



HAL
open science

Comparaison d'images binaires reposant sur une mesure locale des dissimilarités. Application à la classification.

Etienne Baudrier

► To cite this version:

Etienne Baudrier. Comparaison d'images binaires reposant sur une mesure locale des dissimilarités. Application à la classification.. Interface homme-machine [cs.HC]. Université de Reims - Champagne Ardenne, 2005. Français. NNT: . tel-00011570

HAL Id: tel-00011570

<https://theses.hal.science/tel-00011570>

Submitted on 9 Feb 2006

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.

Université de REIMS CHAMPAGNE-ARDENNE
UFR des Sciences Exactes et Naturelles

Thèse en vue de l'obtention du diplôme de docteur
en Traitement de l'image et en Mathématiques Appliquées

Comparaison d'images binaires reposant sur une mesure locale des dissimilarités Application à la classification

Étienne BAUDRIER

Thèse soutenue le vendredi 9 décembre 2005.

Composition du jury :

Jacques	LABICHE	(rapporteur)
Gilles	MILLON	(co-directeur)
Frédéric	NICOLIER	(co-directeur)
Sylvie	PHILIPP-FOLIGUET	(rapporteur)
Alain	RIFFAUD	
Su	RUAN	(directrice de thèse)



Université de REIMS
CHAMPAGNE-ARDENNE



Centre de Recherche en
Sciences et Techniques de
l'Information et la Commu-
nication

Remerciements

Je tiens tout d'abord à remercier mes encadrants, Su Ruan, Frédéric Nicolier et Gilles Millon pour leur disponibilité et leurs conseils avisés. Mais aussi pour leur bonne humeur et leur franchise qui ont donné une ambiance de travail stimulante et productive. Je suis très reconnaissant à Jacques Labiche et Sylvie Philipp-Foliguet d'avoir accepté d'être rapporteurs et de m'avoir éclairé de leurs points de vue pertinents, ainsi qu'à Alain Riffaud qui a accepté d'être examinateur et qui m'a donné des conseils utiles. Ensuite, je tiens à remercier chaleureusement tout ceux qui ont fait de cette thèse une période agréable : les membres de l'IUT de Troyes, en particulier Victor pour sa bonne humeur, Manu et les filles pour le déjeuner, Alice et Ben pour les discussions du soir. Enfin pour tout les Troyens que je ne peux nommer dans le détail... merci !

Table des matières

Introduction générale	1
1 Contexte	1
2 Problématique	1
3 Contribution de la thèse	1
3.1 Méthode de comparaison des images binaires	2
3.2 Exploitation pour la classification	2
3.3 Choix de l'information à comparer	2
4 Organisation de la thèse	3

Partie I Analyse Multirésolution

1	
Cadre général	
1.1 Cas d'un signal unidimensionnel	9
1.1.1 Espaces d'approximation	10
1.1.2 Espaces des détails	10
1.1.3 Algorithmes récursifs	10
1.1.4 Ondelettes biorthogonales	15
1.1.5 Bilan	16
1.2 Cas d'un signal bidimensionnel	16
1.2.1 Matrice de changement d'échelle	16
1.2.2 Axiomatique de base	17
1.3 Bilan	17
1.4 Analyses multirésolution non-linéaires	18
1.4.1 Cas général	18

1.4.2	Analyse multirésolution (2 ^{ième} génération)	19
1.4.3	Ondelettes (2 ^{ième} génération)	19
1.4.4	Transformation en ondelettes rapide	20
1.4.5	Lifting scheme	20
1.5	Bilan	22

2

Morphologie mathématique et analyse multirésolution
--

2.1	Morphologie mathématique	23
2.1.1	Définitions générales	23
2.2	Filtres morphologiques	24
2.2.1	Définition	26
2.2.2	Construction de filtres morphologiques	26
2.2.3	Filtres alternés séquentiels	27
2.2.4	Filtres auto-duaux	27
2.2.5	Bilan	28
2.3	Analyse multirésolution morphologique	28
2.3.1	Décomposition en ondelettes couplée	28
2.3.2	Décomposition en ondelettes découplée	28
2.3.3	Ondelettes de Haar morphologiques en 1d	29
2.3.4	Ondelettes de Haar morphologiques en 2d	29
2.3.5	Bilan	30
2.4	Caractérisation morphologique	31
2.4.1	Granulométrie	31
2.5	Méthodes utilisées	36
2.5.1	Utilisation des détails	39
2.5.2	Étude des filtres	40
2.6	Bilan	42

Partie II Mesures de Dissimilarités

1

La comparaison d'images : panorama

1.1	Les méthodes de comparaison	47
1.1.1	Les descripteurs des images	47

1.1.2	Les mesures de similarité	49
1.1.3	Bilan	49
1.2	Les méthodes spécifiques aux images binaires	49
1.2.1	Introduction	49
1.2.2	Méthodes de comparaison directes	50
1.2.3	Les méthodes indirectes	54
1.2.4	Bilan	56
1.2.5	Vers une nouvelle méthode de comparaison	56
1.3	Distance de Hausdorff (DH) et ses variantes	56
1.3.1	Distance de Hausdorff, généralités	56
1.3.2	Propriétés générales de la distance de Hausdorff	56
1.3.3	Quelques versions modifiées de la distance de Hausdorff	57
1.3.4	Discussion	59

2

Mesure locale de la DH

2.1	Quel sens donner à l'expression « dissimilarité locale » ?	61
2.2	Définition de la DH locale	62
2.2.1	Définition naïve de la DH locale	62
2.2.2	Modification de la définition naïve	63
2.2.3	Distance de Hausdorff locale	64
2.3	Propriétés de la DH locale HD_W	65
2.3.1	Propriétés générales	66
2.3.2	Propriétés dépendant de la taille de la fenêtre W	66
2.4	Une DH locale adaptative et non-paramétrique	66
2.4.1	Caractérisation de la mesure d'une dissimilarité locale	66
2.4.2	Critère pour le calcul du rayon optimal r_{max}	72
2.4.3	Généralisation	72
2.5	Bilan	75

3

Carte des Dissimilarités Locales

3.1	Définition générale	77
3.2	Mise en œuvre	78
3.2.1	Algorithme général	78
3.2.2	Complexité du calcul	78
3.3	Cas de la Distance de Hausdorff	78
3.3.1	Définition	79
3.3.2	Transformation en distance (TeD)	80
3.4	Mise en œuvre de la Carte des Dissimilarités Locales basée sur la distance de Hausdorff	82
3.4.1	Calcul de la distance de Hausdorff globale	82
3.4.2	Calcul de la Carte des Dissimilarités Locales à partir du théorème 3.3.1	82

3.5	Résultats qualitatifs	83
3.5.1	Lignes et carré	83
3.5.2	Lettres « co » et « et »	83
3.5.3	Jeu des dix erreurs	83
3.6	Comparaison aux DH modifiées	84
3.7	Généralisation aux images en niveaux de gris	84
3.7.1	Généralisation à partir de la définition générale de la Carte des dissimilarités locales	88
3.7.2	Généralisation dans le cas de la distance de Hausdorff	88
3.8	Étude sur la fenêtre glissante W	89
3.8.1	Utilité de l'adaptabilité de la fenêtre W	89
3.8.2	Forme de la fenêtre	89
3.9	Bilan	92

Partie III Application à la classification

0.1	La Carte des Dissimilarités Locales au sein d'un processus global de comparaison	97
0.1.1	Généralités	97
0.1.2	Schéma synoptique	97

1

Classification d'impressions anciennes

1.1	Contexte	99
1.2	Projet ANITA	99
1.3	Prétraitement	100
1.3.1	Acquisition des échantillons	100
1.3.2	Binarisation	101
1.4	Bilan	105

2

Classification basée sur la Carte des Dissimilarités Locales

2.1	Classification : généralités	107
2.1.1	Apprentissage	107
2.1.2	Formalisation de la fonction de classification	108
2.2	Différentes approches de classification	110
2.2.1	Méthodes probabilistes	110
2.2.2	k plus proches voisins	111

2.2.3	Réseaux de neurones de type perceptron	111
2.2.4	Séparateurs à vaste marge	112
2.3	Représentation vectorielle de la Carte des Dissimilarités Locales	113
2.3.1	Histogramme de la Carte des Dissimilarités Locales	114
2.3.2	Granulométrie	114
2.3.3	Bilan	117
2.4	Classification directe sur la Carte des Dissimilarités Locales	117
2.4.1	Méthodologie d'évaluation	119
2.4.2	Carte des Dissimilarités Locales basée sur la distance de Hausdorff (CDL_{DH})	119
2.4.3	Résultats en combinant la Carte des Dissimilarités Locales avec l'analyse multirésolution	121
2.4.4	Classification en trois groupes	123
2.4.5	Carte de Différence Simple (CDS)	124
2.4.6	Carte des Dissimilarités Locales basée sur la différence simple $CDL_{DiffSimple}$	124
2.4.7	Comparatif	126
2.4.8	Robustesse	128

3

Autres applications

3.1	Classification sur une base de formes	133
3.2	Classification sur une base de visages	134
3.2.1	Description pratique	134
3.2.2	Analyse	134
3.3	Mesure globale tirant parti de la Carte des dissimilarités Locales dans le cas de la DH	136

4

Bilan

Partie IV Conclusion générale

Annexes

A

Preuves

A.1	Preuve de la proposition 2.3.1 (identité)	147
A.2	Preuve de la proposition 2.3.2 (majoration)	148
A.3	Preuve de la proposition 2.3.3 (croissance)	151

Glossaire	155
Bibliographie	157

Table des figures

1.1	Représentation de l'analyse multirésolution sur trois niveaux	11
1.2	Schéma de l'analyse pour l'algorithme de Mallat.	13
1.3	Schéma de la synthèse pour l'algorithme de Mallat.	13
1.4	Exemple de décomposition d'un signal s avec les ondelettes de Haar sur 5 échelles	14
1.5	Une image et sa décomposition en ondelettes de Haar 2d.	18
1.6	Schéma de l'analyse pour le lifting scheme.	21
2.1	Illustration des 4 opérations morphologiques élémentaires	25
2.2	Schéma de l'analyse des ondelettes de Haar morphologiques en 2d	30
2.3	Décomposition en ondelettes de Haar 2d morphologiques	31
2.4	Une image et son analyse granulométrique par ouverture	32
2.5	AMR morphologique de l'érosion décimée : les approximations et leurs dérivées granulométriques.	34
2.6	AMR morphologique de Haar décimée : les approximations et leurs dérivées granulométriques.	35
2.7	Courbes et dérivées granulométriques pour l'AMR morphologique de la médiane décimée	36
2.8	Image initiale et ses approximations aux échelles 1 et 2 par le filtre de Haar morphologique	37
2.9	Image initiale et ses approximations aux échelles 1 et 2 par l'ouverture morphologique	37
2.10	Diagramme commutatif concernant la décimation et la complémentation illustrant l'égalité (2.39).	38
2.11	Image initiale et ses approximations aux échelles 1 à 4 par le filtre de la médiane morphologique	39
2.12	AMRM de la médiane	40
2.13	AMRM dyadique avec le filtre de la médiane sur une image binaire	41
2.14	AMRM de la médiane par l'algorithme à trous	42
1.1	Deux couples d'images différents donnant les mêmes images de différence simple	52
1.2	Inégalité triangulaire : contre-exemple	53
1.3	Complexité du contour : formes proches mais contours très différents	55
2.1	Illustration de la notion de dissimilarité locale	62
2.2	Deux exemples pour la définition améliorée.	63
2.3	Illustration des sauts de valeurs pour la définition modifiée lors du déplacement de la fenêtre W	64
2.4	Illustration des sauts de valeurs pour la définition modifiée lors de l'agrandissement de la fenêtre W	65
3.1	Une image binaire et ses trois TeD faites avec les trois distances classiques : L_1, L_2, L_∞	81
3.2	Trois formes simples et leurs CDL	83

3.3	Les lettres CO et ET et leur CDL illustrant leurs dissimilarités locales.	84
3.4	Le jeu des dix erreurs	85
3.5	Contre-exemple pour les mesures globales à partir de lignes	86
3.6	Contre-exemple pour les mesures globales à partir de cercles	87
3.7	CDL entre l'image d'un visage et les images érodées	90
3.8	Comparaison entre la CDL et les Cartes des dissimilarités à taille de fenêtre fixe.	91
3.9	Deux images binaires et leur CDL avec des fenêtres de mesure locale asymétriques	92
0.10	Schéma du processus global utilisant l'AMR, la CDL et le module de décision	98
1.1	Exemple d'acquisition initiale pour deux impressions anciennes	100
1.2	Deux illustrations différentes provenant du même tampon	102
2.1	Taux de reconnaissance de la classification des histogrammes basée sur les SVM	115
2.2	Deux CDL différentes ayant le même histogramme	116
2.3	Résultats de la classification des courbes granulométriques basée sur les SVM	118
2.4	CDL_{DH} : taux de classification en fonction de la taille des échantillons d'apprentissage	120
2.5	Exemple d'images de détails recomposés avec une approximation nulle	123
2.6	CDL_{DH} : Taux de classification de la CDL_{DH} pour les 3 classes en fonction du paramètre C du SVM	124
2.7	Représentation de la $CDL_{difSimple}$ pour différentes valeurs du paramètre d'arrêt	125
2.8	$CDL_{difSimple}$: taux de classification en fonction du paramètre d'arrêt	126
2.9	DH partielle : efficacité de la classification en fonction du paramètre p	127
2.10	Comparaison de la CDL_{DH} et de la $CDL_{difSimple}$ sur une exemple	129
2.11	CDL_{DH} : robustesse aux tâches et aux effacements	130
3.1	Quelques exemples de formes de la base.	133
3.2	Exemples de visages de la base ATT, de leur contours et de leurs CDL	135
3.3	Illustration pour la distance globale DH Filtrée (DHF)	137

Liste des tableaux

1.1	Comparaison des méthodes de binarisation, partie 1	104
1.2	Comparaison des méthodes de binarisation, partie 1	105
2.1	Histogramme de la CDL : taux de reconnaissance à la résolution 128×128	114
2.2	CDL_{DH} : taux de reconnaissance en fonction C et du type de noyau.	120
2.3	CDL_{DH} : taux de reconnaissance en fonction de l'échelle de l'AMR	121
2.4	Exemple d'approximations par l'AMR de la médiane morphologique et leur CDL aux différentes résolutions .	121
2.4	suite	122
2.4	Commentaires sur la conservation des structures par L'AMR de la médiane morphologique	122
2.5	CDL_{DH} : taux de reconnaissance pour les détails recomposés de taille 128×128	123
2.6	CDL_{DH} : résultat du test de la classification en trois classes à la résolution 128×128	124
2.7	Comparaison des méthodes avec celles basées sur la CDL	127
2.8	CDL_{DH} : résultat du test circulaire à la résolution 128×128	128
2.9	Récapitulatif des résultats pour toutes les méthodes de comparaison testées.	131
3.1	Résultat de la méthode sur la bases de formes	134
3.2	Résultat de la méthode sur la bases de visages	134

Introduction générale

1 Contexte

L'essor du traitement de l'image et son investissement de secteurs aussi différents que la vie privée ou la médecine ont donné naissance à de nombreuses bases de données et en particulier d'images.

Le support informatique des images permet notamment de les comparer. Cela peut être utile à plusieurs titres :

- pour visualiser les différences entre les images, et les quantifier automatiquement ;
- pour retrouver des images dans une base d'images. En effet, ces dernières sont souvent trop volumineuses pour être exploitées directement par l'homme. Elles nécessitent alors le développement de méthodes d'exploitation automatiques comme la recherche d'images à partir d'une requête. Dans ce cas particulier de la recherche d'images, les méta-données (texte descriptif, ...) sont parfois absentes ou inadaptées. Il est alors nécessaire de se baser sur le contenu des images pour effectuer la recherche. Plusieurs méthodes existent, les unes basées sur la comparaison de descripteurs des images et les autres sur la comparaison directe des images.

Ainsi, la comparaison d'images est un point important dans plusieurs domaines. Les images binaires présentent dans ce cadre des difficultés spécifiques. La pauvreté de leurs attributs, et parfois complexité des images font que si certains outils ont été développés, par exemple dans le cas de formes simples, leur description reste difficile et leur comparaison aussi. Et pourtant, les images binaires sont présentes dans de nombreux processus. Elles peuvent provenir de capteurs binaires, d'une extraction de contours, d'une binarisation... Leur comparaison peut alors être utile. C'est dans le cadre de la comparaison des images binaires que s'inscrit cette thèse.

2 Problématique

Comparer des images binaires soulève plusieurs problèmes. Le premier est le choix de la résolution de comparaison. En effet, pour des applications qui peuvent être coûteuses en temps de calcul, il est intéressant de ne comparer que ce qui est nécessaire. Ce problème n'est pas spécifique aux images binaires mais donne des contraintes particulières qui demandent une étude pour la multirésolution. Le deuxième problème est la manière de comparer les images à proprement parler. Enfin le troisième porte sur l'exploitation de cette comparaison. Ces trois problèmes soulevés par la comparaison d'images binaires ont été abordés dans cette thèse.

3 Contribution de la thèse

Le développement d'une méthode de comparaison d'images binaires constitue la contribution majeure de cette thèse.

3.1 Méthode de comparaison des images binaires

La solution proposée permet une comparaison directe des images. Elle est basée sur la mesure locale des dissimilarités entre les deux images. L'intérêt de cette mesure est double : d'une part, elle s'adapte automatiquement à la taille de la dissimilarité qu'elle doit évaluer localement, et d'autre part, le regroupement de toutes les mesures locales entre deux images donne une carte des dissimilarités locales (abrégée en CDL) qui comporte l'ensemble des mesures et leur distribution spatiale.

Le principe de la carte des dissimilarités locales est général dans le sens où une mesure de dissimilarité locale quelconque peut être appliquée pour peu qu'elle vérifie quelques hypothèses (croissance, majoration). Cela permet une souplesse dans sa mise en œuvre. Par ailleurs, dans le cas où cette mesure locale est dérivée de la distance de Hausdorff, la formule de calcul est simple et rapide et rend compte des dissimilarités avec une bonne fidélité.

Le fait que la carte des dissimilarités locales contienne l'information sur les valeurs des mesures de dissimilarité et leur répartition spatiale est particulièrement intéressant car cela permet de distinguer des types de dissimilarités à partir de la manière dont sont réparties les valeurs dans la carte (si les hautes valeurs sont regroupées ou éparpillées par exemple). Cela est de plus novateur dans le sens où il n'existe pas, à la connaissance de l'auteur, de représentation concernant la distribution spatiale des dissimilarités dans la littérature. Une exploitation de cette information contenue dans la carte des dissimilarités locales concerne la classification des images binaires.

3.2 Exploitation pour la classification

La carte des dissimilarités locales est une représentation d'une mesure des dissimilarités locales entre deux images. Dans la classification, le but est de pouvoir classer les images par classes de similarité. Comme une carte des dissimilarités locales est associée à un couple d'images, cela revient à classer les cartes des dissimilarités locales en deux classes : celles comparant des images similaires et celles comparant des images dissimilaires. Or la carte des dissimilarités locales est en deux dimensions. Son exploitation pour la classification n'est donc pas immédiate. La thèse se propose d'exploiter les données de la carte des dissimilarités locales par un Séparateur à Vaste Marge (SVM). En effet, les Séparateurs à Vaste Marge supportent assez bien les données de grande dimension, ce qui permet de leur soumettre intégralement la CDL pour effectuer la classification.

Cependant toute l'information contenue dans les images peut ne pas être pertinente pour la comparaison. Le choix de cette information s'effectue à l'aide d'une analyse multirésolution (AMR).

3.3 Choix de l'information à comparer

Comme toute l'information contenue dans les images n'est pas pertinente, et que se posent aussi des questions évidentes relatives au temps de calcul, il est intéressant de décomposer de manière contrôlée l'image afin de choisir la partie la mieux adaptée à la comparaison. Le souhait de contrôler la décomposition nous a conduit à nous placer dans le cadre de l'analyse multirésolution. Cependant le cas particulier des images binaires amène à sortir du cadre de l'analyse multirésolution classique pour utiliser celui des ondelettes de seconde génération et considérer les analyse multirésolution non-linéaires basées sur un opérateur morphologique. L'étude dans ce cas est délicate car il n'y a pas d'équivalent (non-linéaire) de l'analyse de Fourier qui fournit au cas linéaire un puissant outil de caractérisation. De ce fait, la caractérisation des filtres et le développement de nouveaux filtres ne sont pas bien maîtrisés. L'apport de la thèse sur ce point est d'avoir fait le point sur ce qui a été déjà proposé, d'avoir testé plusieurs opérateurs (Haar, ouverture, médiane) pour la multirésolution, d'avoir mis en œuvre différents algorithmes dans ce cadre et d'avoir étudié des outils de caractérisation ainsi que la possibilité de nouveaux filtres.

4 Organisation de la thèse

La thèse est organisée en trois parties.

Première partie La première partie porte sur l'analyse multirésolution qui permet de choisir au mieux l'information à comparer dans les images. Le cadre général de l'analyse multirésolution classique y est présenté. Il permet d'introduire celui des ondelettes de deuxième génération. Un point est alors fait sur la morphologie mathématique. Enfin l'analyse multirésolution morphologique est présentée en utilisant les filtres morphologiques dans le cadre des ondelettes de deuxième génération. Cela conduit à choisir l'analyse multirésolution morphologique utilisant le filtre de la médiane.

Deuxième partie La deuxième partie présente un panorama sur les méthodes de comparaison d'images puis situe notre propre démarche. En se basant sur un cas concret (la distance de Hausdorff), le principe de la mesure locale est exposé et ses propriétés démontrées. Cela permet de définir un critère de choix pour la taille de la fenêtre dans laquelle est faite la mesure locale. La mesure des dissimilarités locale en découle dans le cas général et dans le cas de la distance de Hausdorff. Des exemples sont donnés pour illustrer les propriétés de la carte des dissimilarités locales.

Troisième partie Cette dernière partie est dédiée aux applications de la CDL qui sont la classification, et en particulier la classification d'impressions anciennes numérisées, dans le cadre d'une coopération entre le CReSTIC, la MAT et le CEPLECA. La classification est aussi testée sur une base de formes et une base de visages. Puis ses propriétés de visualisation sont exploitées pour l'évaluation des filtres et le calcul d'une mesure globale tirant parti de l'information spatiale contenue dans la carte des dissimilarités locales.

Première partie

Analyse Multirésolution

Introduction

Lors de la comparaison de deux images, il n'est pas toujours nécessaire de comparer l'ensemble des éléments présents dans les images. Par exemple, si la comparaison porte sur les traits grossiers de l'image, il n'est pas utile de conserver une haute définition pour l'image, au risque de surcharger les calculs. Dans cette optique, l'Analyse Multirésolution (AMR) offre un cadre où la décomposition de l'image en détails et approximations est contrôlée puisque la reconstruction est parfaite. Le formalisme classique est présenté section 1. Cependant les images binaires doivent être décomposées de manière non-linéaire afin de donner des images d'approximation et de détail binaires. Le cadre des ondelettes de deuxième génération qui se base sur le cas classique des ondelettes biorthogonales offre la possibilité de développer une AMR non-linéaire. Elles sont présentées dans le paragraphe 1.4. Le cadre de la morphologie mathématique offre la possibilité de filtrer les images binaires et une introduction est faite dans la section 2. Ce type de filtre est utilisé dans le cadre de l'AMR non-linéaire et permet d'obtenir des AMR morphologiques vérifiant les propriétés de reconstruction parfaite. Ce dernier point et le choix de l'opérateur sont présentés dans le paragraphe 2.5.

1

Cadre général

Le but de cette section est de formaliser l'idée d'analyse multirésolution destinée à décomposer une image en approximations et détails à plusieurs échelles. Cela permet de choisir le type de structure à traiter dans la décomposition et par exemple, traiter une image à basse résolution puis augmenter la résolution si nécessaire. Nous présentons dans cette section les bases de l'AMR et les algorithmes (décomposition dyadique et algorithme à trous) qui seront utilisés par la suite. La présentation générale concerne les signaux unidimensionnels puis elle est étendue à la dimension 2. Les ondelettes biorthogonales sont alors introduites afin de servir à la généralisation de l'AMR classique au cas non-linéaire.

La théorie des ondelettes mise en place dans les années 80 et développée dans les années 90 a connu un succès important dans des domaines d'application aussi variés que le traitement de l'image et du signal, la compression et la transmission de données, le débruitage, les solutions numériques des équations différentielles et intégrales. La clef de son succès tient dans le fait que pour une classe de fonctions, la majeure partie de l'information contenue dans la fonction est concentrée dans un petit nombre de coefficients d'ondelettes. La plupart des ondelettes ont été construites dans les années 90, nous faisons référence aux travaux suivants : [Battle, 1987; Lemarié, 1988; Mallat, 1989; Meyer, 1990; Daubechies, 1992; Donoho, 1992; Daubechies, 1993; Cohen and Daubechies, 1993; Herley and Vetterli, 1993].

La section suivante est détaillée dans [Mallat, 2000] et [Truchetet, 1998].

1.1 Présentation dans le cas d'un signal unidimensionnel

Nous nous plaçons dans le cadre de l'espace $L^2(\mathbb{R})$ des fonctions de la variable réelle mesurables de carré intégrable.

1.1.1 Espaces d'approximation

Définition 1.1.1 (Analyse multirésolution). Une suite $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces fermés de $L^2(\mathbb{R})$ est une approximation multirésolution si elle vérifie les 5 propriétés suivantes

$$\forall (j, k) \in \mathbb{Z}^2, \quad f(x) \in V_j \Leftrightarrow f(x - 2^j k) \in V_j \quad \text{invariance par translation,} \quad (1.1)$$

$$\forall j \in \mathbb{Z}, \quad V_{j+1} \subset V_j \quad \text{emboîtement,} \quad (1.2)$$

$$\forall k \in \mathbb{Z}, \quad f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1} \quad \text{changement d'échelle,} \quad (1.3)$$

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \quad (1.4)$$

$$\lim_{j \rightarrow -\infty} V_j = \text{Adhérence}\left(\bigcup_{j=-\infty}^{\infty} V_j\right) = L^2(\mathbb{R}) \quad \text{densité.} \quad (1.5)$$

L'espace V_{j+1} contient des signaux plus grossiers que l'espace V_j et lorsque j tend vers l'infini, (1.4) implique que l'on perd tous les détails de f . On peut montrer l'existence d'une fonction φ dite fonction d'échelle qui par translation engendre une base orthonormée de V_0 . La fonction φ peut alors s'interpréter comme une cellule unitaire de la résolution 0. Les propriétés (1.2) et (1.3) impliquent que par dilatation et translation, φ engendre V_j . Les fonctions de base $(\varphi_{j,n})_{n \in \mathbb{Z}}$ de V_j sont construites à partir de φ selon la relation suivante :

$$\varphi_{j,n}(x) = 2^{-\frac{j}{2}} \varphi(2^{-j}x - n) \text{ avec } n \in \mathbb{Z}. \quad (1.6)$$

1.1.2 Espaces des détails

L'espace des détails W_j est défini comme le supplémentaire (complémentaire orthogonal) de l'espace d'approximation V_j dans V_{j-1} .

Définition 1.1.2 (espace des détails). On appelle espace des détails l'espace vectoriel W_j tel que :

$$V_{j-1} = V_j \oplus W_j, \quad (1.7)$$

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j. \quad (1.8)$$

Les sous-espaces W_j ne sont pas emboîtés, cependant, les propriétés d'échelle et d'invariance par translation sont conservées.

Il est possible de montrer alors l'existence d'une fonction notée ψ appelée *ondelette* qui engendre une base orthonormée de W_0 par translation. Elle engendre alors par translation et dilatation les espaces W_j . Les fonctions de base $(\psi_{j,n})_{n \in \mathbb{Z}}$ de W_j s'expriment alors suivant la relation :

$$\psi_{j,n} = 2^{-\frac{j}{2}} \psi(2^{-j}x - n) \text{ avec } n \in \mathbb{Z}. \quad (1.9)$$

Le schéma de l'analyse multirésolution est présentée dans la figure 1.1.2.

1.1.3 Algorithmes récursifs

Dans la pratique, seule une version discrète de la fonction f est disponible, qui est généralement assimilée à une approximation à une échelle donnée. Principalement, deux algorithmes permettent de décomposer cette donnée en approximation et

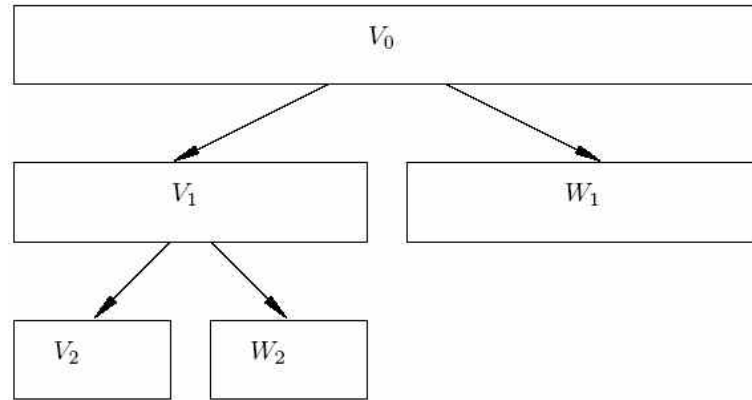


FIG. 1.1 – Représentation de l'analyse multirésolution sur trois niveaux. L'information dans V_0 est partagée à l'échelle 1 entre V_1 et W_1 et ainsi de suite.

détails, ce sont

- l'algorithme de [Mallat, 1989] qui est linéaire et fait appel à un sous-échantillonnage. Il a le défaut de ne pas être invariant par translation,
- l'algorithme à trous [Holschneider *et al.*, 1989] qui conduit à une redondance des données mais qui est invariant par translation.

Le passage à la version discrète a fait l'objet de nombreuses études qui se sont rapidement concentrées sur l'Analyse Multi-résolution (AMR) dyadique. Les données initiales $c(0, k)$ sont les produits scalaires de la fonction f avec la fonction d'échelle φ tradatée au point k :

$$c(0, k) = \langle f, \varphi_{0,k} \rangle \quad (1.10)$$

$$= \int_{-\infty}^{+\infty} f(u)\varphi(u - k)du \quad (1.11)$$

ou, autrement dit, ce sont les coefficients résultant de la projection de f sur la base orthonormée de V_0 engendrée par φ . L'échantillonnage ne correspond donc pas à la valeur de f au point x mais à la moyenne pondérée au point x de f par le translaté de φ .

Algorithme de Mallat (transformée en ondelettes rapide) Les coefficients à une échelle plus grossière sont obtenus de la même manière que (1.10) mais en projetant sur la base $(\varphi_{j,k})_{k \in \mathbb{Z}}$ de V_j . L'équation d'échelle donne la définition d'un filtre numérique qui permet d'obtenir une relation directe entre les coefficients de f à deux échelles successives.

Proposition 1.1.1. *La propriété de causalité des multirésolutions (1.2) impose que $V_j \subset V_{j-1}$. En particulier, $2^{-\frac{j}{2}}\varphi(x/2) \in V_1 \subset V_0$. Comme $(\varphi_{0,k})_{k \in \mathbb{Z}}$ est une base orthonormée de V_0 , on peut faire la décomposition*

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) = \sum_{n=-\infty}^{+\infty} h[n]\varphi(x - n), \quad (1.12)$$

avec

$$h[n] = \left\langle \frac{1}{\sqrt{2}} \varphi \left(\frac{\cdot}{2} \right), \varphi_{0,n} \right\rangle. \quad (1.13)$$

On interprète $h[n]$ comme la réponse impulsionnelle d'un filtre discret. Les coefficients $c(j+1, k)$ de f aux résolutions plus grossières se calculent alors à partir du filtre h et des coefficients à l'échelle j :

$$c(j+1, k) = \left\langle f, \frac{1}{\sqrt{2^{i+1}}} \varphi \left(\frac{\cdot}{2^{i+1}} - k \right) \right\rangle \quad (1.14)$$

en remplaçant avec (1.12)

$$c(j+1, k) = \sum_{n=-\infty}^{+\infty} h[n] c(j, 2k+n) \quad (1.15)$$

Les données sont réduites d'un facteur 2 d'une échelle à l'autre dans cette analyse dénommée *analyse dyadique*.

Si nous notons \tilde{h} le filtre symétrique de h

$$\tilde{h}[n] = h[-n], \quad (1.16)$$

il vient

$$c(j+1, k) = \sum_{n=-\infty}^{+\infty} \tilde{h}[2k-l] c(j, l) \quad (1.17)$$

La somme précédente peut être interprétée comme un produit de convolution entre \tilde{h} et $c(j, \cdot)$ évalué pour un indice sur deux ; ou encore comme le filtrage de la séquence $c(j, \cdot)$ par le filtre de réponse impulsionnelle \tilde{h} suivi d'un sous-échantillonnage de rapport deux.

