Synopsis des différentes approches multivaluées en morphologie mathématique	61
2.14	Formalisation de la segmentation comme une classification	69
2.15	Comparaison entre segmentation et classification	69
2.16	Approches morphologiques supervisées sans prise en compte de l'information spatiale	71
2.17	Processus de segmentation par marqueurs contextuels	72
2.18	Segmentation d'une image couleur guidée par des marqueurs contextuels	73
3.1	Illustration de l'intérêt d'utiliser plusieurs teintes de référence en morphologie mathématique couleur	79
3.2	Segmentation non-supervisée d'images couleur	81
3.3	Segmentation guidée par des marqueurs spatiaux et colorimétriques	81
3.4	Mesures granulométriques spécifiques à certaines couleurs prédéfinies	82
3.5	Intérêt du couplage histogramme – granulométrie en recherche d'image	82
3.6	Intérêt des descripteurs morphologiques multidimensionnels	83
3.7	Architecture du système d'annotation et de recherche MIMAR	85
3.8	Images satellite à différentes résolutions	90
3.9	Intérêt de la classification orientée-objets en télédétection	91
3.10	Architecture proposée dans le projet FoDoMuST	92
3.11	Intérêt d'une segmentation supervisée pour l'interprétation d'images satellite à très haute résolution spatiale	93
3.12	Détection du bâti par des segmentations morphologiques supervisée et semi-supervisée	94
3.13	Détection du bâti avec la transformée en tout-ou-rien	95
3.14	Différentes représentations d'un objet du littoral (trait de côte)	97
3.15	Segmentation guidée par quelques marqueurs pour l'extraction du trait de côte.	97
3.16	Comparaison de différentes méthodes de détection du trait de côte	98
3.17	Détection du trait de côte dans différentes configurations	99
3.18	Exemples d'images astronomiques	104
3.19	Exemple d'une planche astronomique	104
3.20	Représentation multispectrale d'une galaxie	105
3.21	Segmentation d'images astronomiques	106
3.22	Comparaison entre galaxies HSB et LSB	106

3.23	Détection de galaxies LSB avec la transformée en tout-ou-rien	107
3.24	Classification de Hubble	107
3.25	Architecture de la plate-forme PELICAN	113
3.26	Exemple de code source PELICAN	114

Approches multivaluées et supervisées en morphologie mathématique et applications en analyse d'image

Résumé — Ce mémoire présente une synthèse des activités de recherche en morphologie mathématique menées au LSIIT (UMR 7005 CNRS-UDS) depuis 2003. La morphologie mathématique est une théorie introduite il y a quarante ans par les chercheurs français Georges Matheron et Jean Serra. Elle connaît depuis un vif succès dans la communauté de l'analyse et du traitement des images, puisqu'elle permet l'analyse des structures spatiales (le plus souvent au travers d'un voisinage défini par le terme d'élément structurant) dans un cadre non-linéaire. Son application aux images binaires et aux images en niveaux de gris s'effectue très simplement en s'appuyant sur la théorie des ensembles ou mieux, celle des treillis.

Cependant, son extension au cas des images multivaluées (où chaque pixel est représenté par un vecteur et non plus par un scalaire) n'est pas triviale et reste un problème ouvert. Ainsi, nous nous sommes intéressés aux approches morphologiques vectorielles basées sur des ordres totaux (aux fondements théoriques les plus valides), en cherchant à atténuer, à l'aide de différentes méthodes de quantification, leur caractère fortement asymétrique afin de mieux exploiter l'ensemble des données disponibles. Nous avons également étudié une autre stratégie consistant à éviter l'appel explicite à un ordre vectoriel, et à décomposer l'image en un ensemble de composantes binaires ou à niveaux de gris, traitées indépendamment ou conjointement.

Indépendamment de la nature des images considérées, la création des systèmes d'analyse d'image par morphologie mathématique nécessite le plus souvent une expertise du domaine et une connaissance très fine du problème pour pouvoir choisir, combiner, et paramétrer les opérateurs morphologiques à bon escient. De ce fait, les méthodes morphologiques ne peuvent généralement pas être réutilisées dans un contexte différent de celui pour lequel elles ont été élaborées, et ne respectent que très peu la contrainte de généralité souhaitée en analyse d'image. Ce problème n'est bien sûr pas spécifique à la morphologie mathématique et est récurrent en traitement d'image, et nous l'avons abordé selon deux axes principaux. D'une part, nous avons étudié les connaissances pouvant être formalisées au sein des éléments structurants dans le contexte de la détection d'objet. D'autre part, nous avons exploité des procédures de classification supervisée (où des ensembles d'apprentissage sont fournis par l'expert) ou non-supervisée (où seul le nombre d'objets ou de classes d'intérêt est connu) au sein du processus de segmentation d'image en régions.