Un calcul en tout point analogue au précédent permet d'écrire les coefficients de détail : tout d'abord la fonction d'ondelette ψ_1 appartient à V_0 et donc

$$\psi_1 = \sum_{n=-\infty}^{+\infty} g[n] \varphi_{0,n} \quad (1.18)$$

ce qui conduit à l'équation de construction de $g[n]$ suivante

$$g[n] = \langle \psi_1, \varphi_{0,n} \rangle \quad (1.19)$$

On peut montrer que ce filtre est lié au filtre h et qu'il peut être construit à partir de ce dernier. Nous avons alors pour les coefficients de détail

$$d(j+1, k) = \langle f, \psi_{j+1,k} \rangle \quad (1.20)$$

$$d(j+1, k) = \sum_{n=-\infty}^{+\infty} g[n] c(j, 2n+k) d(j+1, k) = \sum_{n=-\infty}^{+\infty} \tilde{g}[2k-l] c(j, l) \quad (1.21)$$

où

$$\tilde{g}[n] = g[-n] \quad (1.22)$$

Comme pour l'approximation, c'est ici encore un produit de convolution qui peut être interprété comme le filtrage de la séquence $d(j, \cdot)$ par le filtre de réponse impulsionnelle \tilde{g} suivi d'un sous-échantillonnage de rapport deux. L'algorithme de Mallat est schématisé sur la figure 1.2.

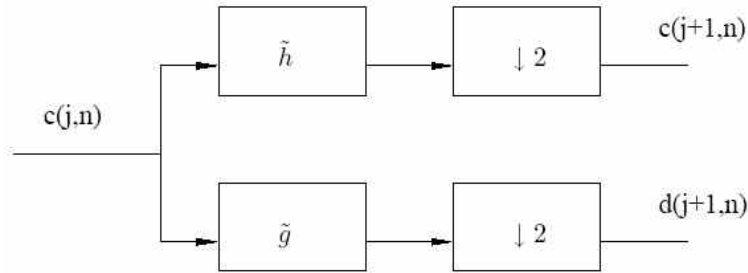


FIG. 1.2 – Schéma de l'analyse pour l'algorithme de Mallat.

Pour un signal initial comportant N coefficients, une représentation complète du signal est obtenue par itération avec $N - 1$ coefficients d'ondelettes et un terme d'approximation. La suite de coefficients obtenue n'est donc pas redondante.

Algorithme de reconstruction Pour reconstruire le signal à partir des coefficients d'approximation et de détail, la méthode est la même que pour l'analyse : l'échelle $j - 1$ (plus fine) est exprimée en fonction de l'approximation et des détails à l'échelle j . Puis en introduisant les filtres h et g , on obtient :

$$c(j-1, k) = \sum_{n=-\infty}^{+\infty} h[k-2n]c(j, n) + \sum_{n=-\infty}^{+\infty} g[k-2n]d(j, n) \quad (1.23)$$

Cette équation est une somme de produits de convolution. Le schéma de l'algorithme de reconstruction est présenté figure 1.3.

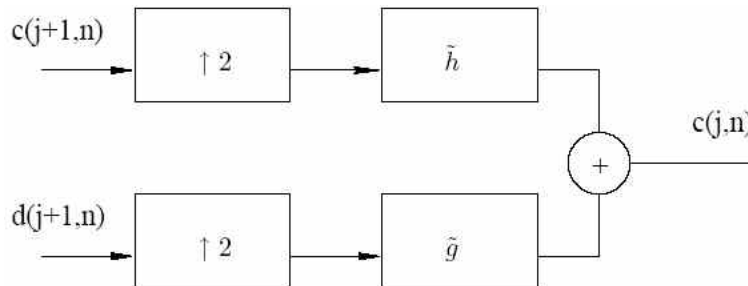


FIG. 1.3 – Schéma de la synthèse pour l'algorithme de Mallat.

Algorithme à trous Cet algorithme [Holschneider *et al.*, 1989; Shensa, 1992] s'implémente avec un banc de filtres semblable à celui de Mallat, mais sans sous-échantillonnage. Il n'y a donc pas décimation, ce qui assure l'invariance par translation,

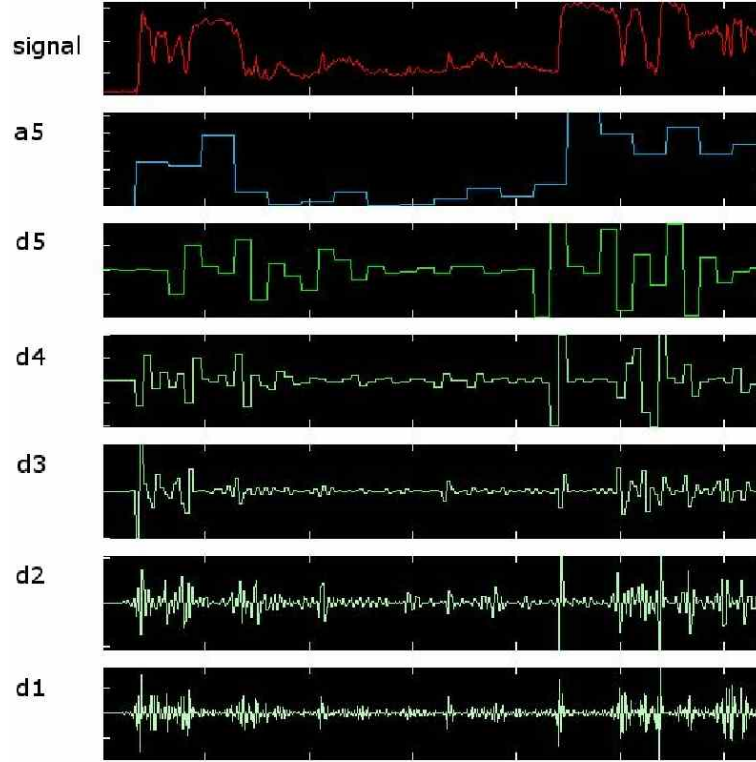


FIG. 1.4 – Exemple de décomposition pour d'un signal s avec les ondelettes de Haar sur 5 échelles : l'approximation a_5 et les détails d_1, \dots, d_5 .

propriété utile dans le traitement et la restauration de l'image [Bijaoui *et al.*, 1994].

Dans l'algorithme à trous, les approximations entrelacées vont être prises en considération pour d'éviter la décimation. De manière générale, l'équation d'échelle à l'échelle $j + 1$ donne le pas à la résolution j :

$$\frac{1}{\sqrt{2^{j+1}}} \varphi\left(\frac{x}{2^{j+1}}\right) = \sum_{n=-\infty}^{+\infty} h[n] \varphi\left(\frac{x}{2^j} - n\right), \quad (1.24)$$

on l'utilise dans la relation suivante

$$c(j+1, k) = \left\langle f, \frac{1}{\sqrt{2^j}} \varphi\left(\frac{\cdot}{2^j} - k\right) \right\rangle, \quad (1.25)$$

et on obtient

$$c(j+1, k) = \sum_{n=-\infty}^{+\infty} h[n] c(j, k + 2^i n) \quad (1.26)$$

$$= \sum_{n=-\infty}^{+\infty} \tilde{h}[l - 2^i n] c(j, l). \quad (1.27)$$

Ceci fournit un pas de 2^j . En pratique, le filtre pour passer de l'échelle 0 à l'échelle j est construit à partir du filtre initial \tilde{h} en insérant 2^j 0 entre ses coefficients. Ce qui revient à faire un sur-échantillonnage de rapport 2^j . Alors, seul un coefficient

sur 2^j est utile dans le calcul et cela donne l'origine du terme « à trous ».

1.1.4 Ondelettes biorthogonales

Dans la pratique, les AMR, que nous venons de présenter, comportent un inconvénient : si l'on souhaite préserver la symétrie gauche-droite dans le traitement (linéarité en phase), les filtres considérés dans les algorithmes ont alors une réponse impulsionnelle infinie. Pour le traitement de l'image, cette symétrie dans le traitement est souhaitable et le nombre important de données à traiter appelle des algorithmes rapides. Le cadre des ondelettes biorthogonales est moins contraignant que l'AMR orthogonale et permet de concilier compacité des supports des filtres et symétrie dans le traitement.

Deux familles de fonctions duales sont considérées, l'une pour l'analyse et l'autre pour la reconstruction.

Définition 1.1.3 (Analyse multirésolution biorthogonale). *Une suite $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces fermés de $L^2(\mathbb{R})$ est une approximation multirésolution si elle vérifie les 4 propriétés suivantes :*

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j, \quad (1.28)$$

$$\lim_{j \rightarrow -\infty} V_j = \text{Adhérence} \left(\bigcup_{j=-\infty}^{\infty} V_j \right) = L^2(\mathbb{R}), \quad (1.29)$$

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \quad (1.30)$$

$$\text{pour tout } j, V_j \text{ admet une base de Riesz fourni par les fonctions d'échelle } \{\varphi_{j,k}, k \in \mathbb{Z}\}. \quad (1.31)$$

Définition 1.1.4 (AMR duale). *Une AMR duale $\overline{\mathcal{M}} = \{\overline{V}_j, j \in \mathbb{Z}\}$ est constituée d'une suite d'espaces \overline{V}_j ayant des bases de Riesz $(\overline{\varphi}_{j,k})$. Ces fonctions d'échelles duales sont biorthogonales aux fonctions d'échelles, au sens où*

$$\langle \varphi_{j,k}, \overline{\varphi}_{j,k'} \rangle = \delta_{k,k'} \quad \text{avec } (k, k') \in \mathbb{Z}^2 \quad (1.32)$$

Les espaces complémentaires W sont définis par

$$W_j \perp \overline{V}_j \quad \text{et} \quad \overline{W}_j \perp V_j \quad (1.33)$$

On a les relations suivantes

$$\begin{cases} V_j = V_{j+1} \oplus W_{j+1} \text{ mais } V_{j+1} \not\subset W_{j+1} \\ \overline{V}_j = \overline{V}_{j+1} \oplus \overline{W}_{j+1} \text{ et } \overline{V}_{j+1} \not\subset \overline{W}_{j+1} \text{ avec } W_j \perp \overline{W}_{k \neq j} \end{cases} \quad (1.34)$$

Les propriétés d'orthogonalité des bases des sous-espaces sont les suivantes :

$$\langle \psi_{j,n}, \overline{\psi}_{i,m} \rangle = \delta_{i,j} \delta_{m,n} \quad \text{et} \quad \langle \varphi_{j,n}, \overline{\varphi}_{j,m} \rangle = \delta_{i,j} \delta_{m,n} \quad (1.35)$$

Les algorithmes récursifs sont analogues à ceux du cas orthogonal, avec pour l'analyse

$$c(j+1, k) = \sum_l \tilde{h}[2k-l]c(j, l) \quad \text{et} \quad c(j+1, k) = \sum_l \tilde{g}[2k-l]c(j, l), \quad (1.36)$$

et pour la reconstruction

$$c(j, n) = \sum_k c(j+1, k)h[n-2k] + \sum_k d(j+1, k)g[n-2k]. \quad (1.37)$$

[Strintzis, 1996] présente une généralisation aux dimensions supérieures des bases biorthogonales d'ondelettes.

1.1.5 Bilan

La notion d'analyse multirésolution a été définie pour un signal en une dimension, avec des décompositions orthogonales de l'espace en espaces d'approximations et de détails. Ceci a permis d'introduire l'algorithme de Mallat rapide et celui à trous qui permettent de décomposer un signal discrétisé. Ils restent valables dans les cadres plus généraux présentés par la suite. L'introduction des fonctions duales aux fonctions de base des espaces d'approximations et de détails permet de définir l'analyse multirésolution biorthogonale, cadre moins contraignant que celui de l'analyse multirésolution orthogonale et qui servira à la définition des ondelettes de seconde génération.

Ces notions se généralisent à la dimension deux sans difficulté dans la mesure où les notions présentées sont vectorielles. Cependant, la relation d'échelle n'est pas vectorielle et demande un travail pour être généralisée en deux dimensions. Ceci est présenté dans la section 1.2.

1.2 Présentation dans le cas d'un signal bidimensionnel

L'application des ondelettes aux images conduit à s'intéresser à l'extension de la théorie à la dimension deux. Nous nous plaçons maintenant dans l'espace $L^2(\mathbb{R}^2)$ des fonctions mesurables de carré intégrable sur \mathbb{R}^2 .

1.2.1 Matrice de changement d'échelle

Les sous-espaces V_j doivent être étendus à 2 dimensions. Pour cela, le facteur d'échelle discret j devient une matrice d'échelle \mathbf{J} permettant de passer d'une résolution à la suivante.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{J} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1.38)$$

Les pixels sont placés aux points de coordonnées entières, la matrice d'échelle \mathbf{J} doit donc transformer des coordonnées entières en coordonnées entières. \mathbf{J} doit donc avoir des coefficients entiers. De plus, la matrice doit assurer une dilatation selon toutes les dimensions, donc ses valeurs propres doivent être de module supérieur à 1. Le facteur de dilatation surfacique d'une échelle à la suivante est donné par $|\det(\mathbf{J})|$.

Le changement d'échelle dans le cas discret conduit à des sous-échantillonnages et des sur-échantillonnages des signaux définis sur des réseaux ou des treillis discrets. Notons Λ le treillis des vecteurs de \mathbb{Z}^2 . Ce réseau se transforme par changement d'échelle pour donner le sous-réseau $\Lambda_{\mathbf{J}}$ formés des vecteurs entiers de la forme :

$$\mathbf{m} = \mathbf{J}\mathbf{k} \text{ où } \mathbf{k} \in \mathbb{Z}^2 \quad (1.39)$$

Nous notons Γ_{v_i} l'ensemble obtenu par translation de $\Lambda_{\mathbf{J}}$ de vecteur v_i .

Il y a $|\det \mathbf{J}|$ ensembles Γ_{v_i} obtenus distincts pour chaque sous-réseau $\Lambda_{\mathbf{J}}$ (ce nombre dépend de la matrice de changement d'échelle \mathbf{J} choisie). Il y a donc $|\det \mathbf{J}|$ vecteurs entiers distincts conduisant à ces échantillonnages (translation nulle comprise).

1.2.2 Axiomatique de base

Elle est transposée directement de la dimension 1.

Définition 1.2.1 (Analyse multirésolution en 2 dimensions). *Une suite $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces fermés de $L^2(\mathbb{R}^2)$ est une approximation multirésolution si elle vérifie les 5 propriétés suivantes*

$$\forall j \in \mathbb{Z}, \forall \mathbf{k} \in \mathbb{Z}^2, f(\mathbf{x}) \in V_j \Leftrightarrow f(\mathbf{x} - 2^j \mathbf{k}) \in V_j, \quad (1.40)$$

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j, \quad (1.41)$$

$$\forall k \in \mathbb{Z}, f(\mathbf{x}) \in V_j \Leftrightarrow f(\mathbf{J}^{-1} \mathbf{x}) \in V_{j+1}, \quad (1.42)$$

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \quad (1.43)$$

$$\lim_{j \rightarrow -\infty} V_j = \text{Adhérence} \left(\bigcup_{j=-\infty}^{\infty} V_j \right) = L^2(\mathbb{R}^2). \quad (1.44)$$

La translation est faite avec un vecteur \mathbf{k} à coordonnées entières. La famille de fonctions dilatées et translatées construites à partir de la fonction mère sont données par :

$$\varphi_{j,n}(\mathbf{x}) = |\det \mathbf{J}|^{-\frac{j}{2}} \varphi(\mathbf{J}^{-j} \mathbf{x} - \mathbf{k}) \text{ avec } \mathbf{k} \in \mathbb{Z}^2, \mathbf{x} \in \mathbb{R}^2, j \in \mathbb{Z} \quad (1.45)$$

La construction des espaces de détail W_j est similaire à celle réalisée en une dimension à cela près que plusieurs signaux de détail sont obtenus : on considère le sous-espace W_j supplémentaire de V_{j-1} dans V_j , tel que

$$V_{j-1} = V_j \oplus W_j \quad \text{avec} \quad V_j \perp W_j \quad (1.46)$$

L'union d'une base orthonormée de V_j et d'une base orthonormée de W_j doit constituer une base orthonormée de V_{j-1} . On note

$$\psi_{j,\mathbf{k}}^i = |\det \mathbf{J}|^{-\frac{j}{2}} \psi^i(\mathbf{J}^{-j} \mathbf{x} - \mathbf{k}) \quad (1.47)$$

une famille orthonormée de W_j . Il y a donc $|\det \mathbf{J}| - 1$ signaux de détail

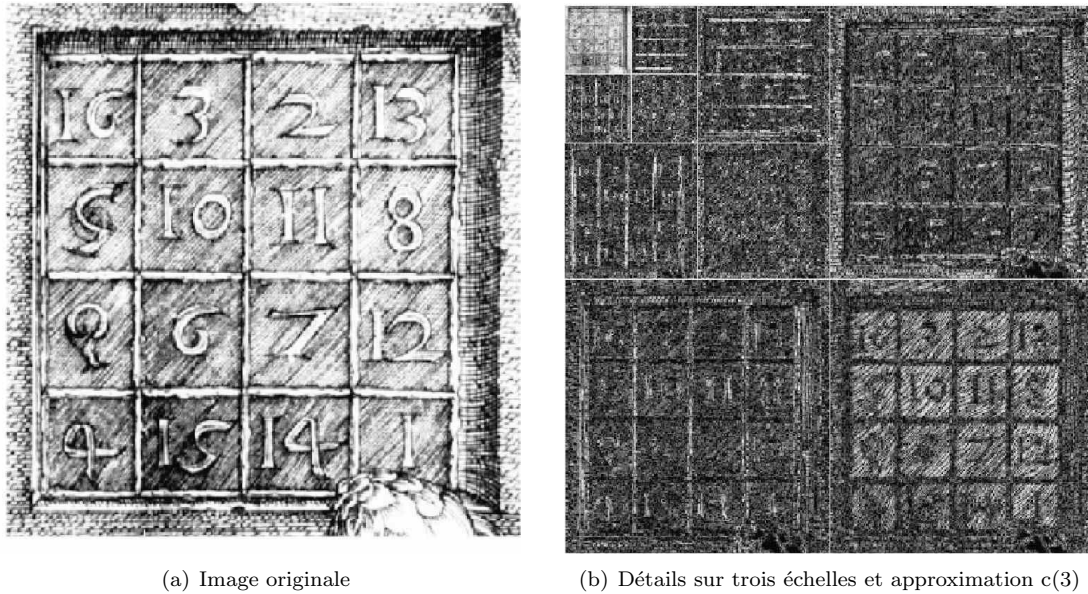
$$d^i(j, \mathbf{k}) = \langle f, \psi_{j,\mathbf{k}}^i \rangle \quad (1.48)$$

pour chaque signal d'approximation $c(j, \mathbf{k})$

Ainsi pour le cas le plus classique où $\mathbf{J} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, il y aura donc 3 signaux de détail.

1.3 Bilan

Le cadre classique de l'analyse multirésolution a été présenté en une puis deux dimensions. Il offre une décomposition de l'espace en approximations et détails. De plus, les changements d'échelles sont obtenus comme produit de convolution entre les coefficients du signal et ceux des filtres d'approximation et de détails. Mais cette simplicité est aussi une limitation pour les images binaires : une convolution entre un signal binaire et un filtre ne donne pas en général un résultat binaire. Comme nous souhaitons travailler sur une décomposition binaire, cela nous conduit à étudier l'analyse multirésolution non-linéaire qui a



(a) Image originale

(b) Détails sur trois échelles et approximation $c(3)$

FIG. 1.5 – Une image et sa décomposition en ondelettes de Haar 2d.

été développée pour sortir des limites de l'analyse classique. Elle permet, entre autre, d'obtenir une analyse multirésolution composée d'approximations et de détails binaires.

1.4 Analyses multirésolution non-linéaires

1.4.1 Cas général

Les ondelettes de première génération reposent sur la dilatation (d'une échelle à la suivante) et la translation (au sein d'une même échelle), opérations qui deviennent des opérations algébriques dans le domaine de Fourier. Ceci permet de construire des ondelettes en utilisant la transformation de Fourier (voir [Daubechies, 1992], etc). Mais elles ont leurs limites, en effet, la transformation de Fourier est bien adaptée aux fonctions définies sur un domaine régulier (\mathbb{R}^2 par exemple) mais réagit mal aux domaines irréguliers. Les ondelettes de première génération satisfont les propriétés suivantes :

P1 Les ondelettes forment une base de Riesz de $L^2(\mathbb{R})$.

P2 Les ondelettes sont soit orthogonales soit leur duales sont connues (dans le cas où elles sont biorthogonales).

P3 Les ondelettes et leurs duales sont localisées en temps et en fréquence. La localisation en fréquence provient de leur régularité et de leurs moments polynômiaux nuls.

P4 La transformation en ondelettes rapide qui permet d'obtenir les coefficients d'ondelettes en un temps linéaire.

Le développement des applications des ondelettes a amené la construction d'ondelettes sortant de ce cadre, on peut citer quatre raisons :

G1 Les problèmes réels sont définis avec un échantillonnage irrégulier alors que les ondelettes de première génération le sont avec un régulier.

G2 L'analyse sur des courbes ou des surfaces nécessitent des mesures pondérées alors que les ondelettes de première génération sont définies sur des espaces munis de mesures invariantes par translation.

G3 Alors que les ondelettes de première génération sont définies sur \mathbb{R}^n , des applications telles que la segmentation nécessite des domaines moins réguliers.

G4 Les données bidimensionnelles binaires nécessitent des filtres non-linéaires.

Le cadre de l'AMR de seconde génération présentée ci-dessous est assez souple pour permettre ces variations.

remarque 1.4.1. *Le dernier point (G4) nous intéresse particulièrement car les images que nous traitons sont binaires. Ceci explique le développement de cette partie dédiée au traitement non-linéaire.*

Les ondelettes de deuxième génération reprennent le principe de l'AMR biorthogonale. Elles le généralisent dans la mesure où les bases ne sont plus construites par dilatation-translation à partir d'une ondelette mère (outre le fait que l'espace de définition L_2 est plus général).

1.4.2 Analyse multirésolution ($2^{\text{ième}}$ génération)

Dans cette section est présentée la version de deuxième génération de l'AMR. Les notations sont identiques à celles utilisées dans la partie 1 pour l'AMR biorthogonale, même si leur sens peut changer. Par exemple, le terme « fonction d'échelle » est conservé bien que celle-ci ne puisse plus être écrite comme combinaison linéaire des versions d'elle-même à la résolution supérieure.

Considérons l'espace général $L_2(X, \Sigma, \mu)$ noté L_2 , où X représente le domaine spatial, Σ une σ -algèbre et μ une mesure non-atomique sur Σ . Cette mesure peut ne pas être invariante par translation. On suppose que (X, d) est un espace métrique. Nous nous plaçons donc dans le cadre de l'AMR biorthogonale avec les espaces d'approximation $(V_j)_{j \in \mathbb{Z}}$, leur espaces duaux $(\bar{V}_j)_{j \in \mathbb{Z}}$.

Par contre l'index de la base est plus général : pour tout j , V_j admet une base de Riesz fournie par les fonctions d'échelle $\{\varphi_{j,k}, k \in \mathcal{K}(j)\}$. On peut voir $\mathcal{K}(j)$ comme un ensemble d'indices général. On suppose que $\mathcal{K}(j+1) \subset \mathcal{K}(j)$.

Pour $f \in L_2$, les coefficients $c(j, k)$ sont alors définis par $c(j, k) = \langle f, \tilde{\varphi}_{j,k} \rangle$

1.4.3 Ondelettes ($2^{\text{ième}}$ génération)

Les ondelettes de première génération sont construites comme fonctions de base de l'espace complémentaire de V_{j+1} dans V_j . Cette idée reste à la base de la définition suivante :

Définition 1.4.1 (ondelettes de $2^{\text{ième}}$ génération). *La suite de fonctions $\{\psi_{j,m}, j \in \mathbb{Z}, m \in \mathcal{M}(j)\}$ avec $\mathcal{M}(j) = \mathcal{K}(j-1) \setminus \mathcal{K}(j)$ est un ensemble de fonctions d'ondelettes si*

- W_j , qui est le sous espace vectoriel fermé engendré par l'ensemble $\{\psi_{j,m}, m \in \mathcal{M}(j)\}$, est le supplémentaire de V_j dans V_{j-1} ,
- l'ensemble $\{\psi_{j,m}, j \in \mathbb{Z}, m \in \mathcal{M}(j)\}$ est une base de Riesz de L_2 .

La base duale est donnée par les ondelettes duales $\tilde{\psi}_{j,m}$ qui sont biorthogonales aux ondelettes

$$\langle \psi_{j,m}, \tilde{\psi}_{j,m'} \rangle = \delta_{k,m} \delta_{k',m'}. \quad (1.49)$$

Pour $f \in L_2$, on a encore $d(j, m) = \langle f, \tilde{\psi}_{j,m} \rangle$. Alors

$$f = \sum_{j,m} d(j, m) \psi_{j,m} \quad (1.50)$$

Leur définition fait que les ondelettes satisfont la relation d'échelle suivante :

$$\psi_{j+1,m} = \sum_l g[j, m, l] \psi_{j,l} \quad (1.51)$$

remarque 1.4.2. *Le filtre g dépend de l'échelle j mais aussi de l'indice de décalage m (qui était le pas de translation k pour les ondelettes de première génération).*

On suppose ces filtres à support finis, ce qui conduit aux ensembles uniformément bornés

$$\mathcal{M}(j, l) = \{m \in \mathcal{M}(j), g[j, m, l] \neq 0\} \quad \text{et} \quad \mathcal{L}(j, m) = \{l \in \mathcal{K}_{j-1}, m \in \mathcal{M}(j, l)\} \quad (1.52)$$

Ainsi, comme $\varphi_{j,l} \in V_{j+1} \oplus W_{j+1}$, il vient

$$\varphi_{j,l} = \sum_k \tilde{h}[j, k, l] \varphi_{j+1,k} + \sum_l \tilde{g}[j, m, l] \psi_{j+1,m}. \quad (1.53)$$

Supposons maintenant que les fonctions d'échelle et d'ondelettes puissent être définies, elles sont biorthogonales au sens suivant :

$$\langle \varphi_{j,k}, \tilde{\varphi}_{j,k'} \rangle = \delta_{k,k'} \quad \langle \psi_{j,m}, \tilde{\varphi}_{j,k} \rangle = 0 \quad (1.54)$$

$$\langle \psi_{j,m}, \tilde{\psi}_{j,m'} \rangle = \delta_{m,m'} \quad \langle \varphi_{j,k}, \tilde{\psi}_{j,m} \rangle = 0 \quad (1.55)$$

1.4.4 Transformation en ondelettes rapide

L'idée, encore une fois, reste la même que pour les ondelettes de première génération : calculer les coefficients d'ondelettes et d'approximation à partir d'un jeu de coefficients à une résolution donnée. La transformation en ondelettes rapide se présente comme suit :

$$c(j+1, k) = \sum_{l \in \tilde{\mathcal{L}}(j,k)} \tilde{h}[j, k, l] c(j, k) \quad \text{et} \quad d(j+1, m) = \sum_{l \in \tilde{\mathcal{L}}(j,m)} \tilde{g}[j, m, l] c(j, m) \quad (1.56)$$

La principale différence avec le schéma de la première génération provient du fait que les coefficients des filtres sont différents pour chaque coefficient du signal. La complexité de la transformation en ondelettes rapide de deuxième génération a été calculée, et dans le cas où les filtres sont à support fini, on a le résultat suivant :

Corollaire 1.4.1. *Dans le cas où les filtres g, h, \tilde{g} et \tilde{h} sont finis, la transformation en ondelettes rapide de deuxième génération est un algorithme en temps linéaire.*

1.4.5 Lifting scheme

Mis au point par Sweldens [Sweldens, 1995], le processus canonique de lifting se décompose en trois étapes, dénommées *partage*, *prédiction* et *mise à jour*. Le but ici est de présenter l'idée de chaque étape et ensuite d'appliquer le schéma aux ondelettes de Haar. Ce paragraphe est inspiré de [Sweldens, 1997].

Partage Lors de la première étape, les données c_0 sont partagées en deux sous-ensembles c_1 et d_1 . d_1 est appelé le sous-ensemble des ondelettes. Aucune contrainte n'est imposée sur la manière dont sont partagées les données, ni sur la taille

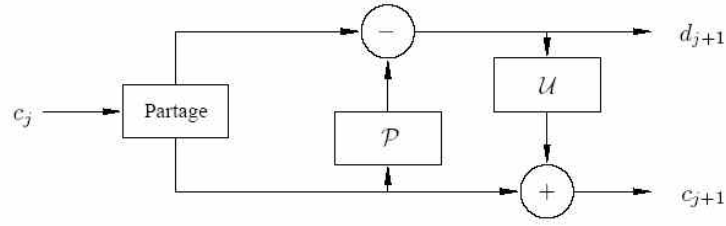


FIG. 1.6 – Schéma de l'analyse pour le lifting scheme.

relative de chacun des sous-ensembles. La seule nécessité est d'avoir une procédure pour obtenir à nouveau c_0 à partir de c_1 et d_1 . Le partage le plus simple consiste à couper simplement les données en deux parties disjointes, ce choix est appelé *les ondelettes paresseuses*.

Prédiction Le but général est d'obtenir une représentation de c_0 plus compacte. Pour ce faire, dans cette deuxième étape, c_1 va être utilisé pour prédire le sous-ensemble d_1 en se basant sur la corrélation présente dans les données originales. La construction de l'opérateur de prédiction repose typiquement sur un modèle des données qui reflète leur structure de corrélation. Bien entendu, l'opérateur de prédiction \mathcal{P} ne peut pas dépendre des données, car cela reviendrait à cacher de l'information dans \mathcal{P} .

En pratique, il n'est pas possible de prévoir exactement d_1 à partir de c_1 . Cependant $\mathcal{P}(c_1)$ peut être assez proche de d_1 . Il est alors envisageable de remplacer d_1 par la différence entre lui et sa valeur prédite $\mathcal{P}(c_1)$. Cette opération de différence (abstraite) est notée avec un signe $-$

$$d_1 \leftarrow c_1 - \mathcal{P}(c_1) \quad (1.57)$$

Le sous-ensemble d'ondelettes reflète maintenant de combien les données dévient du modèle sur lequel \mathcal{P} a été construit.

Cependant, on souhaite souvent conserver des propriétés globales des données originales dans les sous-ensembles d'approximation c_j . Par exemple, dans le cas des images, on souhaite conserver le niveau de gris moyen.

Mise à jour En se basant sur ce qui vient d'être fait, un opérateur de mise à jour \mathcal{U} est construit utilisant le sous-ensemble d'ondelettes d_1 pour conserver les propriétés de c_1 :

$$c_1 \leftarrow c_1 + \mathcal{U}(d_1). \quad (1.58)$$

L'algorithme d'analyse du lifting scheme est schématisée sur la figure 1.6. Ce schéma peut alors être itéré et conduit à l'algorithme de transformation en ondelettes 1.

Algorithm 1 Lifting scheme : transformation en ondelettes

```

for  $j = 1$  à  $n$  do
   $\{c_{j+1}, d_{j+1}\} = \text{partage}(c_j)$ 
   $d_{j+1} \leftarrow c_{j+1} - \mathcal{P}(c_{j+1})$ 
   $c_{j+1} \leftarrow c_{j+1} + \mathcal{U}(d_{j+1})$ 
end for

```

Une fois que cet algorithme est construit, l'algorithme de reconstruction s'obtient aisément, il suffit de permuter les signes $+$ et $-$, on obtient alors l'algorithme de reconstruction 2.

Algorithm 2 Lifting scheme : reconstruction

```
for  $j = n$  à  $1$  do  
   $c_j \leftarrow c_j + \mathcal{U}(d_j)$   
   $d_j \leftarrow c_j - \mathcal{P}(c_j)$   
   $c_{j-1} = \text{assembler}\{c_j, d_j\}$   
end for
```

1.5 Bilan

Ce chapitre a présenté les bases de l'analyse multirésolution sous sa forme classique, ainsi que les algorithmes qui en découlent (algorithmes de Mallat et à trous). Le cas des images est abordé au travers du cas en deux dimensions. Cependant, le cadre linéaire n'est pas adapté au traitement des images binaires. Les ondelettes biorthogonales présentées dans ce cadre servent de base aux AMR de deuxième génération qui sont non-linéaires.

La morphologie mathématique est particulièrement adaptée au traitement des images binaires. Il est intéressant d'avoir une AMR basée sur la morphologie mathématique pour décomposer nos images. Ce cadre existe sans être très développé. Pour l'introduire un point est fait sur le filtrage morphologique dans le chapitre suivant, puis le cadre de l'AMR morphologique est décrit. Ce dernier n'est pas extrêmement fouillé dans la littérature. Nous avons donc fait des études qualitatives concernant l'intérêt de l'algorithme à trous et la caractérisation de ces filtres qui sont non-linéaires à l'aide de la granulométrie. Cela est présenté dans la section [2.5.2](#).

2

Morphologie mathématique et analyse multirésolution

Fondé sur des notions ensemblistes et topologiques, le principe de la morphologie mathématique est d'étudier les caractéristiques morphologiques (forme, taille, orientation,...) des objets par des transformations non-linéaires associées à un objet de référence (élément structurant).

La morphologie mathématique a été développée initialement à l'École des mines de Paris. Elle repose essentiellement sur les travaux de G. Matheron (années 1960-70) puis sur ceux de J. Serra et de son équipe. Depuis, elle constitue un cadre de recherche actif et international.