L'objectif sous-jacent à ces travaux fondamentaux est d'aboutir à terme à des approches morphologiques multivaluées et guidées par les connaissances, aptes à traiter tout type d'information, dans tout contexte. Nous avons donc cherché à appliquer ces développements théoriques et méthodologiques dans différents domaines, en particulier l'analyse d'images couleur (dans un but d'annotation et de recherche par le contenu), ainsi que la télé-détection (à très haute résolution spatiale) et l'imagerie astronomique (où les données peuvent être particulièrement bruitées). Ces domaines d'application, où les images sont de nature multivaluée et où l'intégration de connaissances pour guider les traitements est nécessaire, sont particulièrement pertinents puisque l'information spatiale y est cruciale, la morphologie mathématique prenant alors tout son sens. Les problèmes récurrents rencontrés dans ces différents domaines sont la détection, la segmentation, et la description des images. En complément à ces travaux relatifs à la morphologie mathématique, nous présentons le projet PELICAN, une plate-forme générique et extensible pour le traitement d'image.

Ce mémoire se termine par une présentation de quelques perspectives de recherche envisagées dans le cadre de différentes collaborations. Ainsi, l'apport des propriétés d'invariance et d'imprécision dans le contexte de la morphologie mathématique aurait des répercussions en imagerie du vivant. L'analyse morphologique de séquences vidéo, et l'élaboration de descripteurs morphologiques locaux offriraient des solutions alternatives en indexation multimédia. Enfin, la morphologie mathématique n'étant par définition pas limitée à des données de type image, son application à des données de différentes natures mérite d'être étudiée avec une attention particulière.

Mots-clé — Morphologie mathématique, images multivaluées, intégration de connaissances, analyse d'images couleur, télé-détection, imagerie astronomique, système d'analyse et de traitement d'image

Multivalued and Supervised Approaches within Mathematical Morphology, and Applications in Image Analysis

Abstract — This dissertation contains a summary of research activities in mathematical morphology done in LSIIT (UMR 7005 CNRS-UDS) since 2003. Mathematical morphology is a theory introduced more than 40 years ago by two french researchers, Georges Matheron and Jean Serra. Since then, it has been widely used in the fields of image analysis and processing, due to its ability to analyse spatial structures (most often by means of a neighbourhood called structuring element) in a non-linear framework. Its application to binary and greylevel images is straightforward, relying on the set theory or preferably the lattice theory.

However, its extension to multivalued images (where each pixel is represented by a vector instead of a scalar) is not trivial and is still an open problem. Thus, we focused on vectorial morphological approaches based on (theoretically sound) total orderings, and we have try to alleviate their high asymmetric behavior with various quantification methods in order to rely on the whole set of available data. We have also studied another strategy which avoids the explicit need of a vectorial ordering by decomposing the image into a set of binary or greyscale components, which are then processed independently or jointly.

Regardless of image type, building a morphological image analysis system requires most often some domain expertise and very precise knowledge about the problem under consideration in order to select, combine, and set up adequately the morphological operators to be involved. Thus, morphological methods cannot usually be reused in a different context from the one for which they have been designed, and do not ensure the genericity property expected in image analysis. This problem is of course not limited at all to mathematical morphology (it is frequently encountered in image processing), and we have addressed it following two principal directions. On the one side, we have studied how knowledge could be formalised within structuring elements in the context of object detection. On the other side, we have relied on techniques from supervised classification (where learning sets are provided by the expert) or unsupervised classification or clustering (where only the number of objects or classes of interest is known) within the process of image segmentation.

The overall objective is to obtain multivalued and supervised morphological approaches, able to process every kind of data, in any context. So we have applied these works in several fields, in particular colour image analysis (with the aim of content-based image annotation and retrieval), remote sensing (at very high spatial resolution), and astronomical imaging (where data can be very noisy). These application fields, where images are multivalued and knowledge integration to drive image processing techniques is necessary, are relevant since spatial information is crucial (thus mathematical morphology is worth being used). Recurring problems we were dealing with in these fields are object detection, image segmentation and description. Complementary to these mathematical morphology-related works, we also present the PELICAN project, which is a generic and extensible framework for image processing.

This dissertation ends with indicating a few research directions for future works in the context of several collaborative projects. Ensuring invariance and uncertainty within mathematical morphology would help solving biomedical imaging-related problems. Using mathematical morphology for video sequence analysis and design of local image descriptors would offer some alternative solutions for multimedia indexing. Finally, since mathematical morphology is by no way limited to image data, its application to other types of data should be carefully studied.

Keywords — Mathematical morphology, multivalued images, knowledge integration, colour image analysis, remote sensing, astronomical imaging, image analysis and processing system