La morphologie mathématique a été conçue à l'origine pour l'étude des milieux poreux ; elle est maintenant appliquée dans de nombreux domaines parmi lesquels la reconnaissance de formes [Bangham *et al.*, 1994], l'imagerie satellitaire [Soille, 2005], la multirésolution [Goutsias and Heijmans, 2000; Heijmans and Goutsias, 2000], la biologie [Angulo and Serra, 2003].

Elle s'appuie sur la théorie des ensembles, des treillis et la topologie. Les ouvrages [Serra, 1982; Serra, 1988; Soille, 2003] en font une présentation approfondie. Après avoir présenté des notions générales en 2.1, nous présentons les filtres morphologiques et les principales méthodes pour en construire en 2.2. Ensuite le cadre des ondelettes morphologiques est décrit section 2.3

2.1 Généralités sur la Morphologie mathématique

2.1.1 Définitions générales

Nous présentons ici quelques notions utiles pour la suite. Les trois premières sont généralement recherchées pour les opérateurs morphologiques et guident leur construction.

Dans toutes les définitions suivantes, Ψ désigne une transformation sur des ensembles ou sur des fonctions à valeurs réelles.

Définition 2.1.1 (Croissance). *Une transformation Ψ est croissante si :*

$$\forall X, Y, X \subset Y \Rightarrow \Psi(X) \subset \Psi(Y) \quad \text{ou} \quad \forall f, g, f \leq g \Rightarrow \Psi(f) \leq \Psi(g) \quad (2.1)$$

Définition 2.1.2 (Extensivité). Une transformation Ψ est extensive si :

$$\forall X, X \subset \Psi(X), \quad \text{ou} \quad \forall f, f \leq \Psi(f). \quad (2.2)$$

Elle est anti-extensive si :

$$\forall X, X \supset \Psi(X), \quad \text{ou} \quad \forall f, f \geq \Psi(f). \quad (2.3)$$

Définition 2.1.3 (Idempotence). Une transformation Ψ est idempotente si

$$\Psi \circ \Psi = \Psi \quad (2.4)$$

Définition 2.1.4 (Dualité). Deux transformations Ψ et Φ sont duales l'une de l'autre pour la complémentation si :

$$\forall X, \Psi(X^C) = \Phi(X)^C \quad (2.5)$$

où X^C désigne le complémentaire de X dans \mathbb{R}^n .

$$\forall f, \Psi(M - f) = M - \Phi(f) \quad (2.6)$$

et ce dans le cas où les fonctions sont positives bornées par M (cas des images).

Quatre opérations élémentaires

La morphologie mathématique repose sur une approche ensembliste des images. Les objets ou images binaires sont étudiés via leur relation avec un ensemble fixé, appelé élément structurant. L'élément structurant peut alors interagir avec l'image et donner une image transformée. Les transformations élémentaires ainsi définies sont l'érosion, la dilatation, l'ouverture et la fermeture. Pour leurs définitions et leurs propriétés, voir [Serra, 1988; Soille, 2003]. Leur action est illustrée par la fig. 2.1.

2.2 Filtres morphologiques

En traitement du signal, un filtre est habituellement défini comme une opération linéaire invariante par translation. Ce type de filtre est bien adapté aux problèmes engendrés par des phénomènes linéaires. Les filtres morphologiques sont des filtres non-linéaires prévus pour deux autres types de filtrage. Premièrement ils peuvent être utilisés pour débruiter des images dégradées par certains types de bruits. Leur principal atout est de pouvoir conserver les contours, ce qui reste difficile pour le filtrage linéaire. Deuxièmement, ils peuvent être utilisés pour retirer de manière sélective de l'image des structures ou des objets tout en en préservant d'autres. C'est ce dernier point qui nous intéresse : un filtre qui puisse conserver les traits principaux dans une image binaire et en supprimer les détails. Ce filtre sera utilisé dans le cadre d'une analyse multirésolution morphologique dont le cadre théorique est présenté dans la partie suivante. Nous présentons ici les caractéristiques des filtres morphologiques et leur techniques de construction. Ces filtres seront ensuite utilisés dans le cadre de l'AMR morphologique. La référence pour ce paragraphe est [Soille, 2003].

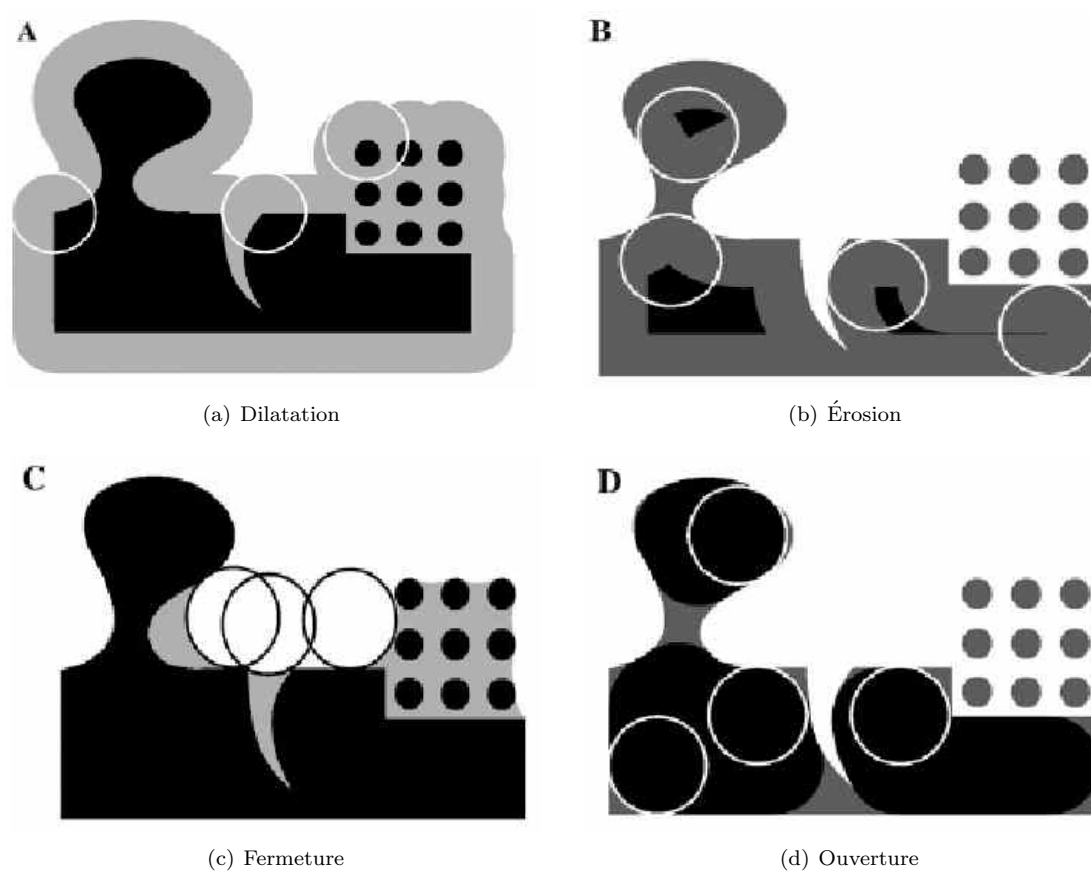


FIG. 2.1 – Illustration des différents filtres présentés. Les zones en gris clair sont ajoutées et celle en gris foncé sont retirées.

2.2.1 Définition

L'idée simple qui sous-tend tout filtre morphologique est de supprimer des structures de l'image, considérées comme gênantes. Mais les structures préservées ne doivent alors pas être modifiées lors d'une éventuelle seconde application : le filtre doit être *idempotent*. Le filtre doit aussi préserver la relation d'ordre entre les images. Ce qui conduit à la définition suivante :

Définition 2.2.1 (filtre morphologique). Ψ est un filtre morphologique $\Leftrightarrow \Psi$ est croissant et idempotent.

Comme nous l'avons déjà vu, l'érosion et la dilatation sont des filtres morphologiques.

2.2.2 Construction de filtres morphologiques

De nouveaux filtres peuvent être construits en combinant des filtres élémentaires, cependant toutes les combinaisons ne conduisent pas à un filtre morphologique. Par exemple, la composition de deux ouvertures ayant un élément structurant différent n'est généralement pas une ouverture ni un filtre.

En pratique, il existe trois méthodes : la combinaison parallèle, séquentielle et itérative.

Combinaison parallèle Une propriété simple peut être exploitée pour construire une nouvelle ouverture à partir d'ouvertures : on applique chaque ouverture sur l'image initiale (parallèlement) puis on prend la réunion des images résultantes de la série d'ouverture. La transformation obtenue est encore une ouverture et la propriété duale est vraie pour les fermetures :

1. Toute union (au sens de valeur maximale au point) d'ouverture est une ouverture,
2. toute intersection (au sens de valeur minimale au point) de fermeture est une fermeture.

Ce type de combinaison s'adapte bien aux images avec des structures directionnelles en utilisant des segments de ligne comme élément structurant.

Combinaison séquentielle On a déjà vu que la composition de deux ouvertures n'est généralement pas une ouverture. Cependant, la composition de deux filtres ordonnés est encore un filtre ordonné. La paire de filtres ordonnés utilisée est souvent une ouverture et la fermeture duale. Ce qui permet d'obtenir un effet similaire sur les parties noires et sur les parties blanches de l'image. Pour les nouveaux filtres ainsi obtenus, on a la relation d'ordre suivante :

$$\gamma \leq \gamma\phi\gamma \leq \gamma\phi, \phi\gamma \leq \phi\gamma\phi \leq \phi. \quad (2.7)$$

où γ est une ouverture et ϕ une fermeture. Le produit d'ouvertures et de fermetures est aussi la base des filtres séquentiels alternés présentés ci-dessous.

Combinaison itérative Pour une ouverture donnée ψ , les ouvertures par itération sont obtenues par itération de la transformation $\lambda = \text{id} \vee \psi$. On peut prouver que

- λ est extensive,
- que $\lambda^{(n)} \leq \lambda^{(n+1)}$,
- que pour tout n , $\lambda^{(n)}$ est plus petite que la plus petite des fermetures plus grandes que ψ , notée $\hat{\phi}$:

$$\hat{\phi} = \bigwedge_{\phi \text{ ouverture}} \{\phi/\psi \leq \phi\}. \quad (2.8)$$

Concernant les images discrètes à support borné, la transformation $\hat{\phi}$ est atteinte pour un nombre fini d'itérations :

$$\hat{\phi} = \lambda^{(i)} \quad (2.9)$$

Les ouvertures par itération sont obtenues par dualité.

2.2.3 Filtres alternés séquentiels

Pour une image très bruitée, dans le sens où il y a du bruit à plusieurs échelles, l'utilisation d'une combinaison séquentielle de filtres avec un élément structurant assez grand ne donne pas des résultats satisfaisants. Il faut alors employer des éléments structurant de taille différente, ce qui conduit aux *filtres séquentiels alternés*.

Définition 2.2.2. Soient γ_i une ouverture et ϕ_i la fermeture duale de taille i . D'après ce qui a été vu ci-dessus, les quatre combinaisons séquentielles suivantes sont des filtres morphologiques :

$$m_i = \gamma_i \phi_i \quad r_i = \phi_i \gamma_i \phi_i, \quad (2.10)$$

$$n_i = \phi_i \gamma_i \quad s_i = \gamma_i \phi_i \gamma_i. \quad (2.11)$$

Un filtre alterné séquentiel est défini comme la combinaison séquentielle de l'un de ces filtres, en commençant la séquence avec un filtre de taille 1 et en la terminant avec un filtre de taille i

$$M_i = m_i \cdots m_1 \quad R_i = r_i \cdots r_1, \quad (2.12)$$

$$N_i = n_i \cdots n_1 \quad S_i = s_i \cdots s_1. \quad (2.13)$$

On peut montrer que les filtres alternés séquentiels sont des filtres morphologiques. Bien que les filtres séquentiels alternés ne soient pas auto-duaux, leur action est plus symétrique que celles des ouvertures et fermetures. Ils sont utilisés particulièrement en débruitage.

2.2.4 Filtres auto-duaux

La plupart des filtres morphologiques se présentent sous forme de paires de filtres duaux dont l'application dépend du fait que l'objet à traiter dans l'image est plus clair ou plus foncé que le fond. Cependant le traitement des images, pour lesquelles la distinction entre le fond et les objets (par exemple, les images satellitaires) n'est pas claire, est délicat avec ce type de filtre. D'autre part, dans le cadre de l'AMR, le filtre est appliqué alternativement avec un sous-échantillonnage. L'idempotence n'a plus d'effet, puisqu'un sous-échantillonnage est effectué entre deux applications du filtre, et si le filtre n'est pas symétrique, l'image aura tendance à s'assombrir ou à s'éclaircir selon que l'un ou l'autre des deux filtres duaux est utilisé. Lorsqu'un comportement strictement symétrique est requis par l'application, il faut faire appel à la notion de filtre auto-dual. Les filtres auto-duaux sont communs dans le traitement de l'image non-morphologique. La convolution normalisée en est un. Le filtre médian est un filtre non-linéaire auto-dual dont il existe une version morphologique qui est présentée dans la partie sur l'analyse multirésolution morphologique.

Définition 2.2.3 (filtre auto-dual). Un filtre Ψ est auto-dual pour la complémentation si

$$\forall X, \Psi(X^C) = \Psi(X)^C. \quad (2.14)$$

2.2.5 Bilan

Dans cette section, la notion de filtre morphologique a été définie et trois modes de combinaison (parallèle, séquentielle et alternée) ont été proposés pour ces filtres qui permettent de conserver certaines propriétés. La classe des filtres alternés séquentiels a été introduite. Enfin, la propriété d'auto-dualité qui est importante en traitement de l'image et nous le verrons, dans le cadre de l'AMR, a été définie.

Les filtres permettent de filtrer de manière itérative les images et en particulier les images binaires. Par contre, ils ne permettent pas à eux seuls d'effectuer une décomposition de l'image en approximation et détail et de savoir ce qui est perdu à chaque itération. Pour avoir un contrôle de la perte, il faut se placer dans le cadre de l'AMR. La section suivante présente les filtres morphologiques dans le cadre de l'AMR.

2.3 Analyse multirésolution morphologique

Cette section présente une version morphologique des ondelettes et du schéma de l'Analyse Multirésolution inspirée de [Heijmans and Goutsias, 2000]. Nous donnerons ensuite la version morphologique de l'ondelette de Haar en une et deux dimensions.

2.3.1 Décomposition en ondelettes couplée

Supposons qu'il existe des ensembles $(V_j)_{j \in \mathbb{Z}}$ et $(W_j)_{j \in \mathbb{Z}}$. V_j est appelé *l'espace d'approximation de niveau j* et W_j *l'espace de détail de niveau j* . Nous prenons les notations suivantes :

$$h_a : V_j \rightarrow V_{j+1} \quad \text{opérateur d'analyse pour l'approximation,} \quad (2.15)$$

$$g_a : V_j \rightarrow W_{j+1} \quad \text{opérateur d'analyse pour les détails,} \quad (2.16)$$

$$\mathbf{s} : V_{j+1} \times W_{j+1} \rightarrow V_j \quad \text{opérateur de synthèse} \quad (2.17)$$

Ces opérateurs doivent conduire à une représentation complète du signal dans le sens où les fonctions $(h_a, g_a) : V_j \rightarrow V_{j+1} \times W_{j+1}$ et $s : V_{j+1} \times W_{j+1} \rightarrow V_j$ doivent être inverses l'une de l'autre. Ce qui donne la condition suivante

$$\forall x \in V_j, \quad \mathbf{s}(h_a(x), g_a(x)) = x \quad \text{condition de reconstruction parfaite,} \quad (2.18)$$

et

$$\begin{cases} \forall (x, y) \in V_{j+1} \times W_{j+1}, h_a(\mathbf{s}(x, y)) = x, \\ \forall (x, y) \in V_{j+1} \times W_{j+1}, g_a(\mathbf{s}(x, y)) = y. \end{cases} \quad (2.19)$$

Ces deux conditions dans (2.19) garantissent que la décomposition n'est pas redondante.

remarque 2.3.1. *L'opération de synthèse n'est pas explicitée, contrairement au cas découpé (ci-dessous) où elle est réalisée comme une addition entre l'approximation et les détails.*

2.3.2 Décomposition en ondelettes découpée

Les notations sont les mêmes que dans le paragraphe précédent. h_a et g_a sont définis de la même manière, cependant \mathbf{s} peut être découpée : il existe deux opérateurs $h_s : V_{j+1} \rightarrow V_j$ et $g_s : W_{j+1} \rightarrow V_j$ et une opération $+$ sur V_j (qui peut varier

avec l'indice j) tels que

$$\forall (x, y) \in V_{j+1} \times W_{j+1}, \mathbf{s}(x, y) = h_s(x) + g_s(y), \quad (2.20)$$

alors les équations de reconstruction parfaite et de non-redondance (2.18) et (2.19) peuvent s'écrire :

$$\forall x \in V_j, h_s(h_a(x)) + g_s(g_a(x)) = x, \quad (2.21)$$

$$\begin{cases} \forall (x, y) \in V_{j+1} \times W_{j+1}, h_a(h_s(x)) = x, \\ \forall (x, y) \in V_{j+1} \times W_{j+1}, g_a(g_s(y)) = y. \end{cases} \quad (2.22)$$

2.3.3 Ondelettes de Haar morphologiques en 1d

La différence majeure avec les ondelettes de Haar linéaires en 1d provient du remplacement du filtre d'analyse par une érosion (ou une dilatation). Les concepts de la morphologie mathématique sont présentés dans la partie précédente. Nous présentons ici une version binaire de l'AMR, on peut aussi l'étendre aux images en niveau de gris en utilisant les filtres morphologiques correspondant. Soient V_j et W_j égaux au treillis des fonctions de \mathbb{Z} dans $\{0, 1\}$: $V_j = W_j = \{0, 1\}^{\mathbb{Z}}$. On se place dans le cadre des ondelettes découplées avec comme addition sur V_j le OU exclusif noté Δ . Les opérateurs d'analyse et de synthèse sont alors définis comme suit :

$$h_a(x)(n) = x(2n) \quad (2.23)$$

$$g_a(x)(n) = x(2n) \Delta x(2n + 1) \quad (2.24)$$

$$h_s(x)(2n) = x(n) \quad \text{et} \quad h_s(x)(2n + 1) = x(n) \quad (2.25)$$

$$g_s(y)(2n) = 0 \quad \text{et} \quad g_s(y)(2n + 1) = y(n) \quad (2.26)$$

Notons que le signal de détail n'est égal à 1 que lorsqu'une transition dans le signal x a lieu à un point pair.

2.3.4 Ondelettes de Haar morphologiques en 2d

Il est possible d'étendre le schéma de décomposition des ondelettes de Haar morphologiques à deux dimensions et plus de manière séparable (en appliquant séquentiellement la décomposition aux colonnes puis aux lignes de l'image 2d) [Mallat, 2000]. Il est aussi possible d'en donner une version non-séparable que nous présentons ici, toujours dans le cadre de la décomposition découplée.

Soit V_j le treillis des fonctions de \mathbb{Z}^2 dans $\{0, 1\}$ et W_j l'ensembles des fonctions de \mathbb{Z}^2 dans $\{0, 1\}^3$. Nous introduisons les notations suivantes $\mathbf{n}, \mathbf{2n}$ désignent les points (m, n) et $(2m, 2n) \in \mathbb{Z}^2$ respectivement, et $\mathbf{2n}_+, \mathbf{2n}^+, \mathbf{2n}_+^\dagger$ les points $(2m + 1, n)$, $(2m, 2n + 1)$ et $(2m + 1, 2n + 1)$, respectivement. Les opérateurs d'analyse sont définis comme suit :

$$h_a(x)(\mathbf{n}) = x(\mathbf{2n}) \wedge x(\mathbf{2n}_+) \wedge x(\mathbf{2n}^+) \wedge x(\mathbf{2n}_+^\dagger) \quad (2.27)$$

$$g_a(x)(\mathbf{n}) = (d_v(x)(\mathbf{n}), d_h(x)(\mathbf{n}), d_d(x)(\mathbf{n})) \quad (2.28)$$

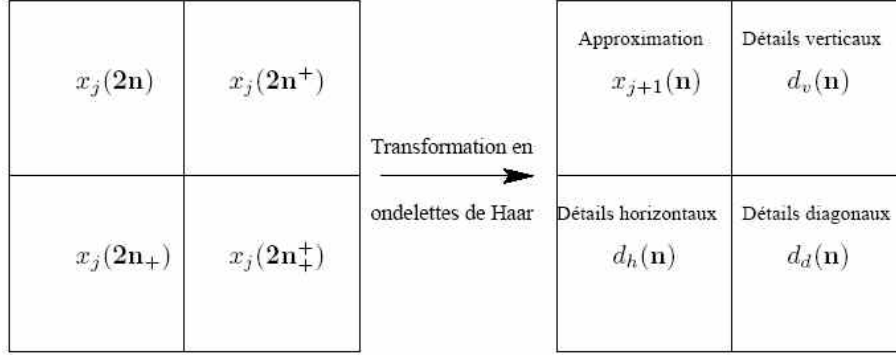


FIG. 2.2 – Schéma de l'analyse des ondelettes de Haar morphologiques en 2d : le signal d'entrée est transformé en un signal d'approximation, et des signaux de détails horizontaux, verticaux et diagonaux.

où d_v, d_h, d_d désignent les opérateurs de détail vertical, horizontal et diagonal, qui sont définis par :

$$d_v(x)(\mathbf{n}) = \frac{1}{2}(x(\mathbf{2n}) - x(\mathbf{2n}^+) + x(\mathbf{2n}_+) - x(\mathbf{2n}_+^+)), \quad (2.29)$$

$$d_h(x)(\mathbf{n}) = \frac{1}{2}(x(\mathbf{2n}) + x(\mathbf{2n}^+) - x(\mathbf{2n}_+) - x(\mathbf{2n}_+^+)), \quad (2.30)$$

$$d_d(x)(\mathbf{n}) = \frac{1}{2}(x(\mathbf{2n}) - x(\mathbf{2n}^+) - x(\mathbf{2n}_+) + x(\mathbf{2n}_+^+)). \quad (2.31)$$

Le schéma de la décomposition en ondelettes de Haar est présenté sur la figure 2.2. Les opérateurs de synthèse sont alors donnés par

$$h_s(x)(\mathbf{2n}) = h_s(x)(\mathbf{2n}_+) = h_s(x)(\mathbf{2n}^+) = h_s(x)(\mathbf{2n}_+^+) = x(\mathbf{n}), \quad (2.32)$$

et

$$g_s(y)(\mathbf{2n}) = (y_v(\mathbf{n}) + y_h(\mathbf{n})) \vee (y_v(\mathbf{n}) + y_d(\mathbf{n})) \vee (y_h(\mathbf{n}) + y_d(\mathbf{n})) \vee 0, \quad (2.33)$$

$$g_s(y)(\mathbf{2n}_+) = (y_v(\mathbf{n}) - y_h(\mathbf{n})) \vee (y_v(\mathbf{n}) - y_d(\mathbf{n})) \vee (-y_h(\mathbf{n}) - y_d(\mathbf{n})) \vee 0, \quad (2.34)$$

$$g_s(y)(\mathbf{2n}^+) = (y_v(\mathbf{n}) - y_h(\mathbf{n})) \vee (-y_v(\mathbf{n}) - y_d(\mathbf{n})) \vee (y_h(\mathbf{n}) - y_d(\mathbf{n})) \vee 0, \quad (2.35)$$

$$g_s(y)(\mathbf{2n}_+^+) = (-y_v(\mathbf{n}) - y_h(\mathbf{n})) \vee (y_v(\mathbf{n}) - y_d(\mathbf{n})) \vee (-y_h(\mathbf{n}) + y_d(\mathbf{n})) \vee 0, \quad (2.36)$$

où $y = (y_v, y_h, y_d)$. Il n'est pas difficile de montrer que les conditions (2.21) et (2.22) sont vérifiées.

2.3.5 Bilan

L'AMR morphologique a été présentée dans cette section, sous sa forme couplée, quand l'opérateur de synthèse n'est pas explicite, et sous sa forme découplée quand il est une somme entre approximation et détail. Puis L'AMR morphologique est illustrée par un exemple de AMR morphologique de Haar où l'opérateur d'analyse, qui dans le cas linéaire est une moyenne, est remplacée ici par un minimum. Ces résultats peuvent être adaptés au cas des images binaires.

Cependant, la plupart du temps, les approximations résultant d'une AMR morphologique sont noircies ou éclaircies ou de mauvaise qualité. La section 2.5 étudie le choix de l'AMR morphologique. Le cas de l'AMR morphologique basée sur le

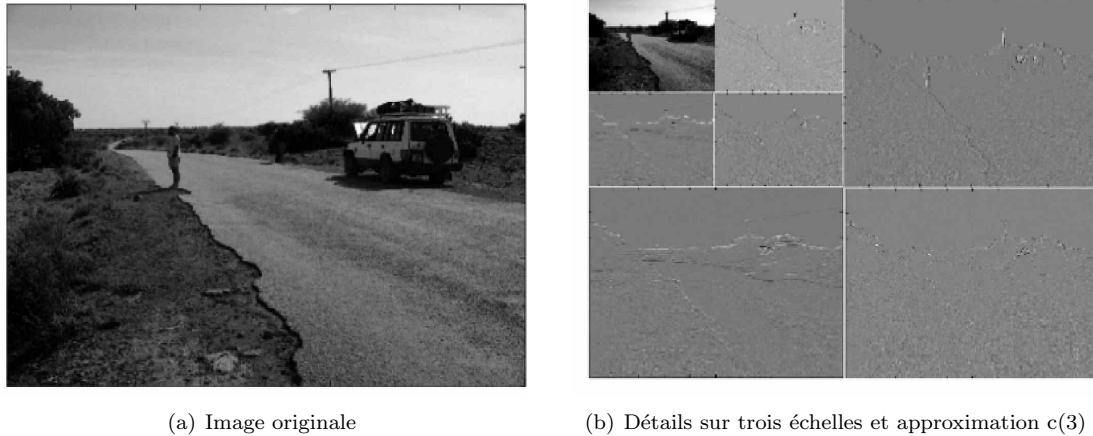


FIG. 2.3 – Décomposition en ondelettes de l'image de gauche avec les ondelettes de Haar 2d morphologiques.

filtre de la médiane dans le cas binaire y est exposé. Mais avant cela, une méthode d'étude de mathématique morphologique pouvant servir pour les filtres morphologiques est présentée section 2.4.

2.4 Méthodes de caractérisation morphologiques

La morphologie mathématique a développé des outils pour l'étude des matériaux poreux ou granuleux qui ont donné naissance à des méthodes de traitement de l'image telle que la granulométrie [Matheron, 1967; Laÿ, 1987] ou l'autocovariance morphologique [Soille, 2003]. Nous les présentons ici car ils ont fait partie de nos études sur la caractérisation des filtres non-linéaires et sur la classification des Cartes de Dissimilarités Locales (ces cartes -CDL- proviennent de la comparaison de deux images binaires et sont présentées dans la partie 2). Les méthodes présentées sont détaillées dans [Soille, 2003].

2.4.1 Granulométrie

Principe

L'analyse d'un matériau granuleux se fait en tamisant des échantillons à l'aide de tamis de taille croissante tout en pesant la masse retenue par chaque tamis. Il est notable que cette méthode ait toutes les propriétés d'une ouverture :

anti-extensivité : ce qui reste dans le tamis est une partie de l'échantillon d'origine,

croissance : le tamisage d'une partie d'un échantillon laisse dans le tamis une partie de ce qui aurait dû rester avec le tamisage de l'échantillon entier,

idempotence réitérer le tamisage avec le même tamis ne change pas le résultat.

Cependant comme la granulométrie implique un série de tamis, elle satisfait une propriété plus forte que l'idempotence qui est appelée *propriété d'absorption* : ce qui reste dans le tamis après avoir tamisé avec deux tamis de taille arbitraire résulte uniquement de la taille du tamis le plus gros. En conséquence, la transposition de la granulométrie au traitement de l'image en appliquant un famille d'ouvertures successives de taille croissante λ ne peut se faire que si la propriété d'absorption est vérifiée :

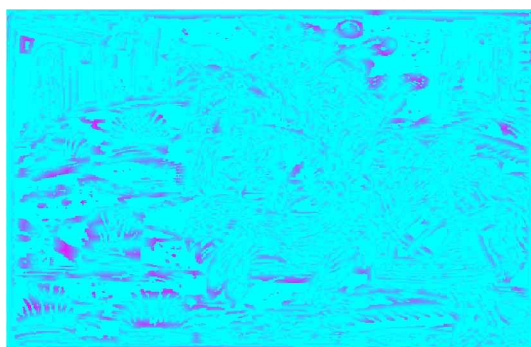
$$\gamma_{\lambda_i} \gamma_{\lambda_j} = \gamma_{\lambda_j} \gamma_{\lambda_i} = \gamma_{\max(\lambda_i, \lambda_j)} \quad (2.37)$$

Le processus est alors appelé *granulométrie par ouverture*. Les éléments structurant utiles dans la pratique et qui vérifient ces propriétés sont les disques et les segments de droite de taille croissante.

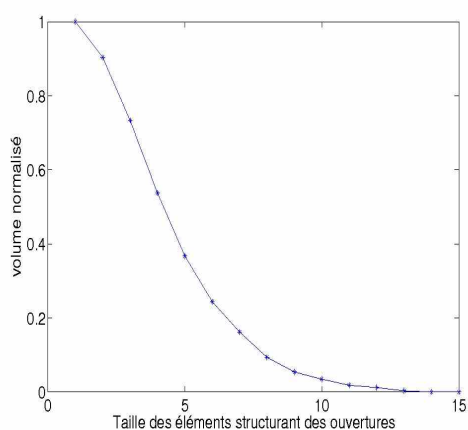
Description

Les *courbes granulométriques* donnent la somme des valeurs des pixels de l'image ouverte en fonction de la taille de l'ouverture. En pratique, ces courbes sont souvent interprétées en calculant leur dérivée discrète. La courbe résultante est appelée *courbe de taille* ou *spectre de forme* car son maximum indique les tailles prédominantes dans l'image.

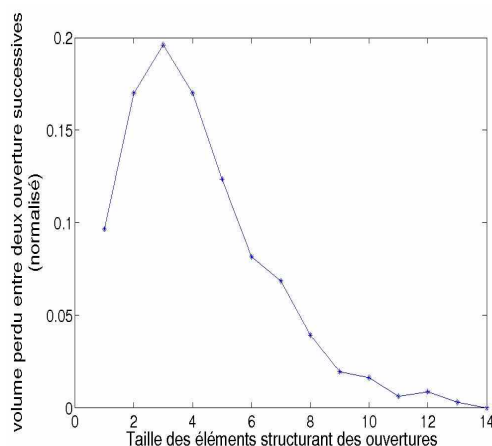
La granulométrie par ouverture donne des informations concernant les structures de l'image plus claires que leur voisinage. Pour obtenir l'information concernant les structures plus foncées, la même démarche est adoptée en remplaçant l'ouverture par la fermeture avec la même famille de filtres. Le procédé est alors appelé *granulométrie par fermeture*.



(a) Image originale (Carte de dissimilarité entre deux impressions anciennes en fausses couleurs)



(b) Courbe granulométrique, les indices négatifs correspondent à des fermetures.



(c) Dérivée de la courbe granulométrique (courbe de taille).

FIG. 2.4 – Une image et son analyse granulométrique par ouverture. Le volume initial est situé en $x=0$ et vaut 1. L'ouverture diminue le volume de l'image jusqu'à 0 et la fermeture l'augmente. La dérivée indique les tailles des structures prépondérantes dans l'image.

Cas des images binaires

La granulométrie par ouverture pour les images binaires peut être accélérée en calculant la *fonction d'ouverture* F_O de l'image binaire en entrée \mathbf{B} . La fonction d'ouverture attribuée à chaque pixel x de l'image binaire \mathbf{B} la valeur de la taille du filtre nécessaire pour le retirer de l'image :

$$F_O(\mathbf{B})(x) = \max\{\lambda, x \in \gamma_\lambda(\mathbf{B})\}. \quad (2.38)$$

Alors, l'histogramme de la fonction d'ouverture est équivalent à la dérivée de la courbe granulométrique de l'image binaire. Cette technique est détaillée dans [Laÿ, 1987] ainsi qu'un algorithme de calcul rapide, similaire à ceux utilisés pour calculer la transformée en distance.

Granulométrie pour évaluer les filtres

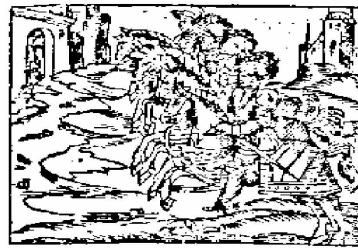
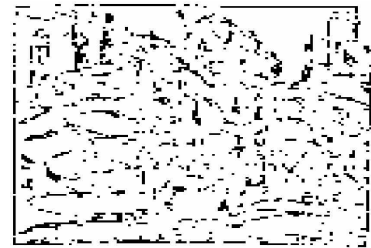
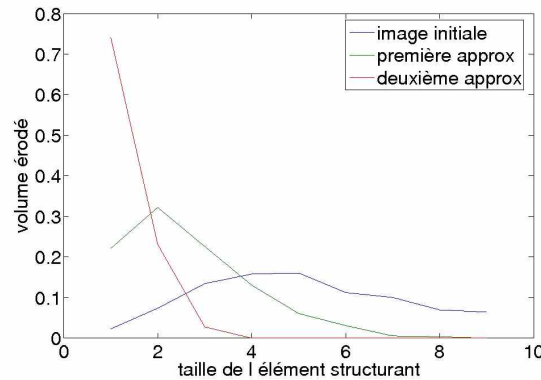
Les filtres linéaires sont caractérisés entièrement par leur réponse impulsionnelle. Heuristiquement, cette propriété est liée à la linéarité qui permet de prévoir leur comportement pour tout signal en écrivant ce signal comme combinaison linéaire d'impulsions. La transformée de Fourier du filtre permet alors de prévoir l'action du filtre dans le domaine fréquentiel (passe-bas, passe-haut, etc.).

Il en est autrement pour les filtres non-linéaires : la réponse impulsionnelle, même si elle est calculable, n'apporte pas d'information sur le comportement du filtre pour un signal autre qu'une impulsion. Il est donc plus difficile d'appréhender leur comportement tant dans le domaine spatial que dans le domaine fréquentiel. Pour ce dernier point, dans le cas des images binaires, la difficulté est encore augmentée par le fait que la représentation fréquentielle n'est pas adaptée aux discontinuités (transitions 0-1) dont sont composées les images. Il existe des propriétés caractérisant leur comportement, telles que l'auto-dualité, qui cadrent leur comportement mais ne permettent pas de l'identifier précisément. Cependant, même pour les filtres non-linéaires, l'expérience permet de cerner leur comportement. Il est donc envisageable de le caractériser en pratique. Dans le cas linéaire, la décomposition fréquentielle sépare les basses fréquences (ce qui est grossier) des hautes fréquences (ce qui est fin). La granulométrie apporte -dans une certaine mesure- des éléments comparables : elle permet grâce au principe de tamisage de séparer ce qui est fin de ce qui est grossier, avec la limitation inhérente au choix de la famille d'ouvertures (ou de fermeture). Elle n'a pas l'assise théorique de la transformation de Fourier, ni les propriétés liées à la structure d'espace vectoriel telles que la décomposition sur une base orthogonale, l'équivalence des énergies dans le domaine spatial et dans le fréquentiel. Néanmoins, c'est une piste intéressante que nous avons en partie étudiée. L'étude est présentée dans la sous-section suivante.

Utilisation de la granulométrie pour l'étude des filtres

Il est possible de déterminer les tailles de structures prépondérantes d'une images avec la dérivée granulométrique. Pour caractériser un filtre, il est intéressant d'étudier le changement de ces tailles entre deux échelles. Les orientations de nos recherches n'ont pas permis d'approfondir cette étude, nous en faisons une présentation rapide. Nous présentons la dérivée granulométrique des approximations de l'AMR pour les filtres morphologiques de l'érosion (fig. 2.5), de Haar (fig. 2.6) et médian (fig. 2.7). L'illustration est faite sur l'image originale et deux échelles.

Commentaires Nous donnons premièrement une interprétation générale qui servira ensuite à commenter les courbes de dérivées granulométriques obtenues.

(a) Image binaire originale (512×512).(b) Approximation à l'échelle 1 par l'AMR morphologique de l'érosion (256×256)(c) Approximation à l'échelle 2 par l'AMR morphologique de l'érosion (128×128)

(d) Dérivées granulométriques des trois approximations de l'AMR morphologique de l'érosion

FIG. 2.5 – AMR morphologique de l'érosion décimée : les approximations et leurs dérivées granulométriques.

Les dérivées granulométriques indiquent le volume perdu en passant à l'élément structurant supérieur. Les maxima indiquent les tailles prépondérantes dans l'images.

Pour les trois figures 2.5, 2.6 et 2.7, les approximations ont été réalisées avec une AMR décimée, et donc les images sont deux fois plus petites en passant à d'une échelle à la suivante. Comme la taille des éléments structurants reste la même, ils sont en proportion par rapport aux approximations deux fois plus gros en passant à l'échelle suivante. Ainsi, pour un filtre qui conserve les structures de l'image, la taille prépondérante dans l'image doit diminuer environ de deux en passant à l'échelle suivante.

D'autre part, la hauteur du maximum indique la proportion d'éléments de cette taille dans l'image. Si ce maximum augmente, c'est que les autres tailles disparaissent.

fig. 2.5 Les courbes ne présentent qu'un seul maximum dont l'abscisse se rapproche de 0 : à l'échelle 0, l'abscisse du maximum vaut 4-5 ; à l'échelle 1, l'abscisse du maximum vaut 2 et à l'échelle 2, l'abscisse du maximum vaut environ 1. Par contre, le pic a une valeur de plus en plus grande, elle passe de 0,15 à 0,3 puis à 0,8. Ceci traduit le fait que seuls les éléments de la taille prépondérante subsistent, les autres disparaissent. L'érosion (avec une décimation) est donc une opération qui diminue la taille des structures tout en brisant la proportion entre les éléments de différentes tailles. L'image est donc simplifiée pour arriver à une image qui ne contient que des éléments de petite taille : elle est alors difficilement interprétable.

fig. 2.6 Dans le cas de Haar, les pics restent autour de 3-5, c'est donc que les éléments prépondérants grossissent puisque

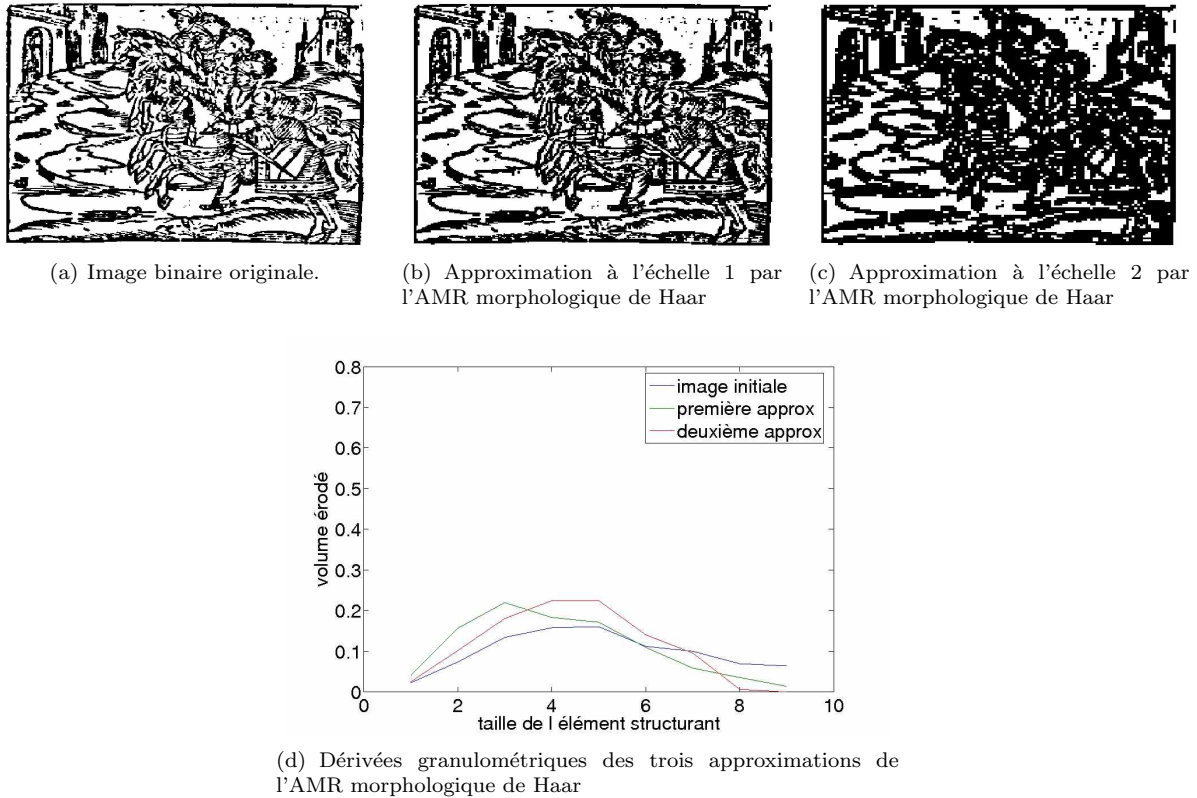


FIG. 2.6 – AMR morphologique de Haar décimée : les approximations et leurs dérivées granulométriques.

l'abscisse des pics devraient diminuer suite au changement d'échelle. D'autre part, la valeur des pics augmente peu, elle reste autour de 0,2 : toutes les tailles bénéficient de l'augmentation de taille. L'image se noircit mais la proportion entre les différentes tailles est mieux conservée. La qualité visuelle est assez bonne.

fig. 2.7 Dans le cas de la médiane, les pics se décalent normalement : la taille prépondérante est respectée. La valeur des pics augmente moyennement : elle passe de 0,15 à 0,25 puis à 0,4. Les structures se concentrent un peu autour de la taille prépondérante, ce qui est normale puisque la décomposition est décimée. Mais cette tendance n'est pas très marquée. Cela reflète la qualité visuelle de l'image qui est meilleure que les précédentes.

Commentaires et pistes La granulométrie permet d'avoir une idée de la manière dont les structures sont transformées par le filtre. Cependant, cette information est limitée par la forme de l'élément structurant (ES). Par exemple, l'action du filtre sur les « traits » de l'image ne sera pas prise en compte par la granulométrie avec un ES carré ou en forme de disque. La granulométrie représente le nombre de points noirs de l'image restant en fonction de la taille de l'ES qui a servi au filtrage. Ou encore, dans le cas de la granulométrie par érosion, le nombre de points noirs de l'images pour lesquels l'ES forme un voisinage dans l'ensemble des points noirs de l'image. Une piste à suivre pour s'affranchir d'un ES est de représenter l'histogramme de la fonction qui à un point noir x associe le nombre de points qu'il y a dans son voisinage V_x . On peut choisir pour V_x le disque unité pour la norme L_1 (connexité 4) ou pour la norme L_∞ (connexité 8). Cette piste reste à étudier.



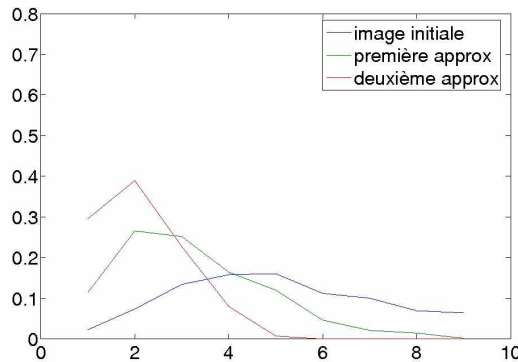
(a) Image binaire originale.



(b) Approximation à l'échelle 1 par l'AMR morphologique de la médiane



(c) Approximation à l'échelle 2 par l'AMR morphologique de la médiane



(d) Dérivées granulométriques des trois approximations de l'AMR morphologique de la médiane

FIG. 2.7 – AMR morphologique de la médiane décimée : les approximations et leurs dérivées granulométriques.

2.5 Méthodes utilisées

Dans cette section, nous présentons les recherches et choix d'AMR faits dans le cadre de la comparaison des images binaires. Le cadre d'utilisation de l'AMR est le suivant : à partir d'une image binaire, une AMR est effectuée pour avoir à disposition l'image à plusieurs résolutions sous forme d'approximations et de détails. Le choix de la résolution est alors fait pour permettre une classification efficace et optimiser le temps de calcul. Ce dernier point conduit à choisir le cadre de l'analyse décimée (dyadique) où la taille des images d'approximation et de détail est divisée par deux en passant d'une résolution à une résolution plus grossière suivant.

Comme le but est de comparer des images binaires, l'AMR doit, d'une part, conserver autant que possible les structures des images et, d'autre part, donner comme résultat une approximation et des détails binaires. Cette dernière contrainte exclut les AMR linéaires qui ne produisent généralement pas de résultats binaires à partir d'une image binaire. Même l'ondelette (linéaire) de Haar qui est très peu régulière puisque discontinue utilise une moyenne dans l'approximation qui conduit, à partir des images binaires, à des valeurs rationnelles du type $\frac{k}{2^n}$. Les AMR non-linéaires particulièrement adaptées aux images binaires sont celles reposant sur la morphologie mathématique. En effet celle-ci a été développée initialement dans le cadre des images binaires et les filtres morphologiques produisent des images binaires. Une présentation de cette théorie a été faite section 2 (p. 23). La section 2.3 (p. 28) présente les bases de l'AMR morphologique que nous mettons ici en pratique. Un exemple simple d'AMR morphologique est celui de Haar déjà présenté en 2.3. Pour estimer l'effet de l'AMR morphologique de Haar sur les structures des images, elle est appliquée sur une image test et l'effet apprécié visuellement. Ainsi, la figure 2.8 présente les approximations de l'AMR morphologique de Haar sur 3 échelles. Nous constatons que les approximations sont



FIG. 2.8 – Image initiale et ses approximations aux échelles 1 et 2 par le filtre de Haar morphologique. Les approximations sont de plus en plus sombres. Ce qui est dommageable pour la conservation des structures de l'image.

de plus en plus sombres, ce qui nuit à la conservation des structures. Pour éviter ce défaut, il faut faire appel à un opérateur respectant mieux la symétrie noir/blanc. L'ouverture est dans ce cas. La figure 2.9 présente les approximations de l'AMR morphologique pour l'ouverture. L'ouverture a tendance à éclaircir l'image. Ainsi, l'ouverture est un filtre morphologique,

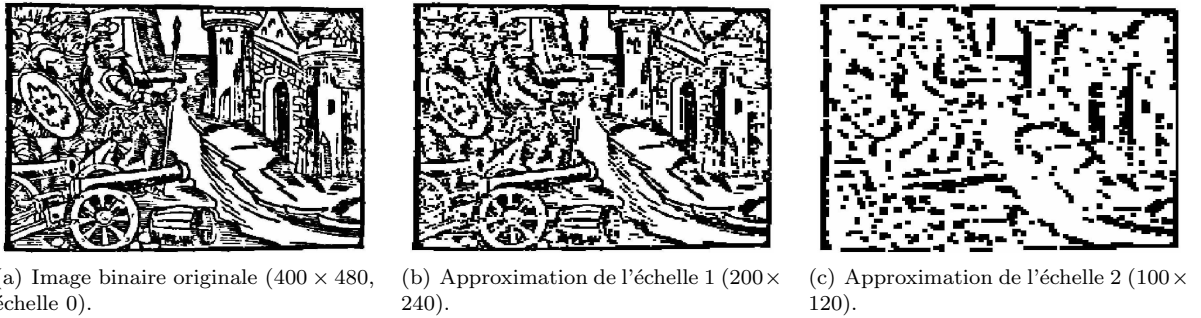


FIG. 2.9 – Image initiale et ses approximations aux échelles 1 et 2 par l'ouverture morphologique. Les approximations sont de plus en plus claires. Ce qui est dommageable pour la conservation des structures de l'image.

et donc idempotent, *i.e.* si elle appliquée deux fois de suite, la deuxième application ne change rien. Dans l'AMR, elle est appliquée plusieurs fois de suite, cependant, entre chaque application, une décimation est opérée. En conséquence l'image sur laquelle est appliqué à nouveau le filtre n'est plus exactement la même. Soit Ψ un filtre morphologique, pour conserver la symétrie dans les échelles, il faut la propriété suivante qui ressemble à l'auto-dualité mais où intervient la décimation :

$$(\downarrow 2) \circ \Psi(X^C) = [(\downarrow 2) \circ \Psi(X)]^C, \quad (2.39)$$

où C désigne la complémentation. Cette propriété revient à l'auto-dualité. En effet, la décimation est une opération auto-duale : décimer revient à conserver un pixel sur 4, et prendre le complémentaire ne dépend que de la valeur du pixel concerné (voir figure 2.10). Ainsi, prendre le complémentaire avant ou après la décimation ne change en rien le résultat. La propriété requise pour le filtre de l'AMR morphologique est donc l'auto-dualité.

Le filtre médian est un filtre auto-dual. Nous présentons ici l'AMR morphologique basée sur ce filtre : soit V_j le treillis des fonctions de \mathbb{Z}^2 dans $\{0, 1\}$ et W_j l'ensembles des fonctions de \mathbb{Z}^2 dans $\{0, 1\}^3$. Avec les notations du paragraphe 2.3, les

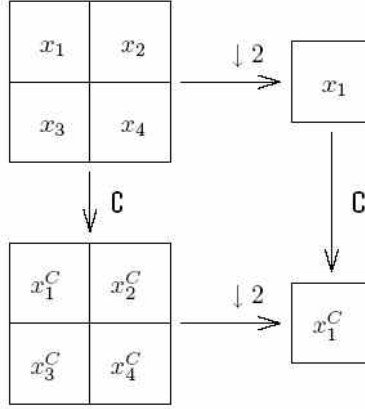


FIG. 2.10 – Diagramme commutatif concernant la décimation et la complémentation illustrant l'égalité (2.39).

opérateurs d'analyse sont définis comme suit :

$$h_a(x)(\mathbf{n}) = \text{médian}(x(\mathbf{2n}), x(\mathbf{2n}_+), x(\mathbf{2n}_+), x(\mathbf{2n}_+^+), x(\mathbf{2n}_+^+)), \quad (2.40)$$

$$g_a(x)(\mathbf{n}) = (d_v(x)(\mathbf{n}), d_h(x)(\mathbf{n}), d_d(x)(\mathbf{n})), \quad (2.41)$$

où d_v, d_h, d_d désignent les opérateurs de détail vertical, horizontal et diagonal, qui sont définis par :

$$d_v(x)(\mathbf{n}) = x(\mathbf{2n}) \Delta x(\mathbf{2n}_+), \quad (2.42)$$

$$d_h(x)(\mathbf{n}) = x(\mathbf{2n}) \Delta x(\mathbf{2n}_+), \quad (2.43)$$

$$d_d(x)(\mathbf{n}) = x(\mathbf{2n}) \Delta x(\mathbf{2n}_+^+). \quad (2.44)$$

En se référant à la figure 2.2, les coefficients dans la matrice $\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}$ sont transformés en $\begin{smallmatrix} t & u \\ v & w \end{smallmatrix}$ avec l'approximation $t = \text{médian}(a, a, b, c, d)$, et les détails $u = a \Delta b, v = a \Delta c, w = a \Delta d$. Le pixel a est représenté deux fois dans le calcul de la médiane : en effet, la médiane doit comporter un nombre impair de valeurs, comme il n'y a que 4 valeurs dans la fenêtre 2×2 , il faut que l'une des valeurs soit représentée deux fois, or a est le point d'application de la fenêtre 2×2 , donc il est représenté deux fois. Il n'est pas difficile de montrer que $a = t \Delta (u \wedge v \wedge w)$. Pour comprendre cela, distinguons deux cas :

1. $u \wedge v \wedge w = 0$: alors, au moins l'un des éléments u, v, w est égal à 0, ce qui implique qu'au moins une des valeurs b, c, d est égale à a . Et donc $t = a$, ce qui correspond à $t = a \Delta 0$.
2. $u \wedge v \wedge w = 1$: alors $u = v = w = 1$ et donc $b = c = d = \bar{a}$. Ce qui donne $t = \bar{a}$ et cela correspond à $a = t \Delta 1$.

Les opérateurs de synthèse sont alors donnés par

$$h_s(x)(\mathbf{2n}) = h_s(x)(\mathbf{2n}_+) = h_s(x)(\mathbf{2n}_+^+) = h_s(x)(\mathbf{2n}_+^+) = x(\mathbf{n}), \quad (2.45)$$

et

$$g_s(y)(2\mathbf{n}) = y_v(\mathbf{n}) \wedge y_d(\mathbf{n}) \wedge y_h(\mathbf{n}), \quad (2.46)$$

$$g_s(y)(2\mathbf{n}_+) = y_h(\mathbf{n}) \triangle (y_v(\mathbf{n}) \wedge y_d(\mathbf{n}) \wedge y_h(\mathbf{n})), \quad (2.47)$$

$$g_s(y)(2\mathbf{n}^+) = y_v(\mathbf{n}) \triangle (y_v(\mathbf{n}) \wedge y_d(\mathbf{n}) \wedge y_h(\mathbf{n})), \quad (2.48)$$

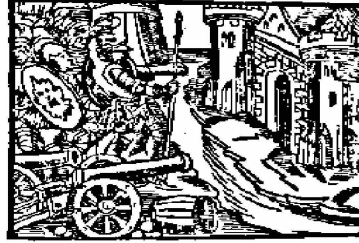
$$g_s(y)(2\mathbf{n}_+^+) = y_d(\mathbf{n}) \triangle (y_v(\mathbf{n}) \wedge y_d(\mathbf{n}) \wedge y_h(\mathbf{n})), \quad (2.49)$$

$$(2.50)$$

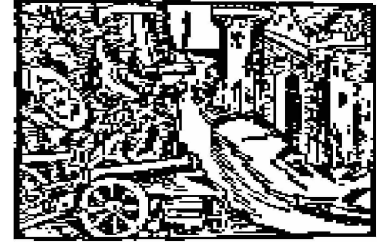
où $y = (y_v, y_h, y_d)$. Les conditions de reconstruction parfaite (2.21) et de non-redondance (2.22) sont vérifiées avec comme + le OU exclusif \triangle . La figure 2.11 illustre l'AMR morphologique de la médiane. L'auto-dualité permet d'obtenir des images



(a) Image binaire originale (400×480 , échelle 0).



(b) Approximation de l'échelle 1 (200×240).



(c) Approximation de l'échelle 2 (100×120).



(d) Approximation de l'échelle 3 (50×60).



(e) Approximation de l'échelle 4 (25×30).

FIG. 2.11 – Image initiale et ses approximations aux échelles 1, 2, 3 et 4 par le filtre de la médiane morphologique.

d'approximation dont les structures sont mieux conservées qu'avec les opérateurs précédents. Les propriétés de cette AMRM de la médiane en ont fait la principale AMR dans les applications de cette thèse. On peut cependant lui reprocher sur l'illustration de ne pas conserver entièrement le cadre (à la résolution 4) alors qu'il a une épaisseur égale sur tout le bord de l'image initiale.

2.5.1 Utilisation des détails

L'AMR morphologique produit à chaque échelle une image d'approximation et trois de détails comme l'illustre la figure 2.12. Pour obtenir une image utilisable à partir des images de détails, il est possible de les recombinaer, en utilisant l'opérateur de reconstruction avec une image à coefficients nuls à la place de l'approximation. La figure 2.13 illustre cette possibilité. Cependant, il est envisageable de travailler sur les trois images de détails séparément afin d'exploiter les informations concernant les orientations verticale horizontale et diagonale.

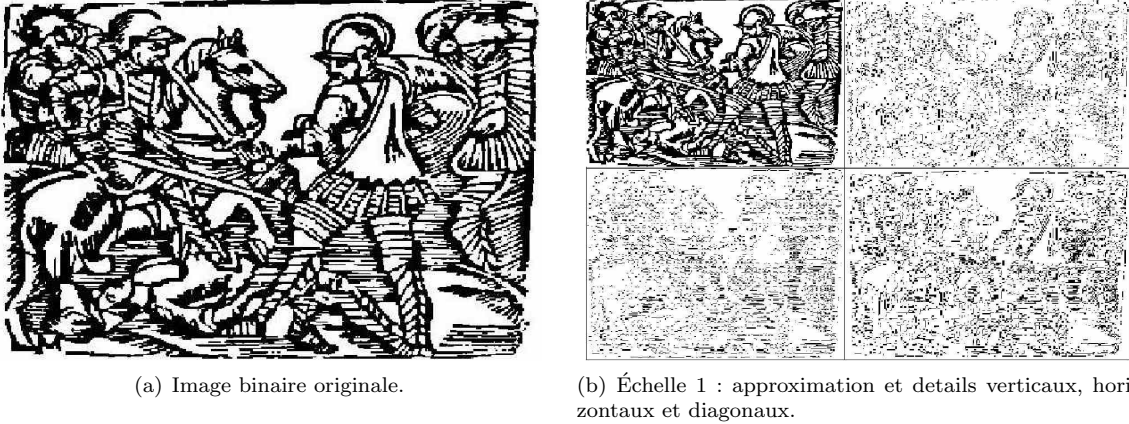


FIG. 2.12 – AMRM de la médiane, les images d’approximation et de détails ont été calculées avec les expressions (2.51) et (2.53)-(2.55).

2.5.2 Étude des filtres

Afin de caractériser les propriétés des filtres non-linéaires concernant la conservations des structures de l’image nous avons suivis plusieurs pistes. L’AMR dyadique décimée n’est alors pas le cadre approprié pour cette étude. En effet, la décimation a le défaut de ne pas donner une AMR invariante par translation et la diminution de la taille des images qui peut être une motivation pour améliorer le temps de calcul ne présente pas d’intérêt ici. Nous avons donc choisi l’algorithme à trous (cf section 1.1) qui conserve l’invariance par translation pour représenter les filtres. L’illustration est faite sur l’image originale et deux échelles. Les courbes concernent les approximations et les détails reconstitués.

La figure 2.14 représente les détails reconstitués de l’algorithme à trous sur une image binaire pour l’AMRM de la médiane. À partir de l’échelle 5, l’image se morcelle et il y a un effet de bord marqué.

Pistes pour d’autres filtres

L’opérateur de la médiane dans le cadre de l’AMR morphologique est utilisé sur un carré de 4 pixels. Le pixel d’application est celui du coin supérieur gauche. Les détails sont alors calculés comme différence entre la médiane faite sur les 4 pixels et la valeur des 3 autres pixels. La position du pixel d’application dans cette fenêtre et la manière dont est calculée la médiane (en répétant 2 fois la valeur du pixel d’application) montre bien que cette fenêtre n’est pas isotropique. Une idée pour obtenir l’isotropie est de calculer la médiane sur une fenêtre 3×3 . Conserver trois images de détail et un pas de 2 donne le schéma d’analyse suivant (en utilisant les notations matricielles pour la fenêtre 3×3 d’analyse) :

$$h_a(x)(\mathbf{n}) = \text{médian}[x(1,1), x(1,2), x(1,3), x(2,1), x(2,2), x(2,3), \\ x(3,1), x(3,2), x(3,3)], \quad (2.51)$$

$$g_a(x)(\mathbf{n}) = (d_v(x)(\mathbf{n}), d_h(x)(\mathbf{n}), d_d(x)(\mathbf{n})), \quad (2.52)$$

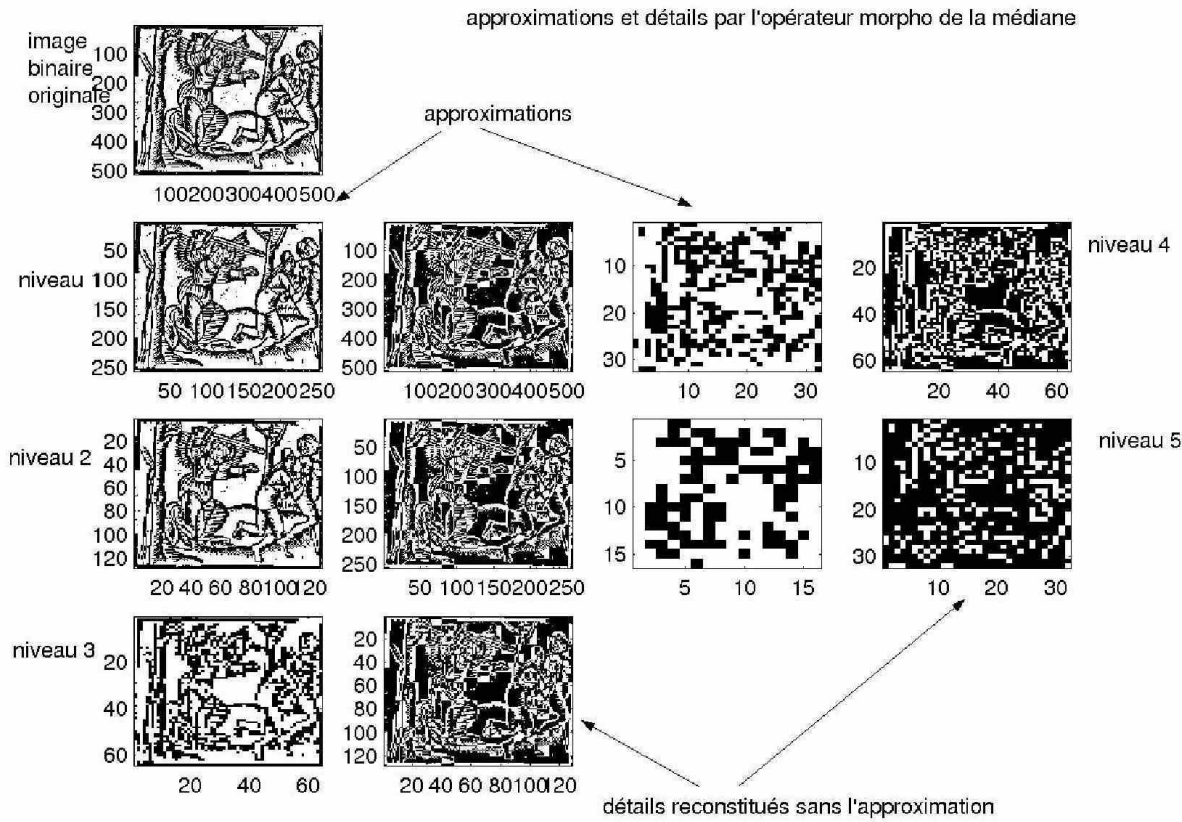


FIG. 2.13 – Image binaire pour laquelle a été faite une AMRM dyadique avec le filtre de la médiane. Sont représentés l'image originale (échelle 0), les approximations et les détails reconstitués aux échelles 1 à 8. Contrairement à l'habitude, les détails sont représentés avec 0=noir et 1=blanc.

où les opérateurs de détail d_v, d_h, d_d sont définis par :

$$d_v(x)(\mathbf{n}) = x(2\mathbf{n}) \triangle x(2,3), \quad (2.53)$$

$$d_h(x)(\mathbf{n}) = x(2\mathbf{n}) \triangle x(3,2), \quad (2.54)$$

$$d_d(x)(\mathbf{n}) = x(2\mathbf{n}) \triangle x(3,3). \quad (2.55)$$

Les opérateurs de détails restent les mêmes que dans le cas de la fenêtre 2×2 . Pour la synthèse, comme le pas est de 2, il n'y a que deux points à reconstruire récursivement. Malheureusement, la condition de reconstruction parfaite n'est pas satisfaite car les fenêtres $\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}$ et $\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{smallmatrix}$ ont la même décomposition selon cet algorithme. L'étude de variations autour de cette analyse (en faisant apparaître plusieurs fois le point d'application dans le calcul de la médiane, par exemple) n'a donc pas permis d'avoir la reconstruction parfaite. Cette étude est difficile, comme souvent dans le cas non-linéaire. Et malgré nos recherches, nous n'avons pas trouver de filtres ayant un meilleur comportement que celui de la médiane morphologique.



FIG. 2.14 – AMRM de la médiane par l'algorithme à trous : détails reconstitués. À partir de l'échelle 5 les images se dégradent car les points utilisés dans l'algorithme pour le calcul sont trop éloignés les uns des autres.

2.6 Bilan

Basée sur la théorie des ensembles, la morphologie mathématique est bien adaptée au traitement des images binaires qui peuvent être considérées comme des ensembles de points. Plusieurs types de filtres morphologiques sont présentés ainsi que certaines propriétés caractérisant ces filtres. Ces filtres morphologiques sont alors présentés dans le cadre de l'analyse multirésolution non-linéaire. La propriété d'auto-dualité est alors soulignée comme permettant la conservation de la proportion noir/blanc pour les images d'approximations aux différentes échelles. Une section est ensuite dédiée à l'étude des filtres et aux moyens de mener cette étude. La piste de la granulométrie est explorée. Le filtre de la médiane se révèle être celui qui donne la meilleure conservation des structures de l'image binaire aux différentes échelles de l'analyse multirésolution.

Deuxième partie

Mesures de Dissimilarités

Introduction

La comparaison d'images intervient dans plusieurs domaines tels que l'évaluation de la qualité d'une image [Wang *et al.*, 2004], la reconnaissance de formes [Veltkamp and Hagedoorn, 1999] ou encore l'indexation. Dans ce dernier cas, la comparaison des images afin de déterminer leur similarité en un sens prédéfini peut prendre plusieurs formes : une méthode générale consiste à extraire des descripteurs des images et à comparer ces descripteurs entre eux par une mesure de similarité. Une autre méthode consiste à effectuer une comparaison directe des images, par exemple lorsque les descripteurs sont trop pauvres. Le cas des images binaires se ramène à cette situation. Dans le cas où les images binaires ne sont pas de simples formes mais des images plus complexes composées éventuellement de plusieurs parties, une possibilité est de développer une méthode de comparaison locale qui puisse rendre compte des dissimilarités locales. C'est cette étude que nous avons menée et qui est présentée dans cette partie. Elle a conduit au développement d'une méthode de comparaison directe d'images binaires qui donne une mesure locale des dissimilarités entre les images comparées.

Dans un premier temps, les différentes méthodes de comparaison d'images sont passées en revue section 1.1 en commençant par les méthodes générales qui sont utilisées en indexation. Puis les méthodes spécifiques aux images binaires sont présentées dans la section 1.2. Enfin les raisons pour développer une nouvelle méthode de comparaison locale déjà esquissées ci-dessus, sont détaillées dans la section 1.2.5.

1

La comparaison d'images : panorama

L'objectif de ce chapitre est de dresser un état de l'art des principales méthodes de comparaison d'images. Étant donné la diversité et le nombre important de travaux disponibles dans le domaine, nous n'avons pas jugé utile de réaliser une étude bibliographique exhaustive. Nous pensons que ce chapitre doit plutôt rassembler les références pertinentes permettant d'appuyer le propos de cette thèse.

Dans ce chapitre, nous décrivons les méthodes de comparaison d'images en couleurs ou en niveaux de gris puis nous montrons en quoi elles ne sont pas transposables aux images binaires. Enfin nous présentons celles dédiées aux images binaires.

1.1 Les méthodes de comparaison

Les méthodes de comparaison d'images reposent sur la sélection de descripteurs discriminants, dans le cadre d'une base d'images donnée. Les descripteurs représentent en général de l'information sur la couleur, la texture, et les formes extraites de l'image [Smeulders *et al.*, 2000]. Leur choix, qui conditionne l'efficacité de la méthode, constitue une étape délicate de l'indexation [Antani *et al.*, 2002]. Ces descripteurs sont aussi utilisés pour les images en niveaux de gris, à l'exception, bien sûr, de celui de la couleur. Une fois les descripteurs sélectionnés, ils sont agencés pour chaque image pour former une signature. Ces signatures constituent l'index de la base et sont utilisées lors de la recherche d'images de la base correspondant à une requête. Les signatures des images de la base sont alors comparées à celle de la requête à l'aide d'une mesure de dissimilarité. Cette dernière est construite en fonction des descripteurs choisis qui peuvent être traités séparément ou agrégés sous forme d'un vecteur. Détaillons maintenant les descripteurs, leur mise en forme et les mesures de dissimilarité.

1.1.1 Les descripteurs des images

La couleur

La couleur a un fort pouvoir discriminant et elle est largement utilisée comme descripteur pour la recherche d'images par le contenu ou pour la discrimination des textures. Les recherches se sont portées sur le choix de l'espace colorimétrique [Stricker and Orengo, 1995], l'étude de l'invariance aux conditions d'acquisition telles que l'illumination et la prise de vue. Historiquement, le premier espace colorimétrique utilisé pour la représentation sur un moniteur est l'espace RVB (acronyme de Rouge Vert Bleu, RGB en Anglais) ; il reste utilisé dans le cadre d'études de recherche d'images par le contenu [Pass *et al.*, 1996; Huang *et al.*, 1997]. Cependant il est sensible aux changements d'illumination et d'autres espaces ont été développés. Ainsi

l'espace TSI (HSV en Anglais) sépare l'information relative à la teinte, la saturation et à l'intensité. La teinte représente la longueur d'onde de la lumière réfléchie, ou transmise par un objet. La saturation indique la pureté ou intensité de la couleur. Enfin, l'intensité désigne la variation d'intensité lumineuse d'une couleur. L'espace TSI est communément utilisé [Smith and Chang, 1996; Stricker and Orengo, 1995]. En se référant au système visuel humain, l'espace CIELab [C.I.E., 1978] a été conçu pour être perceptuellement uniforme. Dans cet espace, les couleurs sont définies par trois valeurs : Luminosité (luminance) codée en pourcentage, a et b correspondent à l'information colorée (chrominance) où la couleur est définie à partir d'un mélange de vert à magenta (a) et un mélange de bleu à jaune (b). Il sert également pour l'indexation [Ciocca and Schettini, 1999]. Notons qu'il est possible de passer d'un espace de représentation à un autre par des fonctions de transfert.

Pour les images binaires, il n'y a qu'un seul paramètre par pixel et il ne peut prendre que deux valeurs (0 ou 1). Dans le cadre de notre thèse, les méthodes ci-dessus n'ont pas lieu d'être utilisées.

Les caractéristiques de formes

Les descripteurs de formes sont complémentaires de la description de la couleur. Des études ont été faites pour les rendre robustes aux transformations géométriques comme la translation, la rotation et le changement d'échelle. Nous distinguons deux catégories de descripteurs de formes : les descripteurs basés sur les régions et ceux basés sur les frontières.

Les premiers font classiquement référence aux moments invariants [Hu62] [Derrode et al.99] et sont utilisés pour caractériser l'intégralité de la forme d'une région. Cependant, il est difficile de relier les valeurs des moments de grand ordre aux caractéristiques des formes et les traits (ou l'information locale) ne peuvent pas être détectés grâce aux moments.

La seconde approche porte sur une caractérisation des contours de la forme (coefficients de Fourier, excentricité, nombre d'Euler, ...) [Veltkamp and Hagedoorn, 1999]. Ces méthodes sont tributaires d'une détection de contour. Comme le principe de la détection de contour est de sélectionner les zones de forte variations. Son choix pour des images peu contrastées ou bruitées peut s'avérer difficile.

Pour les images binaires, les variations de niveaux de gris sont toujours d'amplitude égale à 0 ou à 1. L'information concernant les contours est donc pauvre. La détection se fait alors grâce à la topologie discrète, principalement l'utilisation des 4-voisinages ou 8-voisinages [Leonard, 1991]. Il est possible de distinguer de cette manière l'intérieur et la frontière des objets et de les étiqueter. Cependant, la détection d'objet dans les images binaires n'est pas fiable lorsque les images ne sont pas composées de formes simples, par exemple lorsqu'elles comportent beaucoup de traits. Comme notre objectif est de traiter ce genre d'images, nous n'utiliserons pas ces méthodes. Par contre, le résultat d'une segmentation peut être représenté par une image binaire et son traitement s'inscrit dans le cadre de notre étude.

Les caractéristiques de texture

La texture a plusieurs définitions. La notion de texture est utilisée pour traduire un aspect homogène de la surface des niveaux de gris d'un objet sur une image. Dans [Unser, 1995], l'auteur présente la texture comme une structure disposant de certaines propriétés spatiales homogènes et invariantes par translation. Dans [Gagalowicz, 1983], l'auteur considère la texture comme « une structure spatiale constituée de l'organisation de primitives ayant chacune un aspect aléatoire, donc une structure hiérarchique à deux niveaux ». La prise en compte de ce type de caractéristiques pour représenter globalement ou partiellement une image est courante et discriminante dans de nombreux cas. À titre d'exemple, ce type d'information est utilisé pour l'indexation d'images satellitaires [Li et al.97]. Les descripteurs de texture en indexation font leur apparition dans le système de recherche d'images par le contenu d'IBM QBIC [Niblack et al.93]. Les caractéristiques en question sont le grain, le contraste et l'orientation. Les travaux de Aksoy et Haralick [Aksoy et al.98] sur la texture et en particulier les matrices de co-occurrences et les différents indices qui peuvent en découler ont également servi de base pour l'indexation. De

nombreuses méthodes sont référencées dans la littérature pour la décomposition de l'image et le calcul de caractéristiques dites de texture. Parmi les plus connues, on trouve la décomposition paramétrique Wold d'abord utilisée par Liu et Picard [Liu et al.96] puis reprise notamment dans le système Photobook [Pentland et al.95] ou encore par Stoica et al. [Stoica et al.98].

Contrairement aux images couleurs ou en niveaux de gris, qui peuvent généralement être considérées comme version discrète de fonctions ayant une certaine continuité, les images binaires ne comportent que des variations brusques entre 0 et 1. Donc les méthodes développées ci-dessus qui utilisent une certaine régularité de l'image (vue comme une fonction de \mathbb{R}^2 dans \mathbb{R}) telles que celles basées sur la transformation de Fourier, ou la décomposition par ondelettes, ne peuvent pas être utilisées efficacement sur les images binaires. D'un autre côté, des caractérisations basées sur la fonction d'autocorrélation sont utilisées dans ce cadre [Maddess et al., 2004]. Cependant, les textures ne sont utiles que pour des images binaires texturées, ce qui est un cas bien particulier.

1.1.2 Les mesures de similarité

Une fois les descripteurs discriminants extraits, ils sont ordonnés pour former la signature de l'image. Les signatures servent alors à la comparaison des images [Philipp-Foliguet et al., 2002], et cette comparaison doit rendre compte du degré de similarité entre les images qu'elles représentent. Deux types de structuration de l'information pour former une signature sont possibles : la première synthétise l'information de toute l'image et forme donc une signature globale ; la seconde tient compte des différences au sein même de l'image et rend compte de l'information locale et partielle. L'exemple le plus simple et le plus courant de signature est l'histogramme [Brunelli and Mich, 2001], qui représente une approximation de la densité de probabilité de l'image vue dont l'intensité est vue comme une variable aléatoire.

1.1.3 Bilan

Les méthodes de comparaison d'images en couleur ou en nuance de gris reposent sur l'extraction de descripteurs portant de l'information sur la couleur, la texture et les formes. Ils sont ensuite structurés dans une signature qui peut être globale ou locale. Pour les images binaires, cette méthode est difficilement transposable dans la mesure où la couleur et la segmentation n'ont pas d'équivalent binaire approprié. De même leur structuration sous forme de signature n'est pas adaptée aux images binaires. L'utilisation de la texture est une possibilité, qui ne s'applique pas à tous les types d'images binaires.

Après avoir présenté les méthodes générales de comparaison d'images, nous allons maintenant donner un panorama des méthodes développées pour la comparaison d'images binaires.

1.2 Les méthodes spécifiques aux images binaires

1.2.1 Introduction

Les images binaires sont difficiles à traiter car chaque pixel comporte peu de descripteurs. Cependant, elles font l'objet de nombreuses études. Cela s'explique par leur présence dans beaucoup de méthodes : à partir d'une image quelconque, un prétraitement est effectué, on obtient alors une image binaire qu'il faut traiter ou comparer. L'étude de leur comparaison est donc justifiée. Néanmoins, les images à comparer peuvent être complexes et ne pas entrer dans le cadre de l'étude des formes simples. Pour ces images complexes, certaines zones peuvent être similaires et d'autres complètement différentes. Nous serons donc attentifs à la manière dont les mesures peuvent en tenir compte.

Bien que les images binaires soient dépourvues de la plupart des descripteurs précités, comme la couleur (!), il est possible de tirer parti de la spécificité des images binaires pour les comparer. La première spécificité est de pouvoir distinguer facilement le fond (dont les pixels sont codés par 0) de l'information (codée par 1) et en conséquence de disposer facilement de cette information. Cette information peut être représentée mathématiquement comme un ensemble de points dans le plan. On distingue alors deux classes de méthodes de comparaison d'ensembles de points que nous étudierons par la suite :

- celles qui comparent directement les images binaires.
- Celles effectuant un traitement préalable soit en les transformant (par exemple avec la transformation de Fourier) soit en extrayant de l'information sous forme de descripteurs pour les utiliser pour la comparaison.

Avant de rentrer dans le détail des méthodes nous introduisons ici des notations qui nous serviront dans toute la section.

Notations Soit X l'ensemble des pixels composant l'image, que l'on suppose fini. Une image binaire est une fonction b de X (le support de l'image) dans $\{0, 1\}$. La valeur 1 sera interprétée comme logiquement vraie et vue à l'écran comme noire. L'image binaire b peut-être identifiée à l'ensemble $B = \{x/b(x) = 1\}$, ce que l'on fera par la suite. Pour deux images binaires A et B , nous utiliserons le moins ensembliste $B \setminus A = \{x \in B/x \notin A\}$ et l'opérateur du OU exclusif ensembliste $A \Delta B = (A \setminus B) \cup (B \setminus A)$. Enfin, nous noterons $n(B)$ le nombre de pixels noirs de B et $n(X)$ le nombre total de pixel dans l'image.

1.2.2 Méthodes de comparaison directes

Nous désignons par méthodes directes les méthodes qui effectuent la comparaison entre les images directement à partir des pixels. Les comparaisons directes entre images binaires sont difficiles à mettre en œuvre. Cela vient du fait que contrairement aux images en couleurs, les pixels dans les images binaires comportent peu d'information : l'amplitude de la variation vaut 0 ou 1. Les comparaisons directes peuvent servir à évaluer du bruit (rapport signal sur bruit), de l'erreur de segmentation [Baddeley, 1992] ou à comparer les images de contours de visages pour en évaluer la dissimilarité [Gao and Leung, 2002; Guo et al., 2001; Lin et al., 2003]. La plupart se fondent sur une mesure globale et fournit une mesure de similarité sous forme scalaire (un réel positif), et une se base sur l'étude locale (dans le cadre des images en nuances de gris) [Wang et al., 2004] et fournit un index de similarité en deux dimensions. Nous présentons ces différentes méthodes et montrons leurs limitations.

Comparaisons basées sur la différence pixel à pixel

Ce sont les mesures sur les propriétés statistiques de la comparaison pixel à pixel. Le rapport signal sur bruit binaire est la plus connue : soient A et B deux images binaires,

$$RSB(A, B) = -20 \log_{10} \left(\frac{1}{n(X)^2} \sum_{x \in X} (b(x) - a(x))^2 \right) \quad (1.1)$$

Le PSNR, qui est le rapport signal sur bruit normalisé par le nombre de niveau de gris, revient au RSB dans le cas binaire ($N = 1$) :

$$PSNR(A, B) = 20 \log_{10} \frac{2^N - 1}{\frac{1}{n(X)^2} \sum_{x \in X} (b(x) - a(x))^2} \quad (1.2)$$

Une autre quantité utilisée est le taux de pixels différents :

$$\epsilon(A, B) = \frac{\{x/a(x) \neq b(x)\}}{n(X)}, \quad (1.3)$$

Or dans le cadre d'images binaires, $\{x/a(x) \neq b(x)\}$ est l'ensemble des points qui sont exactement dans l'un des deux ensembles A et B . Ce qui revient à

$$\epsilon(A, B) = \frac{n(A \Delta B)}{n(X)}. \quad (1.4)$$

ϵ a comme avantage théorique d'être symétrique par rapport au rôle des images A et B et de ne pas nécessiter de normalisation. Elle est un cas particulier de la norme L_1 . L'erreur quadratique moyenne est définie par

$$\rho(A, B) = \sqrt{\frac{1}{n(X)} \sum_{x \in X} |b(x) - a(x)|^2}, \quad (1.5)$$

or dans le cas binaire $\forall x \in X, |b(x) - a(x)| = 0, 1$ donc ρ s'écrit

$$\rho(A, B) = \sqrt{\frac{1}{n(X)} \sum_{x \in X} |b(x) - a(x)|}, \quad (1.6)$$

ce qui revient à

$$\rho(A, B) = \text{sqr}t \frac{n(A \Delta B)}{n(X)}. \quad (1.7)$$

Si l'objectif est de mesurer la similarité pixel à pixel, alors, RSB , ρ et ϵ sont des choix raisonnables (par exemple pour évaluer une détection de contours ou pour mesurer du bruit). Cependant pour mesurer la dissimilarité entre les images, alors il est connu que le taux de pixels différents est une mesure pauvre. La dissimilarité entre A et B est mesurée grâce au nombre de pixels différents sans tenir compte des formes. L'erreur due au déplacement d'une frontière touche beaucoup de pixels mais n'affecte pas sévèrement la forme va donner une valeur importante à ϵ . *A contrario*, des dissimilarités qui ne touchent qu'un petit nombre de pixels mais qui affectent réellement (par exemple la disparition ou la création de points, trous, lignes) la forme donnent une faible ϵ (et de même pour le RSB et ρ). Cette mesure est donc globale et peu précise pour les formes. L'inconvénient de ce type de mesure est relié au fait qu'elle est basée seulement sur la différence pixel à pixel, et cela ne prend pas en compte les relations spatiales entre pixels. La sommation qui est faite ensuite contribue à réduire l'information, cependant, si nous considérons l'étape précédent la sommation, c'est-à-dire l'image $|B - A|$ résultant de la soustraction pixel à pixel, l'information spatiale obtenue ne suffit pas à rendre compte des dissimilarités entre les images.

Une illustration est donnée figure 1.1 : les images A et B sont différentes et les images C et D sont proches et ont la même matrice de différence $|B - A|$ entre A et B et $|C - D|$ entre C et D . Cette mesure dans le cadre des images binaires ne permet pas de mesurer les dissimilarités, même en conservant l'information spatiale. Il est donc nécessaire de faire appel à des mesures qui prennent en compte plus d'information que la différence pixel à pixel.

Mesure relative aux ensembles de pixels

Introduction Comme indiqué dans l'introduction 1.2.1, une image binaire est une fonction b de $X \subset \mathbb{R}^2$ dans $\{0, 1\}$ et elle peut être identifiée à l'ensemble B des antécédents dans \mathbb{R}^2 de la valeur 1 : $B = \{x \in \mathbb{R}^2, b(x) = 1\}$. Il existe alors plusieurs mesures mathématiques calculant une distance entre des ensembles de points de \mathbb{R}^2 que l'on peut donc utiliser pour les images binaires. Parmi elles, la différence ensembliste et la distance de Hausdorff sont largement utilisées en traitement de l'image. Ces mesures sont des distances au sens mathématique du terme, *i.e.* elles sont à valeur réelle positive et vérifient les propriétés d'identité, de symétrie et d'inégalité triangulaire. Ces propriétés sont intéressantes dans le cadre du traitement de l'image car elles reflètent des propriétés attendues dans une comparaison d'images :

- Une image doit être identique à elle-même.

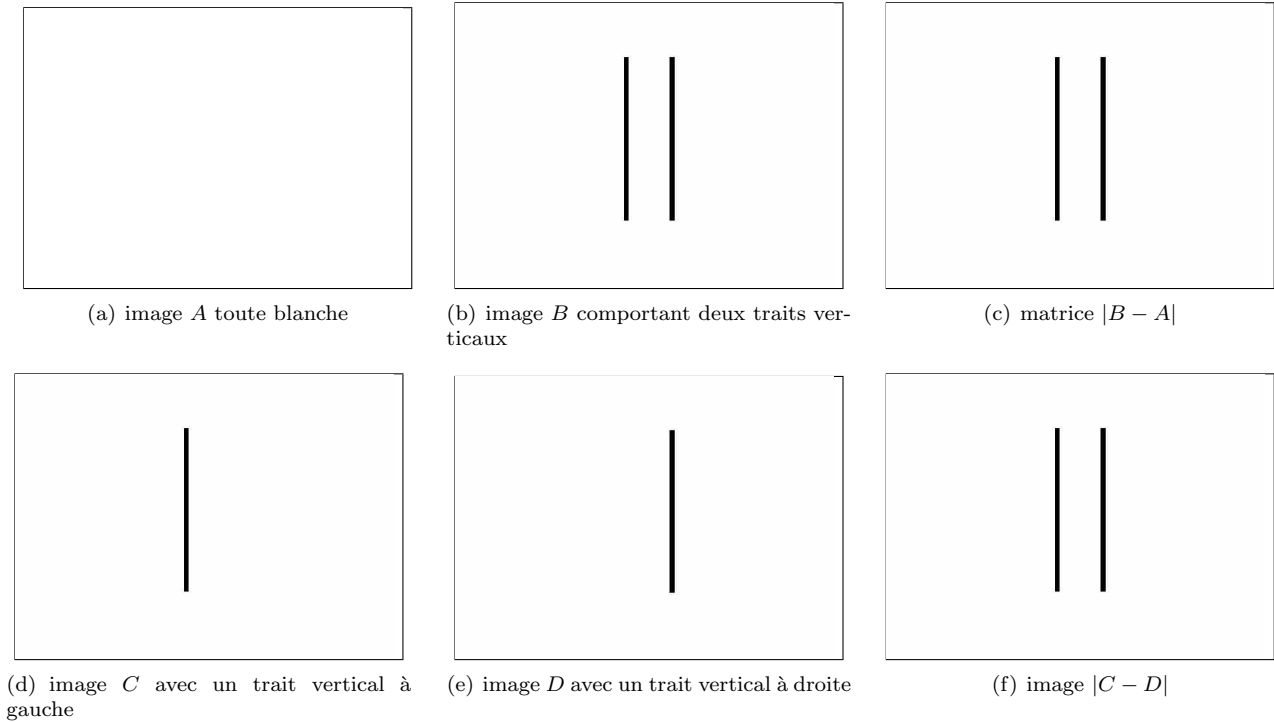


FIG. 1.1 – Les images A et B sont différentes et les images C et D proches. Par contre, les matrices des différences $|B - A|$ et $|C - D|$ sont les mêmes.

- Si une image A ressemble à une image B , le contraire est aussi vrai en général.
- Si une image A ressemble à une image B qui elle-même ressemble à une image C alors l'image A ressemble à l'image C .

Cependant, ces propriétés ne sont pas toujours vérifiées pour le Système Visuel Humain (SVH) [Tversky, 1977], que ce soit la symétrie dont une étude a montré qu'elle n'était pas toujours de mise lors de comparaison d'image par le SVH [Scassellati *et al.*, 1994], ou l'inégalité triangulaire qui ne rend pas compte de la réalité dans le cas de dissimilarités importantes entre les images (voir figure 1.2.2 tiré de [Veltkamp and Hagedoorn, 2000]). Enfin elles peuvent tout simplement ne pas être vérifiées par des fonctions dont les propriétés de mesure sont intéressantes. Ainsi une classe plus large de fonctions ayant des propriétés de mesure intéressantes a été introduite : les mesures de dissimilarité qui contrairement aux distances ne vérifient pas forcément la symétrie et l'inégalité triangulaire.

remarque 1.2.1. : les expressions « mesure de dissimilarité » et « mesure de similarité » sont employées dans la littérature. Il n'y a pas de distinction à notre avis : dans les deux cas, il y a mesure de ce qui est similaire et dissimilaire entre les deux objets (images, vecteurs). La première nous semble plus appropriée puisque la mesure faite est d'autant plus grande que les objets comparés sont différents.

Il est à noter que ces mesures sont parfois dédiées aux formes et peuvent utiliser leurs propriétés, nous préciserons alors ce qui peut être transposable dans un cadre général.

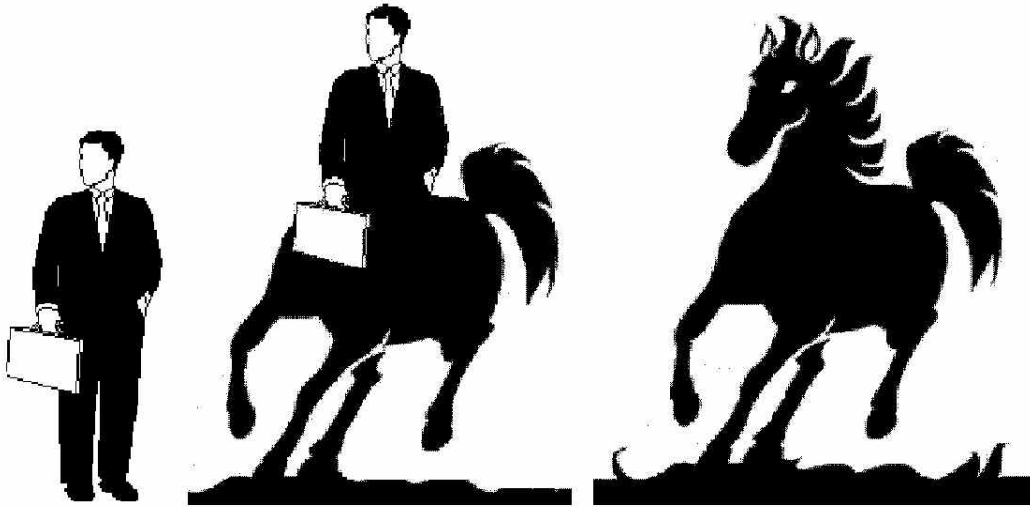


FIG. 1.2 – Illustration pour l’inégalité triangulaire : l’image de gauche est similaire à celle du centre, celle du centre à celle de droite. Si la mesure de similarité respecte l’inégalité triangulaire, celle de gauche est similaire à celle de droite, ce qui n’est pas le cas.

Notations Une distance d entre les pixels d’une même image est introduite, en les considérant comme des points dans le plan $d : (x, y) \in X^2 \mapsto d(x, y) \in \mathbb{R}^+$. Soit $d(x, A)$ la distance minimale du pixel $x \in X$ à $A \subset X$:

$$d(x, A) = \min_{a \in A} (d(x, a)). \quad (1.8)$$

Pour une distance d sur \mathbb{R}^2 classique et pour une grille sur X rectangulaire ou hexagonale, la fonction $x \mapsto d(x, A)$ peut être rapidement calculée grâce à l’algorithme de la transformée en distance. [Borgefors, 1986a; Rosenfeld and Kak, 1982].

Les mesures statistiques point à point Les mesures statistiques point à point sont la distance moyenne

$$\bar{e} = \frac{1}{n(B)} \sum_{x \in B} d(x, A), \quad (1.9)$$

La distance carrée moyenne

$$\bar{e}^2 = \frac{1}{n(B)} \sum_{x \in B} d(x, A)^2, \quad (1.10)$$

et la figure de mérite (Figure Of Merit : FOM) [Pratt, 1977]

$$FOM(A, B) = \frac{1}{\max(n(A), n(B))} \sum_{x \in B} \frac{1}{1 + \alpha d(x, A)^2}, \quad (1.11)$$

avec α une constante d’échelle habituellement fixée à $1/9$ quand d est normalisée pour que la plus petite distance entre deux pixels soit égale à 1. Cette mesure n’est pas symétrique. Elle est cependant très populaire parmi les mesures d’erreur [Baddeley, 1992]. Sa normalisation est faite de telle sorte que $0 \leq FOM(A, B) \leq 1$ avec $A = B \Leftrightarrow FOM(A, B) = 1$. On peut émettre la critique suivante sur cette mesure : la FOM n’est pas sensible aux formes dans la mesure où c’est la moyenne de

la fonction $d(\cdot, A)$ qui est considérée.

La distance de Hausdorff La distance de Hausdorff (DH) se définit comme suit :

Définition 1.2.1 (distance de Hausdorff). Soient deux ensembles non vides de points $A = (a_1, \dots, a_n)$ et $B = (b_1, \dots, b_m)$ de E espace métrique, muni d'une distance d , la DH est donnée par

$$D_H(A, B) = \max(h(A, B), h(B, A)) \quad (1.12)$$

$$\text{où } h(A, B) = \max_{a \in A} \left(\min_{b \in B} d(a, b) \right), \quad (1.13)$$

En utilisant les notations $d(\cdot, A)$ définies ci-dessus, cela devient :

$$D_H(A, B) = \max(\sup_{b \in B} d(b, A), \sup_{a \in A} d(a, B)).$$

La DH est théoriquement intéressante et pratiquement répandue. C'est une distance métrique définie sur l'ensemble des parties de X si X est fini et plus généralement sur l'ensemble des compacts d'un espace métrique localement compact en particulier sur \mathbb{R}^d , pour plus de détails voir [Matheron, 1975]. Cependant, elle n'est pas utilisée en pratique comme mesure de similarité pour les images car elle est très sensible aux points isolés. De nombreuses versions modifiées ont été proposées pour améliorer sa robustesse, avec des applications qui vont de la recherche de forme dans une image [Huttenlocher and Rucklidge, 1993] à la reconnaissance de visages [Zhao et al., 2004; Jesorsky et al., 2001]. Cette distance étant utilisée pour introduire la carte de dissimilarité, une étude détaillée sera présentée section 1.3.

SSIM (Structural SIMilarity index) Récemment, Wang et al ont apporté de l'information spatiale dans leur mesure d'erreur appelée index de similarité structurelle (SSIM) [Wang et al., 2004]. Ils ont pour cela exploité les caractéristiques connues du système visuel humain (SVH) selon lesquelles le SVH est hautement adapté pour extraire l'information structurelle dans une image. Leur index SSIM peut être vu comme une mesure locale de similarité car il est basé sur la luminance, le contraste et la mesure au travers d'une fenêtre 8×8 de la structure. Cependant la taille de la fenêtre de mesure est indépendante du contenu des images comparées et ne peut donc pas être adaptée au type de dissimilarité entre les images à évaluer.

1.2.3 Les méthodes indirectes

Dans cette sous-section sont présentées des méthodes de comparaison utilisées dans le cadre des images binaires qui font appel à une étape intermédiaire. Les premières méthodes présentées utilisent l'extraction de forme pour travailler sur les objets. Les suivantes se fondent sur une transformation de l'image pour représenter l'image dans un nouvel espace (*e.g.* de Fourier).

Méthode sur l'objet global

Méthodes basées sur l'objet : les moments géométriques Les moments géométriques en deux dimensions qui fournissent des descripteurs pour la reconnaissance ont de nombreuses applications en traitement de l'image. En robotique, ils sont utilisés pour la recherche de mouvement le calcul d'orientation et pour le recalage, ils sont aussi utilisés en traitement de

l'image pour des problèmes tels que l'appariement de scène [Wong and Hall, 1978] ou la reconnaissance de caractères [Chim et al., 1999]. Un ensemble de points ou un objet peuvent être décrits par l'ensemble de leurs moments géométriques $m_{p,q}$ d'ordre $(p, q) \in \mathbb{N}^2$

Pour une image b binaire $m \times n$

$$m_{p,q} = \sum_{x=1}^m \sum_{y=1}^n x^p y^q b(x, y), \quad (1.14)$$

où comme précédemment, les pixels du fond ont pour valeur 0. Basées sur ces moments, de nombreuses fonctions appelées invariants de moments peuvent être définies pour être invariantes à certains groupes de transformations (en générale affines). Seuls les moments de bas ordre sont conservés pour servir à la comparaison. Un inconvénient de cette méthode est qu'elle ne permet pas de description locale de l'image comme nous le souhaitons. Ce défaut est partagé par les méthodes s'appuyant sur une description globale de l'image tels que l'aire, l'excentricité, la compacité, le nombre d'Euler, le nombre de formes. De plus, elles sont mal adaptées aux images binaires complexes pour lesquelles un grand nombre d'invariants sont nécessaires, ce qui rend la méthode moins robuste au bruit.

Méthodes basées sur la frontière Ces méthodes telles que les descripteurs de Fourier, l'espace de décomposition du contour courbure-échelle ou la décomposition du contour en composantes principales ont de nombreuses applications en particulier pour le traitement des formes [Veltkamp and Hagedoorn, 1999]. Nous ne nous attarderons pas sur les méthodes basées sur la frontière qui est un descripteur robuste pour une forme, mais qui l'est peu pour une image complexe, qui contient plusieurs éléments, voir figure 1.3.

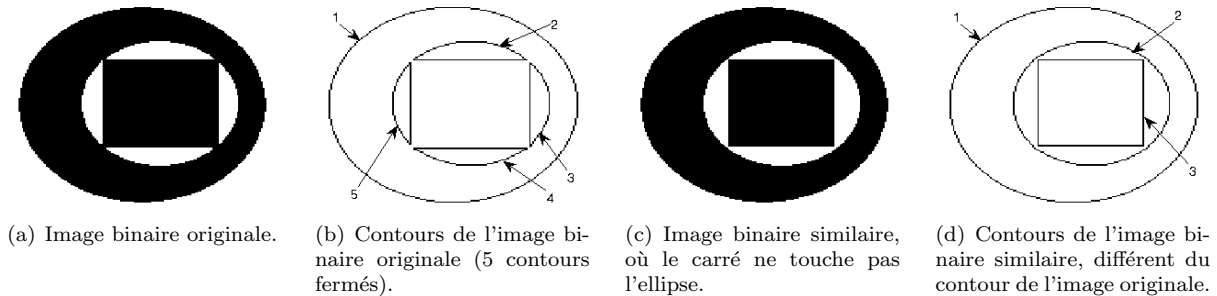


FIG. 1.3 – Illustration de la complexité du contour, le premier comporte 5 contours fermés alors que le deuxième n'en comporte que 3, qui ne ressemblent en rien aux 5 précédents.

Transformée sur l'image globale

Comme nous l'avons déjà expliqué dans le paragraphe 1.1.1 p. 48, les transformées globales connues (transformées de Fourier et transformée en ondelettes) sont mal adaptées au traitement des images binaires car elles sont composées de discontinuités. Les méthodes non-linéaires de filtrage sont adaptées aux images binaires [Heijmans and Goutsias, 2000] et elles font l'objet de la première partie. Cependant, dans le cas du filtrage linéaire, la comparaison entre les transformées des images ne peut se faire qu'après la sélection des coefficients les plus significatifs (par la comparaison de leurs normes), ce qui est beaucoup moins discriminant dans le cas de coefficients binaires.

1.2.4 Bilan

Nous avons présenté dans cette section des méthodes de comparaison d'images directes et indirectes. Les méthodes indirectes sont mal adaptées à la comparaison des images binaires autres que les formes. Les méthodes directes présentent un intérêt à partir du moment où elles mesurent une distance entre des ensembles de points et non plus seulement de pixels à pixels. Cependant elles ne permettent en général pas d'avoir accès à l'information de dissimilarité locale ou alors sans s'adapter au type de dissimilarité à évaluer. Or, en préambule, nous avons indiqué notre intérêt pour que la mesure tienne compte des dissimilarités locales.

1.2.5 Vers une nouvelle méthode de comparaison

Ceci nous a conduit à élaborer une autre méthode qui consiste à comparer directement les images binaires sans en avoir extrait une signature représentative et à travailler ensuite sur le résultat de la comparaison. Cette méthode n'oblige pas à opérer le choix des descripteurs discriminants pour les images. En effet, dans la méthode classique, le choix des descripteurs qui constitue l'étape d'indexation, ainsi que celui de leur structuration ont pour objectif l'objectivité, l'exhaustivité, et l'intégrité de l'index produit. Ceci est conditionné par l'application et la base visées et demande l'intervention d'un expert pour le choix de paramètres (de pondération) de représentation (espace colorimétrique) etc... Dans le cas de notre méthode, la comparaison est directe et ne fait appel à aucun paramètre.

La méthode que nous allons présenter permet une évaluation locale adaptative des dissimilarités tout en conservant leur agencement spatial. Nous allons l'exposer en prenant comme mesure locale de dissimilarité la distance de Hausdorff car les propriétés de la DH permettent de rendre compte d'une notion de dissimilarité locale et rendent de ce fait la présentation plus intuitive. Le principe de la carte de dissimilarité (qui résulte de cette mesure locale) ne repose cependant pas sur les propriétés particulières de la DH et dans la sous-section 2.4.3, une définition générale de la carte de dissimilarités est proposée, ne reposant pas sur la DH mais pouvant utiliser toute mesure d'erreur entre deux images binaires respectant certaines propriétés.

1.3 Distance de Hausdorff (DH) et ses variantes

1.3.1 Distance de Hausdorff, généralités

Nous présentons ici la distance de Hausdorff et quelques propriétés utiles qu'elle vérifie. La définition de la DH peut être trouvée def. 1.2.1.

Rappel de notation Soit $X \in \mathbb{R}^2$ le support de l'image (l'ensemble des pixels composant l'image), que l'on suppose fini. Une image binaire est une fonction de X dans $\{0, 1\}$. La valeur 1 sera interprétée comme logiquement vraie et vue à l'écran comme noire. L'image binaire \mathbf{B} peut-être identifiée à l'ensemble $B = \{x/b(x) = 1\}$, ce que l'on fera par la suite.

1.2.1 (p. 54). Pour les images, nous utiliserons les mêmes notations : $D_H(\mathbf{A}, \mathbf{B}) = D_H(A, B)$.

1.3.2 Propriétés générales de la distance de Hausdorff

L'intérêt de cette mesure vient tout d'abord de ses propriétés métriques : positivité, identité, symétrie, et inégalité triangulaire. Ces propriétés correspondent généralement à notre intuition pour la comparaison d'images. En effet, une forme

image est identique à elle-même, et l'ordre de comparaison n'a pas d'importance en général; dans le cas où l'ordre de comparaison est important, la propriété d'asymétrie de la DH directe peut être exploitée. Enfin, l'inégalité triangulaire évite qu'une image soit similaire à deux images complètement dissimilaires en même temps.

De plus, la distance de Hausdorff ne nécessite pas l'appariement des points pour effectuer la mesure, elle est donc robuste aux déformations locales non rigides. Une autre qualité de la DH provient de la propriété suivante :

Proposition 1.3.1 (Translation). *Soit v un vecteur de \mathbb{R}^2 , T_v la translation de vecteur v et A un ensemble de points non vide, alors*

$$D_H(A, T_v A) = \|v\|. \quad (1.15)$$

Démonstration. Comme $A = T_{-v}(T_v A)$, il est équivalent de montrer $D_H(A, T_v A) = \|v\|$ et $h(A, T_v A) = \|v\|$. Montrons ce dernier point.

– montrons que $h(A, T_v A) \leq \|v\|$.

$$\begin{aligned} \forall a \in A, \min_{b \in B} d(a, b) &\leq d(a, T_v a) \\ &\leq \|v\| \end{aligned}$$

donc $\max_{a \in A} \min_{b \in B} d(a, b) \leq \|v\|$.

– Maintenant, montrons que $h(A, T_v A) \geq \|v\|$. Supposons que $h(A, T_v A) = r < \|v\|$ alors $T_v A \subset \bigcup_{a \in A} B(a, r)$. Absurde! \square

Cela implique que pour de petites translations, la DH sera petite, ce qui correspond à notre attente pour une mesure de dissimilarité.

1.3.3 Quelques versions modifiées de la distance de Hausdorff

La DH classique a de bonnes propriétés mais elle mesure; entre deux ensembles de points; les points les moins bien appariés et en conséquence, elle est sensible au bruit [Paumard, 1997]. En effet, prenons comme exemple deux images contenant la même forme, et rajoutons un point à la première image, éloigné de la forme. Alors la DH va mesurer la distance entre ce point et la forme (de la deuxième image).

De nombreuses modifications de la distance de Hausdorff ont été proposées pour l'améliorer, nous pouvons citer la DH partielle [Huttenlocher et al., 1993], la DH modifiée [Dubuisson and Jain, 1994], la DH censurée [Paumard, 1997], la DH « doublement » modifiée [Takacs, 1998], la DH utilisant les moindres carrés [Sim et al., 1999] et la DH pondérée [Lu et al., 2001]. La DH a aussi été modifiée afin de comparer des images composées de traits [Chen et al., 2003], ou de courbes, nous ne l'aborderons pas car cela sort du cadre de notre étude. Les définitions suivantes sont détaillées dans [Zhao et al., 2004].

DH partielle La distance directe de la DH partielle est définie dans [Huttenlocher et al., 1993] :

$$h_K(A, B) = K_{a \in A}^{i\grave{e}me} d(a, B) \quad (1.16)$$

où $K_{a \in A}^{i\grave{e}me}$ désigne la K ième valeur par ordre croissant de $d(a, B)$. La DH partielle a été utilisée pour la mise en correspondance d'objet avec occlusion éventuelle. Elle donne de bons résultats en présence de bruit impulsionnel car un choix judicieux du paramètre $f = \frac{K}{N_a}$ permet de ne pas prendre en compte les points marginaux. Ce paramètre est compris entre 0 et 1. Expérimentalement, quand f vaut 0,6, les résultats obtenus pour la mise en correspondance sont bons. Cependant le paramètre f doit être fixé par l'utilisateur en fonction du type d'images et il se peut qu'il n'y ait pas de meilleur choix pour f .

DH modifiée La distance directe de la DH modifiée est définie dans [Dubuisson and Jain, 1994] :

$$h_{DHM}(A, b) = \frac{1}{N_a} \sum_{a \in A} d(a, B) \quad (1.17)$$

où $N_a = \text{card}(A)$. La DH modifiée ne nécessite pas de paramètre contrairement à la DH partielle. Les auteurs ont comparé son comportement avec des variantes de la DH dont la DH partielle en présence de bruit. Leur conclusion est que même la DH partielle manque de robustesse face au bruit, mais que la DH modifiée a un meilleur comportement pour des applications réelles.

DH pondérée La distance directe de la DH pondérée est définie dans [Zhao et al., 2004] :

$$h_{DHP}(A, B) = \frac{1}{N_a} \sum_{a \in A} w(a) \cdot d(a, B) \quad (1.18)$$

où $\sum_{a \in A} w(a) = N_a$. Dans une image peuvent se distinguer plusieurs zones qui n'ont pas la même importance (dans une reconnaissance de visages par exemple, les zones des yeux et de la bouche ont plus d'importance que la zone du front), leur poids est alors différent dans la DH pondérée. Cette distance a été utilisée pour la reconnaissance de caractères chinois [Lu et al., 2001; Lu and Tan, 2002] et pour la reconnaissance de visages [Guo et al., 2001].

DH censurée Soit $(\alpha, \beta) \in [0, 1]^2$ et $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ où les x_i sont rangés par ordre croissant de valeur. On note $N_\alpha = E(\alpha \times n)$ où E désigne la partie entière et $Q_\alpha = x_{N_\alpha}$. La distance directe de la DH censurée est définie dans [Paumard, 1997] :

$$h_{\alpha, \beta}(A, B) = Q_{1-\beta}\{d_\alpha(a, B), a \in A\}, \quad (1.19)$$

$$\text{avec } d_\alpha(a, B) = Q_\alpha\{d(a, b), b \in B\}. \quad (1.20)$$

où typiquement $\alpha = 1\%$ et $\beta = 10\%$. Maintenant, si $h_{\alpha, \beta}(A, B) = R$, cela signifie que 90% des points de A ont 1% des points de B dans une boule de rayon R . Comme la DH censurée classe les distances entre les points, l'effet du bruit impulsionnel est réduit. Cependant, comme pour la DH partielle, cette distance fait appel à deux paramètres.

DH doublement modifiée La distance directe de la DH doublement modifiée est définie dans [Takács, 1998] :

$$h_M = \frac{1}{N_a} \sum_{a \in A} d(a, B) \quad (1.21)$$

avec $d(a, B) = \max(\mathbf{I} \min_{b \in N_B^a} d(a, b), (1 - \mathbf{I})P)$

où \mathbf{I} indique s'il existe un point $b \in N_B^a$, N_B^a est un voisinage du point a dans l'ensemble B et P est une pénalité. La notion de similarité présentée par la DH doublement modifiée est locale : les images A et B sont proches si tout point de A est près d'un point de B et *vice et versa*. Elle requiert que les points en correspondance se retrouvent dans le voisinage l'un de l'autre. Si ce n'est pas le cas, une pénalité P est alors attribuée pour assurer que les images trop différentes soient pénalisées. L'application à laquelle est destinée la DH doublement modifiée est la reconnaissance de visages, l'auteur suppose donc que les déformations sont locales, ce qui justifie l'usage des voisinages et de la pénalité.

DH de moindres carrés La distance directe de la DH de moindres carrés est définie dans [Sim et al., 1999] par une combinaison linéaire de statistiques d'ordre :

$$h_{mc}(A, B) = \frac{1}{K} \sum_1^K d(a, B)_{(i)} \quad (1.22)$$

où K désigne $f \times N_A$, avec le paramètre f défini comme dans le cas de la DH partielle, et où $d(a, B)_{(i)}$ représente la i ème valeur de distance dans la suite rangée par ordre croissant $d(a, B)_{(1)} \leq d(a, B)_{(2)} \leq \dots \leq d(a, B)_{(N_a)}$. La mesure $h_{mc}(A, B)$ est minimisée du fait que seules les faibles valeurs sont conservées alors que les plus grandes sont éliminées. Ainsi, même si l'objet est caché ou dégradé, la méthode de comparaison donne de bons résultats. Si f est fixé à 1, la DH de moindres carrés revient à la DH modifiée.

1.3.4 Discussion

Nous devons souligner le fait que, à l'exception de la DH modifiée, il y a au moins un paramètre à fixer pour que les méthodes fonctionnent : K pour la DH partielle, la taille du voisinage N_B^a pour la DH doublement modifiée, la fonction de poids pour la DH pondérée ou le rang pour la DH de moindres carrés. Le paramètre, ou la série de paramètres, doit être fixé dans le but de rendre la mesure aussi discriminante que possible. Ceci, bien sûr, en fonction du type d'image traité, voire en fonction des images au sein d'une même application (par exemple, des images plus ou moins sombres ou plus ou moins bruitées pourront nécessiter des réglages différents). La DH modifiée ne nécessite pas le réglage d'un paramètre, cependant, ses performances en comparaison d'images ne sont pas aussi bonnes que celles de la DH partielle et de la DH censurée, à cause de l'opération de sommation sur toutes les distances parmi lesquelles des points éloignés peuvent intervenir.

De plus, ces mesures sont globales et ne permettent pas de distinguer les dissimilarités locales. En effet, la DH est une distance max-min. Cela signifie que la valeur de la DH entre deux images est atteinte pour au moins un couple de points. *Mais cela ne dit pas si cette valeur est atteinte en plusieurs endroits ou si elle ne concerne qu'un seul couple de points*, ce qui correspond à plusieurs degrés de similarité. Par exemple, deux images qui sont identiques sauf pour un détail (ou pour une série de détails localisés dans une zone des images) peuvent donner la même mesure globale que deux images présentant des dissimilarités partout. Ces constats ont motivé la définition d'une DH locale non paramétrique dans le chapitre suivant.

2

Mesure locale de la DH

Dans ce chapitre, la notion de dissimilarité locale est tout d'abord précisée, puis une définition naïve de la mesure locale de la DH est donnée. Elle est une simple adaptation de la mesure de la DH globale à une fenêtre locale. Cependant la DH n'est pas définie si l'un des deux ensembles est vide, ce qui peut arriver dans le cas de la mesure dans une fenêtre locale. Cela amène à une modification de la définition naïve. D'autre part, lors de l'utilisation de la DH locale, la fenêtre va se déplacer, et sa taille va éventuellement varier. Pour que les valeurs prises par la DH restent cohérentes lors de ces variations, une dernière modification est opérée et finalement, cela conduit à une définition de la mesure locale de la DH cohérente.

Notations : dans tout le chapitre, A et B désignent deux ensembles non vides de points de \mathbb{R}^2 et W une partie convexe fermée de \mathbb{R}^2 .

2.1 Quel sens donner à l'expression « dissimilarité locale » ?

Faire localement une mesure de dissimilarité implique de comparer deux images localement. Cela peut être fait grâce à une fenêtre glissante : les parties de chaque image vues au travers de cette fenêtre locale sont alors comparées avec une mesure de dissimilarité.

La taille de la fenêtre glissante joue un rôle important car elle doit être ajustée à la taille de la dissimilarité locale de telle manière que la mesure puisse l'évaluer. Ici, nous avons besoin de préciser ce que l'on entend par dissimilarité locale pour savoir comment ajuster la fenêtre. Malheureusement, cette notion fait appel à la sémantique, et elle est bien difficile à préciser. De plus, notre étude se situant à bas-niveau, il n'est pas nécessaire d'approfondir cette notion haut-niveau. Au mieux pouvons-nous en donner une idée générale : si les pixels situés dans la fenêtre glissante appartiennent à des traits grossiers, la fenêtre devra avoir une taille suffisante pour englober la dissimilarité résultant de la comparaison de ces traits. D'un autre côté, pour des zones où les traits sont fins, il faut que la fenêtre ne soit pas trop grosse par rapport à l'échelle des traits, sinon, elle risque de prendre en compte des dissimilarités qui ne sont plus locales et qui vont fausser la valeur de la mesure dans la fenêtre. Ainsi, il est nécessaire d'adapter la taille de la fenêtre de sorte à obtenir toute l'information sur la dissimilarité locale, et elle seule. Nous voyons figure 2.1 un exemple de comparaison simple : on souhaite mesurer localement la dissimilarité au centre des deux images A et B . Avec la petite fenêtre et la moyenne, les extraits de A à travers W sont entièrement noir ceux de B sont entièrement blanc. Ces extraits ne peuvent rendre compte de la taille de la dissimilarité au centre des deux images. La plus grande des fenêtres est la fenêtre adéquate pour avoir toute l'information pour faire la mesure locale. Le choix de la taille de la fenêtre est un point important de la mesure et il est détaillé dans le paragraphe 2.4.2.

Le sens du mot « locale » dans l'expression « dissimilarité locale » mérite lui aussi quelques précisions. Nous faisons l'hy-

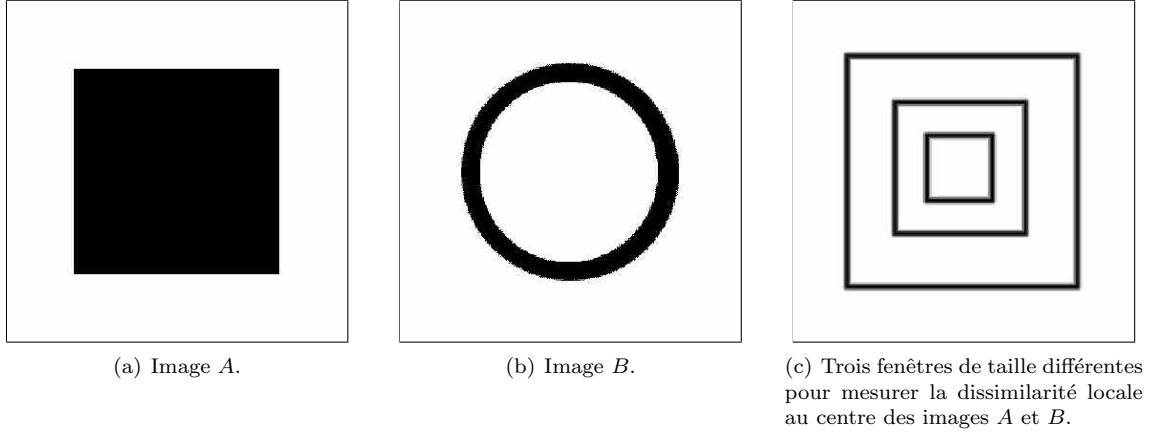


FIG. 2.1 – Illustration de la notion de dissimilarité locale : les images A et B sont dissimilaires et les fenêtres petites et moyennes ne permettent pas d’appréhender la dissimilarité locale au centre des images car elles sont trop petites.

pothèse qu’une dissimilarité locale doit concerner l’information ayant trait aux pixels centraux dans la fenêtre. La définition de pixel central dépend à la fois de la taille minimale de la fenêtre et du pas de translation (qui sont aussi reliés, dans la mesure où le pas ne peut pas être pris plus grand que la taille minimale de fenêtre, au risque de passer à côté de parties de l’image). Dans notre étude, nous avons fait le choix de prendre un pas de 1 pixel et une taille minimale de fenêtre égale à 1×1 pixel, ce qui amène naturellement à ne prendre qu’un seul pixel central. Notre choix se justifie de la manière suivante : un pas plus grand que 1 entraîne une perte d’information implicite si la taille de la fenêtre initiale est inférieure au pas et si cette taille est adaptée, elle donne une fenêtre initiale asymétrique. Par exemple, pour un pas de 2, la fenêtre initiale devra être de taille 2×2 avec le pixel central en haut à droite, ce qui conduit à une mesure anisotrope et aussi plus grossière. Nous préférons pouvoir contrôler la perte d’information en réalisant un sous-échantillonnage explicite (voir partie 1) avec une AMR et prendre un pas de 1 pixel ainsi qu’une fenêtre initiale 1×1 pixel.

2.2 Définition de la DH locale

2.2.1 Définition naïve de la DH locale

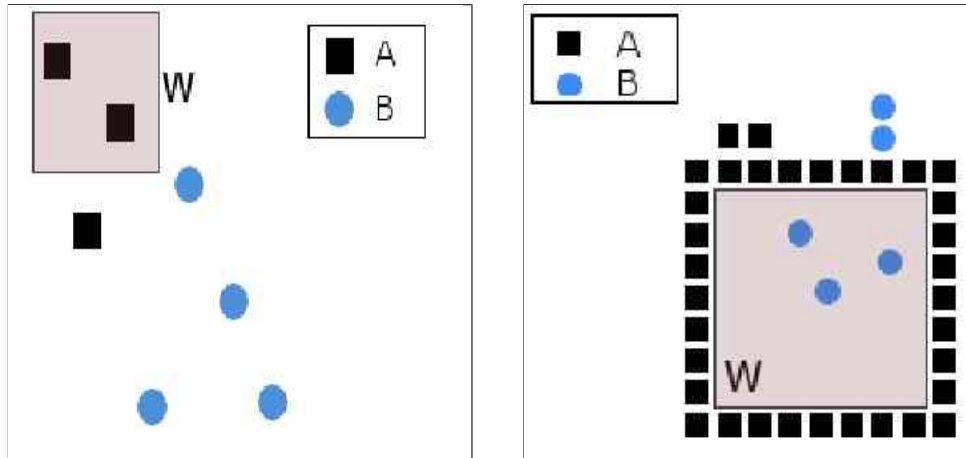
Nous souhaitons mesurer la DH non plus entre les deux images complètes, mais entre deux extraits par une fenêtre. Il faut donc définir la mesure de la DH dans une fenêtre. Cela revient à modifier la définition de la mesure globale (def. 1.2.1) en introduisant une restriction des ensembles de points à la fenêtre :

Définition 2.2.1 (DH restreinte à une fenêtre (naïve)).

$$HD_W(A, B) = \max(h_W(A, B), h_W(B, A)) \quad (2.1)$$

$$\text{où } h_W(A, B) = \max_{a \in A \cap W} \left(\min_{b \in B \cap W} d(a, b) \right). \quad (2.2)$$

Cette définition est naïve dans la mesure où tout est fait comme si ce qui est extrait à travers la fenêtre constituait encore une image. Mais il se peut que l’une des images n’ait aucun point (noir) dans la fenêtre (voir figure 2.2(a)). Or la distance de Hausdorff n’est pas définie pour des ensembles vides. Dans un premier temps (en 2.2.2) nous allons donner une définition



(a) Exemple de cas critique pour la DH locale. Les deux images a et b ont été superposées. L'ensemble A est représenté par des carrés noirs et l'ensemble B par des ronds bleus. Dans sa position actuelle, la fenêtre W (dans le coin supérieur gauche) ne comprend que des points de A et on a $B \cap W = \emptyset$. La mesure naïve n'est pas définie dans ce cas car elle comporte un maximum et un minimum sur B .

(b) Exemple pour l'amélioration de la définition naïve. Quelle valeur attribuer à la mesure lorsque l'un des ensembles de points est vide dans la fenêtre? Cette valeur dépend de la disposition des points à l'extérieur. Le disposition qui donne la valeur la plus faible est celle présentée ci-dessus.

FIG. 2.2 – Deux exemples pour la mise en forme de la définition améliorée.

modifiée, valable lorsque l'un des deux ensembles n'a pas de point dans la fenêtre W . Ceci ne sera pas suffisant : la fenêtre va glisser sur les images et à chaque position, éventuellement, s'agrandir. Avec la définition modifiée, cela va engendrer des variations brusques de valeurs, indésirables car elles ne reflètent pas la réalité. En conséquence, dans un deuxième temps (en 2.2.3), nous adapterons la définition pour que le comportement de la mesure soit cohérent lors du déplacement et de l'agrandissement de la fenêtre dans les images.

2.2.2 Modification de la définition naïve

Nous avons ici à résoudre le cas où l'un des deux ensembles de points contenus dans la fenêtre W est vide. Il n'est pas possible de se référer à la définition globale qui exclut ce cas. Le problème est d'attribuer une valeur à une mesure entre un ensemble non vide de points et un ensemble vide.

- La valeur 0 ne convient pas car les deux ensembles sont différents.
- La valeur infinie ne convient pas non plus car les ensembles sont extraits de deux ensembles de points plus grands (les deux images) sur lesquels on obtient une mesure finie.

Comme la mesure est faite sur un ensemble de points (de A et de B confondus) plus petit, la mesure doit être inférieure à la mesure globale. Cependant, cette mesure globale dépend de la disposition des points à l'extérieur de la fenêtre. Comme il n'est pas possible de tenir compte de toutes les dispositions éventuelles des points extérieurs à la fenêtre, une solution est de se placer dans le cas où la disposition des points extérieurs à la fenêtre est la plus défavorable, dans le sens où elle donnerait (si les points de l'extérieur étaient pris en compte) la valeur la plus faible. Ce cas, qui est illustré par la figure 2.2(b), est celui où les points de l'ensemble qui est absent de l'intérieur de la fenêtre, bordent entièrement la fenêtre.

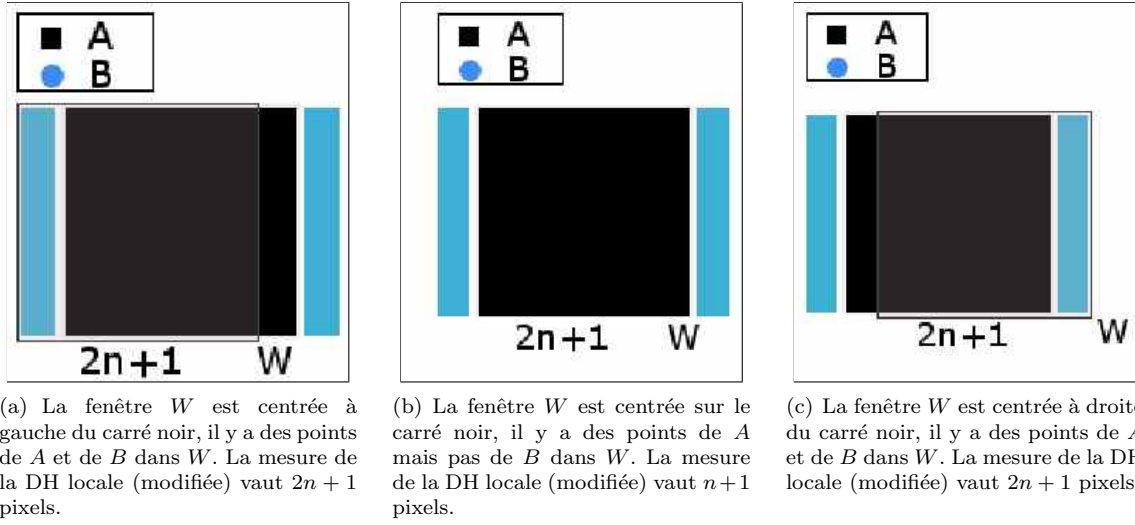


FIG. 2.3 – Illustration des sauts de valeurs pour la définition modifiée lors du déplacement de la fenêtre W , La valeur mesurée passe successivement de $2n$ à $n + 1$ puis à $2n$ pixels. Pour la version finale de la distance, on obtient n , $n + 1$ et n .

Cette solution présente l'avantage de fournir une valeur inférieure à celle de la DH globale entre les deux images. Et elle revient simplement à prendre en compte la distance au bord de la fenêtre pour traduire notre ignorance de ce qu'il y a en dehors de la fenêtre. Cela conduit à la définition suivante :

Définition 2.2.2 (amélioration).

$$HD_W(A, B) = \max(h_W(A, B), h_W(B, A)) \quad (2.3)$$

$$\text{où } h_W(A, B) = \begin{cases} \max_{a \in A \cap W} (\min_{b \in B \cap W} d(a, b)) & \text{si } A \neq \emptyset \text{ et } B \neq \emptyset \\ \max_{a \in A \cap W} (\min_{w \in Fr(W)} d(a, w)) & \text{si } A \neq \emptyset \text{ et } B = \emptyset \\ 0 & \text{si } A = \emptyset \end{cases} \quad (2.4)$$

La distance globale est toujours le maximum des deux distances directes, et la distance directe présente deux cas particuliers correspondant aux éventuels ensembles vides. La frontière Fr provient de la topologie définie par la distance spatiale d . Dans les applications (partie 3), les distances utilisées seront les distances associées aux normes L_2 et L_∞ .

2.2.3 Distance de Hausdorff locale

Le cas où l'un des deux ensembles est vide est maintenant clarifié. Cependant la définition dans le cas général doit être modifiée pour rester cohérente lorsque la fenêtre glisse sur les images. En effet, avoir deux cas avec des mesures de distance différentes (distance au bord dans un cas et DH dans l'autre) dans la définition apporte des variations brusques de la valeur de la distance lors du déplacement de la fenêtre (figure 2.3) ou lors de son agrandissement (figure 2.4) qui ne correspondent pas à l'intuition .

Il est donc nécessaire que la distance à la frontière de W soit aussi prise en compte dans la définition générale.

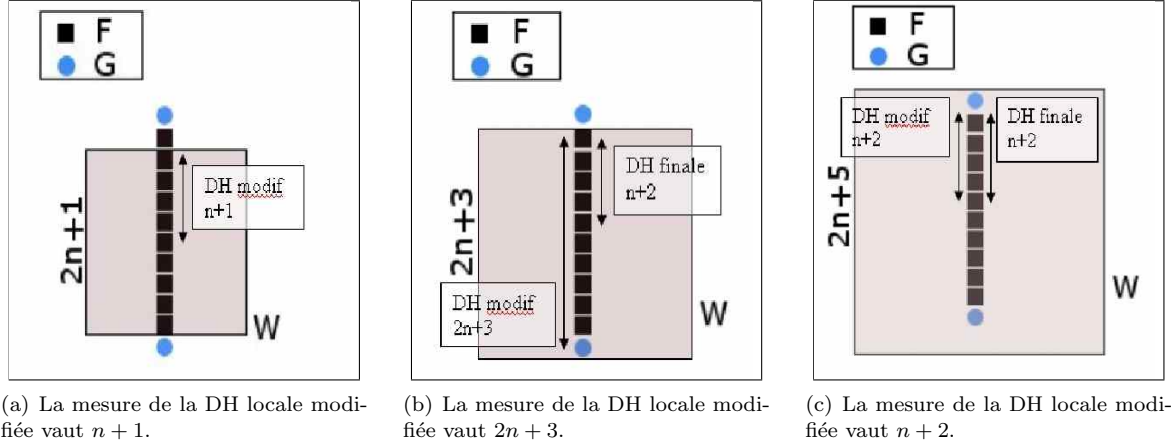


FIG. 2.4 – Le cas 2.4(a) où A n'a pas de point dans W donne comme mesure la distance au bord et le cas 2.4(b) où les deux ensembles ont des points dans W donne comme mesure la DH. Dans la cas 2.4(c), un autre point de B est pris en compte et la valeur de la DH locale rebaisse à $n + 2$. Cela donne un saut de valeur contre-intuitif.

Avec la version finale de la mesure, on obtient $n + 1$, $n + 2$ et $n + 2$.

Définition 2.2.3 (Version finale : la distance de Hausdorff locale). Soient A, B deux ensembles bornés de points de \mathbb{R}^2 .

$$HD_W(A, B) = \max(h_W(A, B), h_W(B, A)) \quad (2.5)$$

$$\text{où } h_W(A, B) = \begin{cases} \max_{a \in A \cap W} [\min(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w))] & \text{si } A \neq \emptyset \text{ et } B \neq \emptyset \\ \max_{a \in A \cap W} [\min_{w \in Fr(W)} d(a, w)] & \text{si } A \neq \emptyset \text{ et } B = \emptyset \\ 0 & \text{si } A = \emptyset \text{ et } \forall B \end{cases} \quad (2.6)$$

remarque 2.2.1. – Dans le cas où il n'y a pas de point de A ni de B dans W , chacune des deux distances directes est nulle et la distance globale aussi par conséquent. Ce qui est cohérent avec le fait que les parties extraites des deux images sont égales.

– Dans le cas où exactement l'un des deux ensembles n'a pas de points dans W , alors l'une des deux distances directes est nulle et l'expression de l'autre tient compte de la distance au bord.

2.3 Propriétés de la DH locale HD_W

Dans cette section, des propriétés utiles de la DH locale sont démontrées. Le critère développé dans la section 2.4.2 qui doit permettre d'ajuster la taille de la fenêtre W s'appuie sur ces propriétés :

- HD_W est comprise entre 0 and $HD(A, B)$,
- si l'on considère une suite croissante de fenêtres emboîtées $W_1 \subset W_2 \subset \dots \subset W_n$, alors la suite de valeurs $(HD_{W_i})_{1 \leq i \leq n}$ est croissante.

2.3.1 Propriétés générales

HD_W est positive et symétrique par définition (définition 2.2.3).

Proposition 2.3.1 (Identité). *Soit A, B deux ensembles bornés de points de \mathbb{R}^2 , et W une partie convexe fermée de \mathbb{R}^2 .*

$$HD_W(A, B) = 0 \iff A \cap W = B \cap W \quad (2.7)$$

Démonstration. voir sec. A.1, p. 147. □

Ainsi dire que la valeur de la DH locale est nulle revient à dire que les deux images sont identiques dans la fenêtre locale. Par ailleurs, la construction de la DH locale permet d'avoir une propriété de majoration de sa valeur par la DH globale qui est présentée ci-dessous.

Proposition 2.3.2 (Majoration). *Soit $x \in \mathbb{R}^2$ et $r > 0$, et soit $W = B(x, r)$ un disque fermé de rayon r et de centre x alors*

$$HD_W(A, B) \leq HD(A, B).$$

Démonstration. voir sec. A.2, p. 148. □

Ainsi, quand la fenêtre W parcourt les deux images, les valeurs dans la carte des dissimilarités locales qui en résulte restent comprises entre 0 et $HD(A, B)$. Considérons maintenant les propriétés liant la DH locale et l'agrandissement de la fenêtre locale.

2.3.2 Propriétés dépendant de la taille de la fenêtre W

Proposition 2.3.3 (croissance). *Soit $V = B(x_v, r_v)$ et $W = B(x_w, r_w)$ deux disques fermés tels que $V \subset W$ alors $HD_V(A, B) \leq HD_W(A, B)$.*

Démonstration. voir sec. A.3, p. 151. □

2.4 Une DH locale adaptative et non-paramétrique

En nous appuyant sur le paragraphe 2.1, nous allons préciser la notion de dissimilarité locale mathématiquement afin de donner un critère pour le choix de la taille de fenêtre optimale. Ceci est l'objectif des prochains paragraphes.

2.4.1 Caractérisation de la mesure d'une dissimilarité locale

Dans tout le paragraphe, A et B désignent deux ensembles de points de \mathbb{R}^2 non vides. La proposition suivante spécifie les conditions sur A et B pour que $HD_W(A, B)$ soit maximum. Puis, la notion de dissimilarité locale au sens de la distance de Hausdorff est définie. Ce qui permet ensuite de définir la meilleure taille pour la fenêtre $W = B(x, r)$ en fonction de A et B pour mesurer au mieux la dissimilarité locale.

Lemme 2.4.1 (valeur maximale). *Soit $x \in \mathbb{R}^2$ et $r > 0$, et soit une fenêtre $B(x, r) = \{y / \|x - y\| < r\}$ alors*

$$\sup_{A, B} (HD_{B(x, r)}(A, B)) = r$$

Démonstration. Montrons tout d'abord que $HD_{B(x,r)}(A, B) \leq r$.

Par définition, nous savons que

$$HD_{B(x,r)}(A, B) = \max(h_{B(x,r)}(A, B), h_{B(x,r)}(B, A)), \quad (2.8)$$

donc il suffit de montrer que les distances directes sont inférieures à r . Considérons $h_{B(x,r)}(A, B)$. D'après la définition 2.2.3, trois cas se distinguent

1 $A \cap B(x, r) \neq \emptyset$ et $B \cap B(x, r) \neq \emptyset$

alors

$$h_{B(x,r)}(A, B) = \max_{a \in A \cap B(x,r)} \left[\min \left(\min_{b \in B \cap B(x,r)} d(a, b), \min_{w \in Fr(B(x,r))} d(a, w) \right) \right]. \quad (2.9)$$

Soit $a \in A \cap B(x, r)$, considérons

$$\min_{w \in Fr(B(x,r))} d(a, w) = d(a, Fr(B(x, r))),$$

par définition de $B(x, r)$,

$$\forall y \in B(x, r), d(y, Fr(B(x, r))) \leq r,$$

donc

$$\min_{w \in Fr(B(x,r))} d(a, w) \leq r$$

en remplaçant dans 2.9, on obtient $h_{B(x,r)}(A, B) \leq r$ et en remplaçant dans 2.8, on obtient le résultat.

2 $A \cap B(x, r) \neq \emptyset$ et $B \cap B(x, r) = \emptyset$

alors, on a

$$h_{B(x,r)}(A, B) = \max_{a \in A \cap B(x,r)} \left[\min_{w \in Fr(B(x,r))} d(a, w) \right] \quad (2.10)$$

Comme la seule distance en jeu est la distance au bord, il est clair que l'on a :

$$h_{B(x,r)}(A, B) \leq r \quad (2.11)$$

3 $A = \emptyset$

Alors

$$h_{B(x,r)}(A, B) = 0 \quad (2.12)$$

d'où le résultat. Il suffit maintenant de trouver pour x et r donnés, A et B tels que $HD_{B(x,r)}(A, B) = r$. Soit x et r fixés, définissons $A = x$ et $B = \emptyset$, considérons la distance directe :

$$h_{B(x,r)}(A, B) = \max_{a \in A \cap B(x,r)} \left[\min_{w \in Fr(B(x,r))} d(a, w) \right]$$

or $A \cap B(x, r) = \{x\}$ et $\forall w \in Fr(B(x, r)), d(x, w) = r$, donc

$$\min_{w \in Fr(B(x, r))} d(x, w) = r$$

et donc

$$h_{B(x, r)}(A, B) = r$$

et par 2.8,

$$HD_{B(x, r)}(A, B) = r.$$

□

Proposition 2.4.1 (valeur maximale). *Soit $x \in \mathbb{R}^2$ et $r > 0$, et soit une fenêtre $W = B(x, r)$ alors*

$$\sup_{A, B} (HD_{B(x, r)}(A, B))$$

est atteinte si et seulement si $x \in A$ (resp. $x \in B$) et aucun point de B (resp. A) excepté peut-être sur la frontière $Fr(W)$ de W . Nous avons alors $HD_{B(x, r)}(A, B) = r$.

Démonstration. Nous avons déjà par le lemme 2.4.1 que

$$\begin{aligned} \sup_{A, B} (HD_{B(x, r)}(A, B)) &= \max_{A, B} (HD_{B(x, r)}(A, B)) \\ &= r. \end{aligned}$$

Il suffit donc de montrer que le maximum est atteint dans le cas cité et uniquement dans ce dernier. Montrons tout d'abord qu'il est atteint dans le cas cité.

Les rôles de A et B étant symétriques, nous démontrerons seulement le cas où *il y a exactement un point de A au centre de W et aucun point de B excepté peut-être sur la frontière $Fr(W)$ de W* . Nous nous plaçons maintenant sous cette hypothèse. Nous savons alors que le maximum va être atteint par la distance directe $h_W(A, B)$ et non par $h_W(B, A)$ (car nous avons déjà fait le calcul...), calculons donc

$$h_W(A, B) = \max_{a \in A \cap W} \left[\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right].$$

Par hypothèse, $\exists! a \in A \cap W$ et $a = x$, donc

$$h_W(A, B) = \min \left(\min_{b \in B \cap W} d(x, b), \min_{w \in Fr(W)} d(x, w) \right),$$

intéressons nous à chacun des deux minima en jeu

– pour le premier, comme $B \cap W \subset Fr(W)$, si $B \cap W \neq \emptyset$ alors

$$\min_{b \in B \cap W} d(x, b) = r$$

et si $B \cap W = \emptyset$, ce minimum n'est pas défini.

– Pour le deuxième, comme $W = B(x, r), \forall w \in Fr(W), d(x, w) = r$ donc

$$\min_{w \in Fr(W)} d(x, w) = r$$

en remplaçant les minima par leur expressions dans la définition de $h_W(A, B)$, on obtient

$$h_W(A, B) = r$$

et donc

$$DH_W(A, B) = r.$$

Montrons maintenant que le maximum n'est atteint que dans ce cas, pour cela, prenons la négation de l'hypothèse et on montre que le maximum n'est pas atteint.

La proposition initiale est

$$[(x \in A) \& (B \cap W \subset Fr(W))] \vee [(x \in B) \& (A \cap W \subset Fr(W))] \quad (2.13)$$

sa négation est donc

$$[(x \notin A) \vee (B \cap W \not\subset Fr(W))] \& [(x \notin B) \vee (A \cap W \not\subset Fr(W))] \quad (2.14)$$

il y a donc quatre cas à étudier dans la proposition 2.14 :

1. $(x \notin A) \& (x \notin B)$
2. $(x \notin A) \& (A \cap W \not\subset Fr(W))$
3. $(x \notin B) \& (B \cap W \not\subset Fr(W))$
4. $(B \cap W \not\subset Fr(W)) \& (A \cap W \not\subset Fr(W))$

Comme les rôles de A et B sont symétriques, il est possible de regrouper les cas 2. et 3. en ne traitant que l'un des deux. Il reste donc 3 cas. Nous traitons ces trois cas séparément pour montrer que dans chaque cas, la distance maximale r ne peut pas être atteinte par $DH_W(A, B)$

1. $(\mathbf{x} \notin \mathbf{A}) \& (\mathbf{x} \notin \mathbf{B})$

Nous avons par définition

$$HD_W(A, B) = \max(h_W(A, B), h_W(B, A)),$$

Il faut donc que les deux distances directes soient inférieures à r . Nous allons montrer que $x \notin A \Rightarrow (h_W(A, B) < r)$, et nous aurons alors, de la même manière $x \notin B \Rightarrow (h_W(B, A) < r)$, ce qui implique le résultat. Montrons que $x \notin A \Rightarrow (h_W(A, B) < r)$

– Si $A = \emptyset$, alors par définition (def 2.4, p. 64)

$$h_W(A, B) = 0,$$

– Si $A \neq \emptyset$,

$$\forall a \in A \cap W, \forall w \in Fr(W), d(a, w) < r$$

et comme A est fini, nous avons donc

$$\forall a \in A \cap W, \min_{w \in Fr(W)} (d(a, w)) < r$$

or

$$h_W(A, B) \leq \max_{a \in A \cap W} \left(\min_{w \in Fr(W)} d(a, w) \right). \quad (2.15)$$

de la même manière,

$$x \notin B \Rightarrow (h_W(B, A) < r)$$

Donc

$$h_W(A, B) < r$$

2. $(x \notin A) \& (A \cap W \not\subset Fr(W))$

Nous avons alors, de la même manière que dans le cas précédent,

$$(x \notin A) \Rightarrow (h_W(A, B) < r).$$

Nous allons maintenant prouver que

$$(A \cap W \not\subset Fr(W)) \Rightarrow (h_W(B, A) < r).$$

Supposons donc que $A \cap W \not\subset Fr(W)$ (donc $A \cap W \neq \emptyset$)

– Si $B \cap W = \emptyset$, alors, par définition de $h_W(B, A)$, $h_W(B, A) = 0 < r$

– Si $B \cap W \neq \emptyset$, alors

$$h_W(B, A) = \max_{b \in B \cap W} \left[\min \left(\min_{a \in A \cap W} d(a, b), \min_{w \in Fr(W)} d(b, w) \right) \right].$$

nous allons utiliser l'un ou l'autre des deux minima selon que le point de B est au centre de la fenêtre ou non :

si $x \in B$, pour $b = x$, comme $A \cap W \not\subset Fr(W)$, $\exists a_0 \in A \cap W, a_0 \notin Fr(W)$ et donc

$$d(a_0, b) < r,$$

d'où

$$\min_{a \in A \cap W} d(a, b) < r,$$

et donc

$$\min \left(\min_{a \in A \cap W} d(a, b), \min_{w \in Fr(W)} d(b, w) \right) < r.$$

si $b \in B \neq x$, alors

$$\min_{w \in Fr(W)} d(b, w) < r$$

et donc

$$\min \left(\min_{a \in A \cap W} d(a, b), \min_{w \in Fr(W)} d(b, w) \right) < r$$

En regroupant ces deux cas, nous avons donc

$$\forall b \in B \cap W, \min \left(\min_{a \in A \cap W} d(a, b), \min_{w \in Fr(W)} d(b, w) \right) < r,$$

d'où

$$\max_{b \in B \cap W} \left[\min \left(\min_{a \in A \cap W} d(a, b), \min_{w \in Fr(W)} d(b, w) \right) \right] < r$$

i.e.

$$h_W(B, A) < r.$$

Ce qui permet de conclure pour le cas 2.

3. $(\mathbf{B} \cap \mathbf{W} \not\subset \mathbf{Fr}(\mathbf{W})) \& (\mathbf{A} \cap \mathbf{W} \not\subset \mathbf{Fr}(\mathbf{W}))$ Ce dernier cas se déduit facilement de la démonstration faite pour le cas 2. En effet, dans le cas 2., nous avons montré que

$$(A \cap W \not\subset Fr(W)) \Rightarrow (h_W(B, A) < r).$$

Comme les rôles de A et B sont symétriques nous avons de même

$$(B \cap W \not\subset Fr(W)) \Rightarrow (h_W(A, B) < r),$$

d'où

$$\max(h_W(B, A), h_W(A, B)) < r$$

et donc

$$DH(A, B) < r.$$

□

Le but ici est de fournir un critère permettant de fixer la taille de la fenêtre glissante. Comme cela est écrit dans le paragraphe 2.1, on fait l'hypothèse qu'une dissimilarité locale doit faire intervenir des traits dont fait partie le pixel central de la fenêtre W . Nous souhaitons aussi que -dans la mesure du possible- d'autres dissimilarités n'interviennent pas dans la mesure faite dans la fenêtre W (i.e. relatives à d'autres traits dans les images). Ainsi, la mesure doit concerner

- un point central : si le point central n'est pas impliqué, la fenêtre W peut être déplacée pour que un des points concerné par la mesure soit en son centre,
- et un point de la frontière de W : si aucun de ces points n'est impliqué la fenêtre peut être réduite.

Le lemme 2.4.1 montre que dans ce cas, la mesure dans la fenêtre atteint sa valeur maximale. D'où la définition :

Définition 2.4.1 (mesure locale). *On dit que la fenêtre $W = B(x, r)$ donne une mesure locale quand la mesure de la DH dans la fenêtre $B(x, r)$ est maximale : $HD_{B(x, r)}(A, B) = r$.*

Nous voulons maintenant savoir s'il y a une mesure locale *maximale*. Donc, soit $x \in \mathbb{R}^2$ et $r > 0$, Définissons :

Définition 2.4.2 (L'ensemble de mesure locale). *L'ensemble de mesure locale \mathcal{R} est déterminé par*

$$\mathcal{R} = \{r > 0 / HD_{B(x, r)}(A, B) = r\}. \quad (2.16)$$

Quand \mathcal{R} n'est pas vide, il est bien entendu majoré par $HD(A, B)$ (prop 2.3.3), et il a donc une borne supérieure r_{max} . D'où la définition :

Définition 2.4.3 (Mesure locale maximale). *Pour $x \in \mathbb{R}^2$ fixé, si \mathcal{R} n'est pas vide, $r_{max} = \sup(\mathcal{R})$ est appelé le rayon optimal et pour ce rayon, $HD_{B(x, r_{max})}$ donne la mesure locale maximale r_{max} .*

Proposition 2.4.2 (Expression mathématique de la mesure locale maximale). *Pour $x \in \mathbb{R}^2$ fixé, si \mathcal{R} n'est pas vide,*

$$\begin{aligned} HD_{B(x, r_{max})} &= r_{max} \\ &= \max(d(x, A), d(x, B)) \end{aligned} \quad (2.17)$$

2.4.2 Critère pour le calcul du rayon optimal r_{max}

Le but ici est premièrement de fournir un critère pour calculer le rayon optimal r_{max} et deuxièmement de définir l'algorithme pour effectuer le calcul de $HD_{B(x, r_{max})}(A, B)$ numériquement.

Proposition 2.4.3 (nature de l'ensemble \mathcal{R}). *Soit $x \in \mathbb{R}^2$, il y a deux possibilités :*

- $\mathcal{R} = \emptyset$
- $\mathcal{R} = [0, r_{max}]$.

Démonstration. C'est une conséquence du lemme 2.4.1.

- Si x n'appartient ni à A ni à B ou si x appartient à A et B , alors par le lemme 2.4.1, HD_W n'atteindra jamais sa valeur maximale.
- Si x appartient seulement à l'un des deux ensembles A et B , par exemple A . Alors, comme B est fini, $\exists b_{min} \in B \cap W$, le point le plus proche de x (pour la distance d). Par le lemme 2.4.1, nous savons que pour $0 \leq r \leq d(x, b_{min}) = r_m$, W donnera une mesure locale et pour $r > r_m$, comme $\min_{b \in B \cap W}(d(x, b)) = r_m < r$, HD_W ne peut pas atteindre sa valeur maximale. En conséquence, \mathcal{R} est l'intervalle $[0, r_m]$ et $r_m = r_{max}$.

□

Critère 2.4.1. *Soit A et B deux ensembles de points de \mathbb{R}^2 finis non-vides et $x \in \mathbb{R}^2$, en ce qui concerne la mesure locale de la DH au point x , le rayon optimal r_{max} pour la fenêtre $W = B(x, r)$ est*

$$r_{max} = \max_{r > 0}(\{r / HD_{B(x, r)}(A, B) = r\}). \quad (2.18)$$

Nous avons maintenant un critère pour trouver la taille maximale de $W = B(x, r)$ qui permette de mesurer la dissimilarité locale. Ainsi l'étude pour un point fixe et une taille de fenêtre variable se termine, nous allons maintenant nous intéresser au résultat du calcul lorsque celui-ci est effectué pour tous les points du plan ou en pratique, lorsqu'il est effectué pour deux images complètes.

2.4.3 Généralisation

La méthode qui a été utilisée dans cette partie pour la mesure de la DH locale est liée par certains aspects aux propriétés de la distance de Hausdorff. En particulier, la définition du critère d'arrêt de croissance de la fenêtre locale est basée sur le fait que la DH est une distance *max-min*. Cependant, le principe de la mesure locale ne repose pas sur les propriétés de la DH, et il est possible de définir de manière générale la mesure locale à partir d'une mesure quelconque sous réserve qu'elle vérifie certaines propriétés. Malheureusement, cette généralisation est en général paramétrique : il est nécessaire de fixer un paramètre pour définir le critère d'arrêt et les propriétés qui permettent de s'en passer sont liées à la DH. Nous allons maintenant détailler cette généralisation.

Propriétés nécessaires à la généralisation

La méthode de la mesure locale comprend deux aspects :

- la propriété de croissance de la valeur de la mesure dans la fenêtre lorsque la fenêtre croît et celle de majoration de la valeur de la distance par une valeur indépendante de la fenêtre,
- un critère d'arrêt fixant la plus grande valeur pour une mesure de la dissimilarité locale.

Lorsque ces deux aspects sont vérifiés, il est possible de faire une mesure locale de dissimilarité. Pour cela, il reste à définir une mesure au travers d'une fenêtre.

Définition 2.4.4 (mesure de dissimilarité dans une fenêtre (ou "localisée"?)). Soit \mathfrak{B} l'ensemble des images binaires qui vont du support X dans $\{0, 1\}$, soit $\mathcal{F} \subset \wp(X)$ un sous-ensemble de l'ensemble des parties de X . Une mesure de dissimilarité dans une fenêtre est définie par :

$$\Phi^\Delta : \begin{cases} \mathfrak{B} \times \mathfrak{B} \times \mathcal{F} \longrightarrow \mathbb{R}^+ \\ (a, b, F) \longmapsto \Phi_{a,b}^\Delta(F) \end{cases} \quad (2.19)$$

Avec Δ la mesure locale de dissimilarité. Pour un couple d'images fixé, Φ^Δ associe à une fenêtre F la valeur de la mesure locale Δ dans cette fenêtre. Notons bien que cette mesure locale de Δ n'est pas forcément la mesure de Δ entre $a|_F$ et $b|_F$ mais une version adaptée à la mesure dans une fenêtre de Δ . Dans le cas de la DH, Δ est la DH locale et \mathcal{F} est l'ensemble des boules $B(x, r)$ de centre x et de rayon r pour une certaine distance spatiale d . Lorsque le contexte le permet le Δ sera omis et nous noterons simplement Φ .

Proposition 2.4.4. Si Φ vérifie les propriétés suivantes,

$$\forall(a, b) \in \mathfrak{B}^2, \quad \Phi_{a,b}(F) = 0 \Leftrightarrow a|_F = b|_F, \quad (2.20)$$

$$\forall(a, b) \in \mathfrak{B}^2, \exists K > 0, \forall F \in \mathcal{F}, \quad \Phi_{a,b}(F) \leq K, \quad (2.21)$$

$$\forall(a, b) \in \mathfrak{B}^2, \forall(F, G) \in \mathcal{F}^2, \quad F \subset G \Rightarrow \Phi_{a,b}(F) \leq \Phi_{a,b}(G), \quad (2.22)$$

alors, il est possible de définir un critère pour fixer la mesure locale des dissimilarités. Un critère possible est le suivant

Critère 2.4.2. Soit un paramètre $\alpha \in [0, 1]$, si la mesure dans la fenêtre est supérieure à $\alpha \times$ la valeur maximale dans la fenêtre et que la valeur maximale K n'est pas atteinte, alors la fenêtre est agrandie.

Démonstration. la preuve de cette proposition vise simplement à montrer la possibilité d'un critère d'arrêt. En effet, ce dernier ne peut pas être défini avec précision sans faire référence à la mesure de distance utilisée. Et dans le cadre général, elle n'est pas précisée. La démonstration reste donc générale.

La méthode suivie pour la mesure locale est de prendre une fenêtre de taille initiale fixée et de l'agrandir tant que le critère d'arrêt n'est pas satisfait. La première exigence est que la distance entre les deux extraits des images par la fenêtre initiale soit nulle si les extraits sont identiques. Ceci pour que le critère d'arrêt soit satisfait et que la croissance de la fenêtre cesse. Ce qui justifie la propriété 2.20.

Ensuite, la proposition 2.21 est nécessaire pour avoir la certitude que la croissance s'arrête : si la valeur de la distance atteint sa borne supérieure, il est inutile de poursuivre la croissance de la fenêtre.

Enfin, la propriété 2.22 est nécessaire pour pouvoir disposer d'un critère d'arrêt fiable : si la propriété de croissance n'est pas vérifiée, le critère peut être vérifié pour certaines tailles de fenêtre et pas pour d'autres. La valeur retenue dépend alors du pas d'agrandissement de la fenêtre et de sa taille initiale. De plus, si le critère d'arrêt dépend d'un paramètre, de faibles variations de ce dernier peuvent alors engendrer des variations importantes dans les valeurs obtenues. Ceci dit, ce choix n'est

pas anodin : il suppose que la valeur mesurée augmente et cela exclut les valeurs moyennées. Il est néanmoins possible de les intégrer en considérant les moyennes cumulées.

Beaucoup de critères différents sont envisageables et qui dépendent de la distance choisie. Avec ces trois propriétés, il est possible de définir un critère, nous en avons déjà vu un avec la DH : si la mesure dans la fenêtre est maximale et que la valeur maximale K n'est pas atteinte, alors la fenêtre est agrandie.

Ce critère qui convient bien pour une distance *max-min* peut être assoupli pour des distances qui atteignent moins facilement la valeur maximale en introduisant un paramètre $\alpha \in [0, 1]$: si la mesure dans la fenêtre est supérieure à $\alpha \times$ la valeur maximale et que la valeur maximale K n'est pas atteinte, alors la fenêtre est agrandie. \square

Définition 2.4.5 (mesure locale de dissimilarité). *Soit Φ une mesure de dissimilarité dans une fenêtre (basée sur une mesure d) vérifiant les propriétés 2.20, 2.21 et 2.22, soit $(a, b) \in \mathfrak{B}^2$ et $F \in \mathcal{F}$ une fenêtre initiale, alors la valeur obtenue en faisant croître la fenêtre F par un critère du type 2.4.2 est appelée mesure locale de dissimilarité et elle est notée $\delta_d^{a,b}(F)$ ou $\delta_d(F)$ s'il n'y a pas d'ambiguïté sur les images.*

Dans le cas où la fenêtre initiale est de taille 1×1 (comme pour la DH), nous notons x la fenêtre initiale $F = \{x\}$.

Exemple applicatif

La mesure de dissimilarité que nous allons utiliser ici est basée sur la somme des différences pixel à pixel. Cette mesure n'est pas très intéressante comme cela a été montré dans la partie 1.2.2 (p. 50) car elle ne prend pas en compte de distance entre les ensembles de points, mais elle permet d'illustrer le cas général. La définition de la mesure est donnée ici :

$$\forall (a, b) \in \mathfrak{B}^2, \forall F \in \mathcal{F}, \Phi_{DS}(a, b, F) = \sum_{x \in F} |a(x) - b(x)|. \quad (2.23)$$

Proposition 2.4.5. *La mesure locale de différence simple par une fenêtre Φ_{DS} vérifie les trois propriétés de la proposition 2.4.4.*

Démonstration. Nous allons montrer successivement que Φ_{DS} vérifie les trois propriétés.

Soient $(a, b) \in \mathfrak{B}^2$ et $F \in \mathcal{F}$,

(propriété 2.20)

$$\Phi_{DS}(a, b, F) = 0 \Leftrightarrow \sum_{x \in F} |a(x) - b(x)| = 0, \quad (2.24)$$

$$\Leftrightarrow \forall x \in F, |a(x) - b(x)| = 0, \quad (2.25)$$

$$\Leftrightarrow \forall x \in F, a(x) = b(x), \quad (2.26)$$

$$\Leftrightarrow a|_F = b|_F. \quad (2.27)$$

Ainsi, la propriété 2.20 est vérifiée par Φ_{DS} .

(propriété 2.21)

$$\Phi_{DS}(a, b, X) = \sum_{x \in X} |a(x) - b(x)|, \quad (2.28)$$

$$= \sum_{x \in X \setminus F} |a(x) - b(x)| + \sum_{x \in F} |a(x) - b(x)|, \quad (2.29)$$

$$\text{or } \sum_{x \in X \setminus F} |a(x) - b(x)| \geq 0, \quad (2.30)$$

$$\text{donc } \Phi_{DS}(a, b, X) \geq \sum_{x \in F} |a(x) - b(x)|. \quad (2.31)$$

Ainsi nous avons

$$\Phi_{DS}(a, b, F) \leq \Phi_{DS}(a, b, X) \quad (2.32)$$

Ainsi la borne supérieure est $K = \Phi_{DS}(a, b, X)$ et Φ_{DS} vérifie la propriétés 2.21.

(propriété 2.22) Soit $G \in \mathcal{F}$ tel que F soit inclus dans G . Alors pour les mêmes raisons que ci-dessus, nous avons :

$$\Phi_{DS}(a, b, F) \leq \Phi_{DS}(a, b, G). \quad (2.33)$$

Donc le point 2.22 est aussi vérifié par Φ_{DS} .

□

Il est donc possible d'appliquer pour Φ_{DS} le critère 2.4.2.

Si le paramètre α vaut 1, cela signifie que la fenêtre ne s'agrandie que si la valeur de la différence est maximale dans la fenêtre, autrement dit, quand le pixel sont différents 1 à 1 dans la fenêtre. Ce qui est très contraignant. De plus, il peut y avoir une dissimilarité locale sans que chaque pixel soit différent. Il est donc préférable de prendre $\alpha < 1$. L'expérience montre que $\alpha = 0,6$ donne les meilleurs résultats. Cependant, comme nous le verrons dans les illustrations au chapitre suivant, ils sont moins bons que lorsque la DH est utilisée.

2.5 Bilan

L'état de l'art concernant la comparaison d'images nous a orienté, dans le cas des images binaires, vers une comparaison directe. Plusieurs mesures de dissimilarité entre images binaires existent mais elles présentent l'inconvénient d'être globales et, de ce fait, de ne pas permettre l'évaluation des dissimilarités locales. Dans ce but, une mesure locale au travers d'une fenêtre est proposée. L'exemple de la distance de Hausdorff est choisi et la mesure de la DH dans une fenêtre nécessite alors des aménagements. Cette mesure locale dépend de la taille de la fenêtre dans laquelle elle est faite. Les propriétés de croissance et de majoration de la mesure locale sont démontrées afin de pouvoir fixer la fenêtre à une taille permettant d'évaluer la dissimilarité locale et elle seule. Cela conduit à la définition d'un critère de choix pour la taille de la fenêtre. Le cas particulier de la distance de Hausdorff qui a servi de support dans cette étude laisse alors place à une généralisation avec une mesure locale quelconque vérifiant uniquement les propriétés d'identité, de croissance et de majoration.

Dans cette étude, la fenêtre avait une position fixe, la section suivante poursuit l'étude lorsque les mesures sont faites en plusieurs positions.

3

Carte des Dissimilarités Locales

3.1 Définition générale

La carte des dissimilarités locales regroupe l'ensemble des mesures de dissimilarités locales faites en différents endroits des images. Ceci permet d'avoir accès à leur répartition spatiale.

Définition 3.1.1 (Carte des dissimilarités locales (CDL)). *Soit a et b deux images binaires (ou A et B deux ensembles finis de points de \mathbb{R}^2), et Φ^Δ une mesure de dissimilarité dans une fenêtre vérifiant les hypothèses de la proposition 2.4.4, la carte des dissimilarités locales CDL est définie par*

$$CDL \begin{cases} \mathfrak{B} \times \mathfrak{B} \times \mathcal{F} \longrightarrow \mathbb{R}^+ \\ (a, b, F) \longmapsto \delta_{a,b}^\Delta(F) \end{cases} \quad (3.1)$$

Cette définition est valable pour toute fenêtre qui est incluse dans X . D'après l'algorithme qui fixe la taille de la fenêtre. La fenêtre a une taille initiale et est agrandie jusqu'à atteindre la taille permettant la mesure de la dissimilarité. Dans le cas présent, aucune hypothèse n'a été faite sur la taille initiale de la fenêtre. Si la taille initiale est supérieure à celle permettant la mesure, alors la fenêtre n'est pas agrandie car le critère d'arrêt est vérifié. La valeur attribuée n'est alors pas forcément le reflet d'une mesure locale.

Cependant, il est possible de restreindre cette application à un sous-ensemble \mathcal{F}_{init} de \mathcal{F} , constitué de fenêtres initiales, dans le sens où elles sont petites par rapport à X (ou aux dissimilarités à évaluer) de même taille et uniformément réparties sur X . Ainsi, pour une fenêtre initiale F_{init} la CDL lui associe la valeur de la dissimilarité locale. Le cas particulier où l'ensemble des fenêtres initiales \mathcal{F}_{init} est constitué de l'ensemble des pixels, $\mathcal{F}_{init} = \{\{x\}, x \in X\}$, sera souvent utilisé par nous, il donne :

$$CDL \begin{cases} \mathfrak{B} \times \mathfrak{B} \times X \longrightarrow \mathbb{R}^+ \\ (a, b, x) \longmapsto \delta_{a,b}^\Delta(x) \end{cases} \quad (3.2)$$

La CDL peut alors être représentée sur le même support X que les images binaires a et b .

3.2 Mise en œuvre

3.2.1 Algorithme général

L'algorithme pour calculer la CDL est proposé ci-dessous avec une taille de fenêtre initiale $k \times k$. Cet algorithme est général et est valable dans tous les cas de figure. Des algorithmes plus simples sont proposés ci-dessous dans le cas de la DH. Dans cet algorithme, α est un paramètre compris entre 0 et 1 qui peut éventuellement intervenir. Il n'est pas utile dans le cas

Algorithm 3 Mise en œuvre de la CDL

```

for all pixel  $x$  do
   $n := k$  {initialisation de la côté de la fenêtre}
  while  $\Phi(a, b, F_n) \geq \alpha \max_{(a,b) \in \mathfrak{B}^2}(\Phi(a, b, F_n))$  et  $n \leq m$  do
     $n := n + 1$ 
  end while
   $CDL(x) := \Phi(a, b, F_{n-1})$ 
end for

```

où la CDL est basée sur la distance de Hausdorff mais il est présent dans le cas où la CDL est basée sur la différence simple.

3.2.2 Complexité du calcul

Considérons deux images $m \times m$. Le calcul de la CDL est composée d'une boucle qui est effectuée pour calculer en chaque point la mesure locale de dissimilarité (alg. 3) :

- pour chaque pixel, la taille de la fenêtre W est incrémentée de 1 depuis la taille initiale jusqu'à la taille finale $r_{max} \times r_{max}$ avec $r_{max} \leq m$ où m est le côté du support X .
- Le calcul est effectué pour les $m \times m$ pixels.

À chaque étape, le calcul est fait uniquement pour les nouveaux pixels situés sur le bord de la fenêtre. En conséquence, le calcul à partir d'une fenêtre initiale est fonction de l'aire de la fenêtre $F_{max} = r_{max}^2$ qui est inférieure à m^2 .

Ainsi pour chaque pixel, il y a $O(m^2)$ opérations. Comme il y a m^2 pixels, la complexité pour la CDL est $O(m^4)$.

remarque 3.2.1. r_{max} est une variable qui est majorée par m , mais la plupart du temps ce majorant n'est pas atteint pour une raison simple : pour les points de fond communs aux deux images, la fenêtre n'est pas agrandie.

Ce majorant est mauvais car très général, dans les cas particuliers (notamment pour la DH) il est meilleur.

3.3 Cas de la Distance de Hausdorff

Dans le cas de la distance de Hausdorff, le calcul de la CDL (notée CDL_{DH} en cas d'ambiguïté) est bien simplifié car il revient essentiellement à calculer les transformées en distance. Cette simplicité permet un calcul rapide et de plus, elle permet une généralisation élégante aux images en niveaux de gris. Enfin, malgré sa simplicité de calcul, la CDL_{DH} offre une très bonne efficacité dans les applications.

3.3.1 Définition

Définition 3.3.1 (Carte des dissimilarités locales (CDL)). Soit A et B ensembles finis de points de \mathbb{R}^2 , la carte des dissimilarités locales CDL est définie par

$$\forall x \in \mathbb{R}^2, CDL(x) = \begin{cases} DH_{B(x, r_{max})}(A, B) & \text{si } \mathcal{R} \neq \phi \\ 0 & \text{si } \mathcal{R} = \phi. \end{cases} \quad (3.3)$$

Théorème 3.3.1 (expression mathématique de la CDL pour la DH).

$$\forall x \in \mathbb{R}^2, CDL(x) = \max(d(x, A), d(x, B)) \mathbb{1}_{A\Delta B}(x) \quad (3.4)$$

Démonstration. Distinguons deux cas :

$x \notin A\Delta B$ alors $a(x) = b(x)$ et $CDL(x) = 0$, or $\mathbb{1}_{A\Delta B}(x) = 0$ donc la formule est vérifiée.

$x \in A\Delta B$ dans ce cas l'ensemble de mesure locale n'est pas vide et par la propriété 2.4.2 (p. 72), on a donc $CDL(x) = \max(d(x, A), d(x, B))$. Or $\mathbb{1}_{A\Delta B}(x) = 1$ donc la formule est vérifiée. □

Ce théorème est central parce qu'il provient de la définition de dissimilarité locale : la valeur obtenue représente la mesure de la dissimilarité locale mesurée dans une fenêtre de taille adéquate. Et il donne une expression mathématique simple de la CDL. Cette expression est une fonction mathématique qui est valable pour deux ensembles compacts de points A et B quelconque du plan. Mais l'expression est aussi directement utilisable pour les images binaires, comme cela est présenté dans le paragraphe suivant.

Les valeurs contenues dans la CDL sont majorées par la valeur de la mesure globale par construction. La proposition suivante énonce que la valeur de la mesure globale (*i.e.* la DH entre les deux images) est atteinte dans la CDL_{DH} .

Proposition 3.3.1 (Valeur maximale dans la CDL). La valeur $v = DH(A, B)$ est atteinte au moins une fois dans la CDL :

$$\max(CDL(A, B)) = DH(A, B). \quad (3.5)$$

Démonstration. $DH(A, B) = \max(h(A, B), h(B, A))$, donc le maximum est atteint pour l'une des deux distances directes, par exemple pour $h(A, B)$:

$$h(A, B) = v.$$

par définition,

$$\begin{aligned} h(A, B) &= \max_{a \in A} \left(\min_{b \in B} d(a, b) \right), \\ \exists a_0 \in A, h(A, B) &= \min_{b \in B} d(a_0, b) \\ \exists b_0 \in B, h(A, B) &= d(a_0, b_0). \end{aligned} \quad (3.6)$$

Montrons maintenant que pour $x = a_0$, $CDL(x) = v$.

Nous savons que

$$\forall b \in B \cap B(a_0, v), d(a_0, b) = v,$$

en effet, par 3.6 nous avons

$$\min_{b \in B} d(a_0, b) = d(a_0, b_0).$$

D'autre part a_0 est au centre de $B(a_0, v)$, et enfin, $b_0 \in B \cap B(a_0, v)$, donc le lemme 2.4.1 implique que pour a_0 , $r_{max} = d(a_0, b_0) = v$. D'où le résultat. \square

La définition de la CDL_{DH} fait intervenir la distance d'un point à un ensemble $x \mapsto d(x, A)$. Cette fonction a été étudiée sous le nom de transformation en distance. Elle est présentée dans la sous-section suivante.

3.3.2 Transformation en distance (TeD)

Définition

La Transformation en Distance (TeD, Distance Transform ou DT en Anglais) joue un rôle important en vision par ordinateur, en reconnaissance de forme ou en robotique. Elle représente l'information de la distance à l'ensemble des pixels noirs dans une image binaire :

Définition 3.3.2 (transformée en distance). *Soit A un ensemble de points de \mathbb{R}^2 la transformée en distance de A est donnée par*

$$TeD_A(x) = d(x, A) \text{ pour } x \in \mathbb{R}^2 \quad (3.7)$$

En robotique, par exemple, si les pixels noirs représentent un obstacle, la TeD donne l'information de la distance qui sépare le robot de l'objet. Ce qui est utile lorsque l'on essaie de déplacer l'objet en évitant les obstacles. D'autre part, elle est une étape importante dans le calcul de la distance de Hausdorff. La plupart des études qui lui sont dédiées ont pour but d'améliorer son temps de calcul. Nous en faisons ici une présentation.

État de l'art

Elle a été introduite par Rosenfeld [Rosenfeld and Pfalz, 1966; Rosenfeld and Kak, 1982], puis dans [Borgefors, 1986b], l'auteur présente une méthode de calcul rapide de la TeD pour la distance euclidienne en utilisant des masques. Plus récemment, plusieurs méthodes ont été présentées pour le calcul rapide et exact de la TeD euclidienne par exemple [Huang and Mitchell, 1994] qui utilise une technique reposant sur des opérateurs morphologiques en nuances de gris, et Breu [Breu et al., 1995] qui utilise le diagramme de Voronoï pour accéder à un temps de calcul de la TeD linéaire. D'autre part, des études ont été faites aussi sur la distance L_1 [Kovács and Guerrieri, 1992]. Brown présente un algorithme rapide pour obtenir la TeD pour la distance L_∞ [Brown, 1994]. Enfin, dans [Arlandis and Perez-Cortes, 2000], les auteurs présentent une généralisation de la TeD aux images en nuances de gris sur laquelle repose une extension de la CDL proposée § 3.7.2.

Propriétés

Le calcul de la TeD dépend bien sûr de la distance d choisie, prenons $x = (x_1, x_2)$, $y = (y_1, y_2)$, les choix classiques pour d sont

- la distance euclidienne issue de la norme L_2 ,

$$d_2(x, y) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2},$$

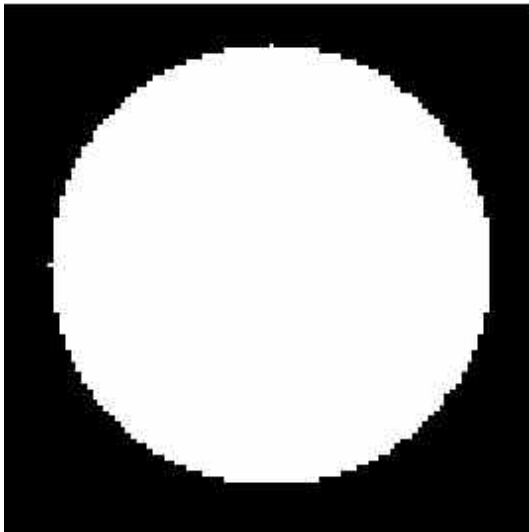
- la distance issue de la norme L_1 , qui produit le 4-voisinage,

$$d_1(x, y) = |y_1 - x_1| + |y_2 - x_2|,$$

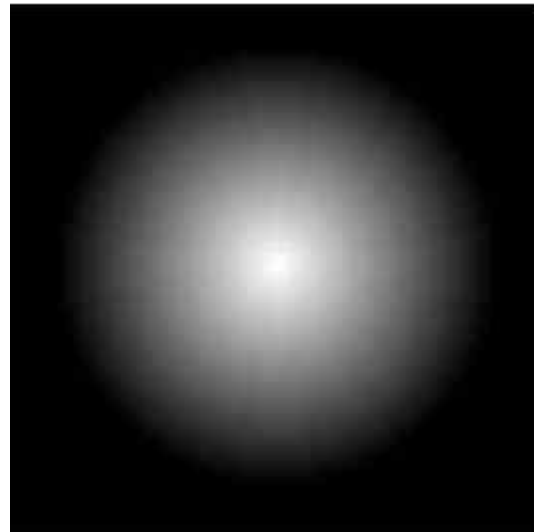
- la distance d_∞ issue de la norme L_∞ , qui produit le 8-voisinage,

$$d_\infty(x, y) = \max(|y_1 - x_1|, |y_2 - x_2|).$$

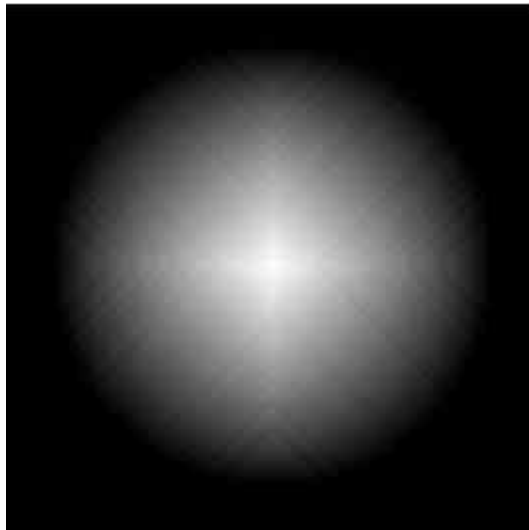
L'influence du choix de la distance est illustrée figure 3.1.



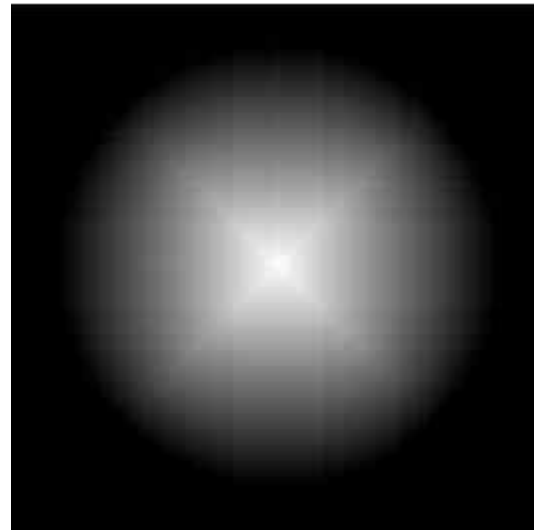
(a) Image binaire originale.



(b) *TeD* utilisant la distance euclidienne.



(c) *TeD* basée sur la norme L_1 .



(d) *TeD* issue de la norme L_∞ .

FIG. 3.1 – Une image binaire et ses trois *TeD* faites avec les trois distances classiques. La forme des voisinages se retrouve dans la *TeD* : rond pour la distance euclidienne, losange pour L_1 , et carré pour L_∞ .

3.4 Mise en œuvre de la Carte des Dissimilarités Locales basée sur la distance de Hausdorff

Cette partie concerne l'aspect pratique du calcul de la CDL. Il est simple dans le cas de la CDL_{DH} et un peu plus lourd dans le cas général.

3.4.1 Calcul de la distance de Hausdorff globale

À partir de la TeD de chacun des deux ensembles, il est possible de calculer la DH :

Proposition 3.4.1 (Des TeD à la DH). *Soient deux ensembles non vides de points $A = (a_1, \dots, a_n)$ et $B = (b_1, \dots, b_m)$ de E , muni d'une distance d , La DH entre A et B est donnée par*

$$DH(A, B) = \max \left(\max_{a \in A} (TeD_B(a)), \max_{b \in B} (TeD_A(b)) \right). \quad (3.8)$$

Une fois la TeD calculée, le calcul de la DH revient à un simple calcul de maximum. L'avantage est double :

- c'est la même TeD qui sert pour calculer le maximum en chaque point de l'ensemble,
- alors que les algorithmes précédents amenaient à refaire tout le calcul pour chaque comparaison d'images, s'il y a plusieurs images à comparer, le calcul des TeD est fait une seule fois et chaque TeD sert pour toutes les comparaisons entre son image et les autres.

3.4.2 Calcul de la Carte des Dissimilarités Locales à partir du théorème 3.3.1

Expression

Dans le cadre de la DH, le théorème 3.3.1 produit avec la transformation en distance une mise en œuvre très simple de la CDL : pour un x donné, le théorème 3.3.1 fournit l'expression suivante

$$CDL(x) = \max(d(x, A), d(x, B)) \mathbb{1}_{A \triangle B}(x) \quad (3.9)$$

or avec l'expression de la TeD (définition 3.7), cela devient

$$CDL(x) = \max(TeD_A(x), TeD_B(x)) \mathbb{1}_{A \triangle B}(x) \quad (3.10)$$

Cette expression est directement calculable pour peu que le calcul de la TeD soit acquis. Elle n'utilise pas de fenêtre pour accéder à l'information de dissimilarité locale, même si la méthode qui a permis d'y accéder repose sur la mesure dans une fenêtre. Cela vient de la définition de la DH.

Complexité

Le calcul de la TeD est linéaire, comme il y a m^2 pixels, la complexité du calcul de la TeD est en $O(m^2)$. Le calcul du maximum et la multiplication par l'indicatrice sont aussi de complexité $O(m^2)$, donc la complexité du calcul de la CDL basée sur la distance de Hausdorff est un $O(m^2)$. Cet ordre de grandeur est à comparer à celui de l'algorithme général qui est un $O(m^4)$. Le calcul dans le cas de la CDL basée sur la distance de Hausdorff est beaucoup rapide, ce qui est vérifié dans la pratique.

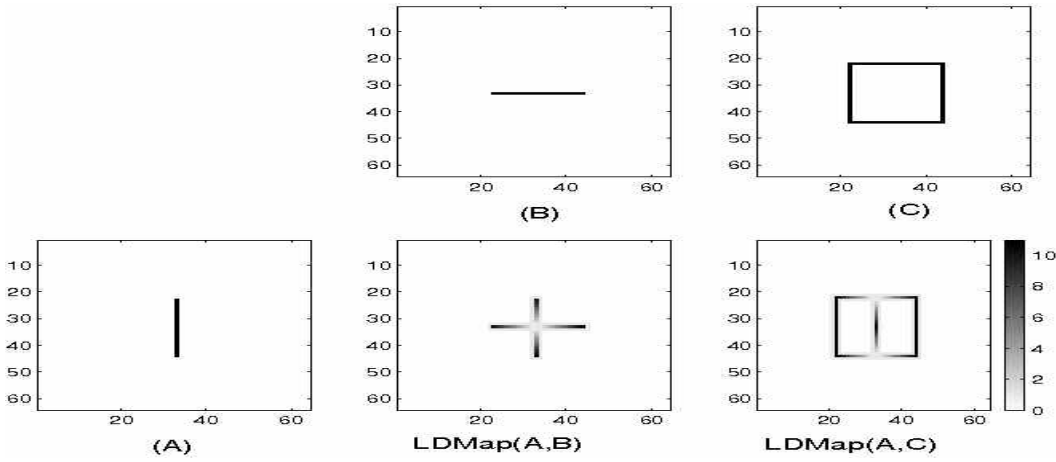


FIG. 3.2 – Trois formes simples et leurs CDL. Une ligne verticale, une horizontale et un carré et leurs CDL. Plus le pixels est foncé, plus grande est la valeur de la distance. La valeur de la DH globale est 11 pour les deux comparaisons présentées

3.5 Résultats qualitatifs

Des images et leurs CDL sont présentées dans ce paragraphe afin de mettre en valeur les propriétés de la CDL.

3.5.1 Lignes et carré

La figure 3.2 contient des formes simples (une ligne verticale, une horizontale et un carré) et les CDL correspondantes. L'échelle des valeurs pour les CDL va de blanc pour la valeur 0 à noir pour la plus haute valeur.

Commentaires

Pour la comparaison entre la ligne verticale et celle horizontale, le pixel situé à leur croisement est commun aux deux lignes, il a donc pour valeur 0 dans la CDL. Pour les pixels contenus dans une des lignes, plus ils sont éloignés de l'autre ligne, plus la distance qui leur est attribuée dans la CDL est grande, et donc, plus ils sont foncés. La valeur de la DH globale est 11 pour les deux comparaisons présentées. la comparaison entre la ligne verticale et le carré donne la même valeur pour la DH globale, mais l'organisation spatiale de la CDL montre que cette valeur est atteinte pour de nombreux pixels (appartenant aux côtés verticaux du carré) alors qu'elle n'est atteinte que quatre fois pour la comparaison des deux lignes (en leurs extrémités).

3.5.2 Lettres « co » et « et »

La figure 3.3 illustre la notion de dissimilarité locale. Chacune des deux images contient deux lettres. La première lettre de chaque image (« c » et « e ») ressemble à l'autre, et les grandes valeurs de distance (en noir) sont situées à l'endroit du trait droit du « e », trait que ne possède pas le « c ». La comparaison entre le « o » et le « t » illustre le même comportement.

3.5.3 Jeu des dix erreurs

La figure 3.4 propose un moyen sophistiqué de gagner du temps durant les vacances. En effet, les images proviennent d'un jeu des dix erreurs proposé dans un journal d'été : la seconde image est une copie de la première, mais avec dix différences

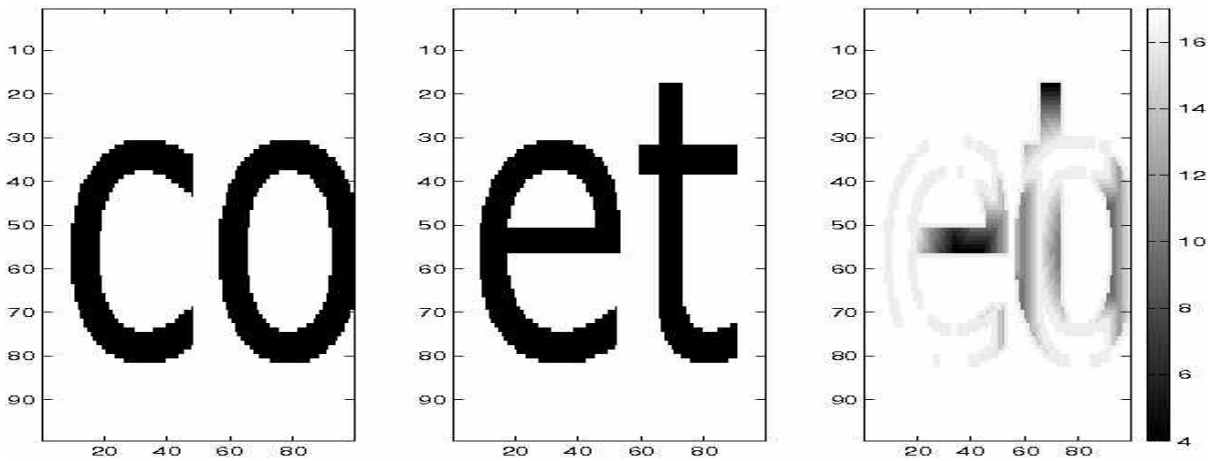


FIG. 3.3 – Les lettres CO et ET et leur CDL illustrant leurs dissimilarités locales.

(les dix erreurs). Elles ont été acquises de deux points de vue légèrement différents et une comparaison directe $C = |B - A|$ ne permet pas de retrouver les dix erreurs (image C). La CDL (image D) permet de visualiser la plupart des erreurs (les dix erreurs ont été entourées par nos soins), à l'exception de la bande noire sur la pantalon de survêtement (qui se transforme en deux bandes noires dans la deuxième image).

La raison de la faible valeur de la distance pour cette erreur est son anisotropie : l'erreur a lieu sur une bonne longueur mais la largeur des bandes est faible. Lorsqu'elle est sur les pixels d'une des bandes, la fenêtre arrête sa croissance aussitôt qu'elle rencontre les pixels d'une bande de l'autre image et de ce fait, elle est aveugle à la longueur de l'erreur.

3.6 Comparaison aux DH modifiées

La CDL permet d'avoir accès aux mesures locales des dissimilarités. Dans le cas de la DH, elle donne accès au nombre de fois où les valeurs sont atteintes. Ces informations qui sont difficiles à prendre en compte avec une seule valeur interviennent par exemple dans la DH modifiée où les valeurs sont moyennées. Cependant, la valeur de la DH modifiée ne permet pas de savoir comment sont réparties les valeurs autour de la moyenne et comment elles sont réparties spatialement. Les figures 3.5 et 3.6 donnent ici des exemples où les valeurs de la DH et des versions modifiées qui ont été proposées sont estimées.

3.7 Généralisation aux images en niveaux de gris

La CDL est destinée à mesurer les dissimilarités locales entre des images binaires que l'on peut assimiler à des ensembles de points dans le plan. Comme les images en niveaux de gris sont nombreuses, il est souhaitable de pouvoir calculer une CDL entre images en niveaux de gris. Cependant, elles ne peuvent être assimilées à un ensemble de points, la généralisation n'est donc pas immédiate. Il y a au moins deux possibilités pour le faire qui sont présentées dans cette section : en considérant les images en niveaux de gris comme des surfaces (de points) ou en utilisant l'expression de la CDL_{DH} avec la transformation en distance.

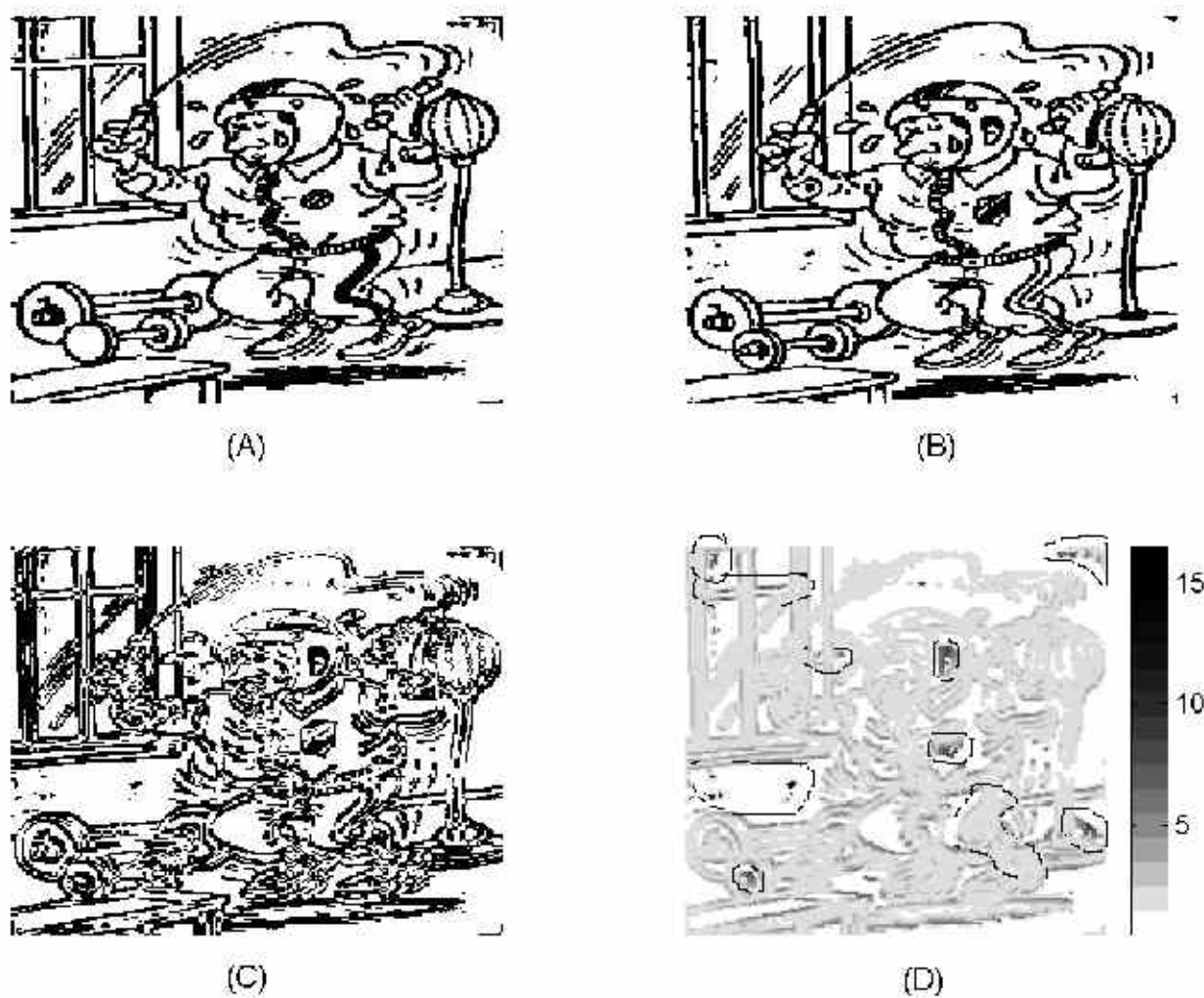
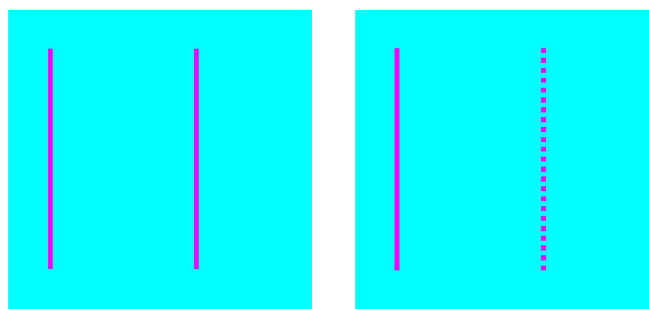
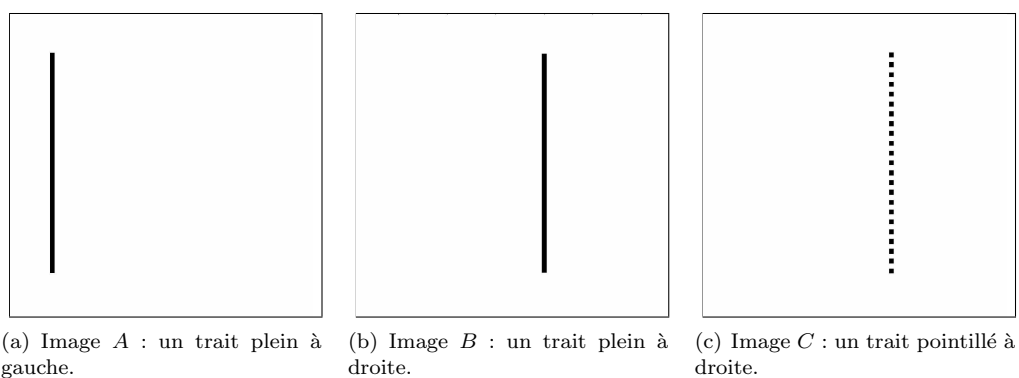


FIG. 3.4 – Le jeu des dix erreurs. Images *A* et *B* : les deux images à comparer, image *C* : la différence absolue $C = |B - A|$ et image *D* : leur CDL où les dix erreurs ont été entourées de noir par nos soins.



(d) CDL entre l'image A et l'image B . Il y a deux valeurs de distance dans la CDL : 0 (en bleu) et 12 (en rose) qui est la valeur de la DH globale. (e) CDL entre l'image A et l'image C . Les valeurs de distance sont représentées par les mêmes couleurs que dans la CDL 3.5(d).

FIG. 3.5 – La dissimilarité entre A et B est différente de celle entre A et C . Cependant, la mesure de la DH globale est la même et celles des DH modifiées sont identiques ou proches. Les CDL permettent d'accéder aux différences à la fois dans la répartition spatiale des dissimilarités et dans la répartition de leurs valeurs.

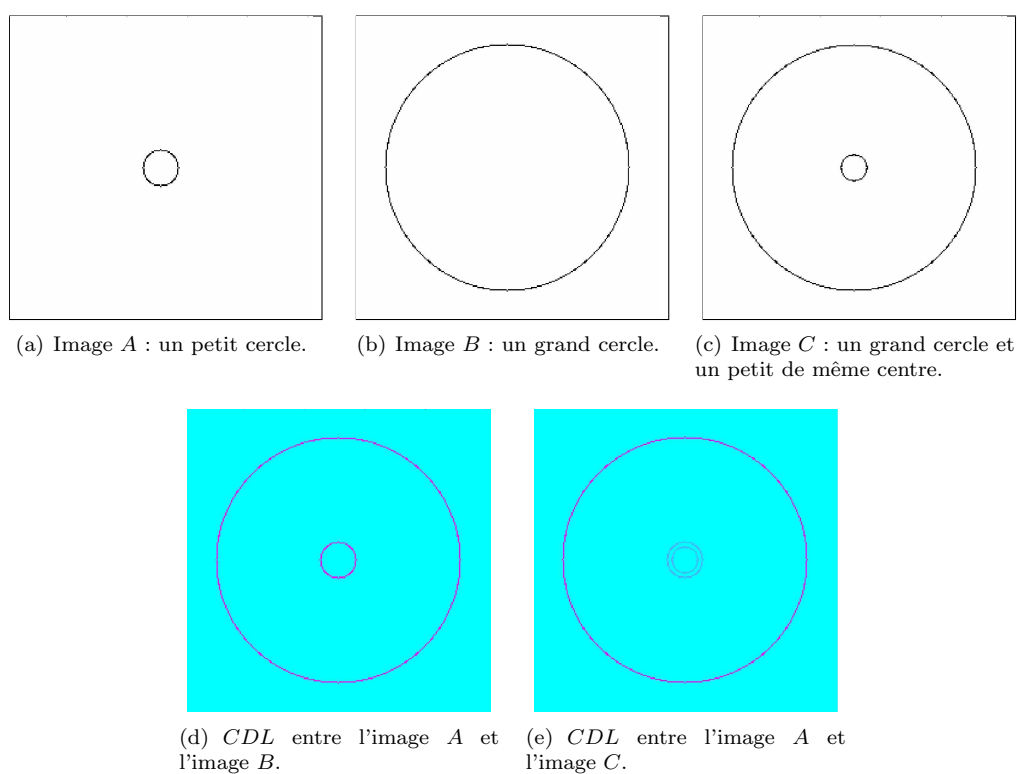


FIG. 3.6 – La dissimilarité entre *A* et *B* est différente de celle entre *A* et *C* (où les deux cercle centraux sont proches). Cependant, la mesure de la DH globale est la même et celles des DH modifiées sont identiques ou proches. Les CDL permettent d'accéder aux différences à la fois dans la répartition spatiale des dissimilarités et dans la répartition de leurs valeurs.

3.7.1 Généralisation à partir de la définition générale de la Carte des dissimilarités locales

La méthode générale qui permet d'obtenir une CDL repose sur une mesure de dissimilarités Δ à partir de laquelle est définie une mesure de dissimilarités par une fenêtre Φ^Δ . Si cette dernière vérifie les hypothèses de la proposition 2.4.4, alors la mesure locale de dissimilarités δ^Δ peut être définie. La généralisation se fait alors simplement à partir d'une mesure de dissimilarités Δ_g entre images en niveaux de gris : il suffit qu'elle vérifie les hypothèses de la proposition 2.4.4, et la mesure locale de dissimilarités δ^{Δ_g} peut être définie.

Exemple applicatif

Cet exemple est la généralisation de celui présenté en 2.4.3. Soit \mathfrak{G} l'ensemble des images en niveaux de gris qui vont du support X dans $\{0, N - 1\}$, Δ_g est définie de la manière suivante :

$$\forall (f, g) \in \mathfrak{G}^2, \Delta_g(f, g) = \sum_{x \in X} |f(x) - g(x)|. \quad (3.11)$$

Alors, la mesure de dissimilarité dans une fenêtre est donnée par :

$$\forall (f, g) \in \mathfrak{G}^2, \forall F \in \mathcal{F}, \Phi_{DS}(f, g, F) = \sum_{x \in F} |f(x) - g(x)|. \quad (3.12)$$

Proposition 3.7.1. *La mesure de la différence simple par une fenêtre Φ_{DS} vérifie les trois propriétés de la proposition 2.4.4.*

Démonstration. La preuve est identique à celle de la proposition 2.4.5. □

La mesure locale de dissimilarité δ_{DS} ainsi définie souffre cependant du même défaut que la mesure de la différence simple Δ dont elle provient : elle mesure la différence pixel à pixel et va mesurer de fortes valeurs là où il y a une légère translation. Ce qui est atténué par la mesure locale.

3.7.2 Généralisation dans le cas de la distance de Hausdorff

Généralisation par la méthode de la fenêtre croissante

Pour les images binaires, la méthode de la fenêtre croissante (en deux dimension) pour la DH repose sur le principe suivant : à partir d'une fenêtre initiale centrée en un point d'une image, la fenêtre est agrandie jusqu'à rencontrer un point de l'autre image. La notion de point n'a plus de sens lorsque les images sont en niveaux de gris. Cependant, il est possible de considérer une image en niveaux de gris comme une surface dans un espace en trois dimensions : à chaque point x du support, la position de la surface est indiquée par la valeur du niveau gris de l'image au point x . Une voie possible de généralisation de la méthode de la fenêtre croissante est de partir d'une boîte centrée en un point d'une des surfaces et de l'agrandir jusqu'à rencontrer un point de l'autre surface. Le « rayon » de la boîte donne alors la valeur de la distance directe.

Cette méthode, même si elle est lourde à mettre en œuvre, a le mérite de généraliser de manière simple la méthode dédiée aux images binaires. Cependant, elle présente un inconvénient majeur parce qu'elle met sur le même plan la distance spatiale (celle en 2 dimension entre les pixels) et la différence de niveau de gris. Or ces deux mesures n'ont en général pas la même signification : hormis dans le cas des images de profondeur, le niveau de gris ne retranscrit pas une distance. De plus, les niveaux de gris ont une échelle qui dépend du type d'image (en général entre 0 et 255) et la mesure de leur différence dépend de cette échelle. Ainsi, cette méthode présente une faiblesse puisqu'elle met sur le même plan distance spatiale et différence de niveaux de gris.

Généralisation par la méthode de la Transformée en Distance

Il est possible de contourner cette difficulté en utilisant la transformée en distance. En effet, la transformée en distance a été généralisée aux images en niveaux de gris par [Arlandis and Perez-Cortes, 2000]. Or, dans le cas où la mesure locale est la DH, l'expression de la CDL (notée CDL_{DH}) est donnée dans le théorème 3.3.1 et fait intervenir les transformées en distance des images et le OU exclusif Δ . La généralisation de $A\Delta B$ se fait par $|A - B|$ car on souhaite conserver la valeur zéro aux points où les deux images A et B ont la même valeur. L'expression de la CDL est donnée dans la proposition suivante.

Proposition 3.7.2 (généralisation de la CDL_{DH} par la TeD aux images en niveaux de gris). *Dans le cas de la CDL_{DH} , la généralisation aux images en niveaux de gris se fait par l'expression suivante :*

$$CDL_{DH}(A, B) = |B - A| \max(TeD_A, TeD_B) \quad (3.13)$$

L'expression 3.13 permet de calculer une CDL entre images en niveaux de gris. Comme la mesure des dissimilarités est locale, cette CDL n'a de sens que pour des images proches. Une illustration est donnée sur une image en niveaux de gris tirée de la base de visages ATT figures 3.7. La comparaison est faite avec la différence simple entre les deux images.

3.8 Étude sur la fenêtre glissante W

3.8.1 Utilité de l'adaptabilité de la fenêtre W

Le choix que nous avons fait d'adapter localement la taille a été justifié par la volonté de mesurer la dissimilarité locale. Il est néanmoins possible de prendre une fenêtre de taille fixe (c'est ce qui a été fait chronologiquement en premier). Cependant, comme nous allons le voir, une taille fixe de fenêtre empêche d'avoir une bonne précision sur toutes les échelles. L'exemple porte sur une comparaison entre une forme de lapin et une de requin (figure 3.8(a) et 3.8(b)), leur DH globale vaut $DH(a, b) = 12$ (pour pouvoir l'appréhender, il faut une fenêtre de taille supérieure à 24×24). Leur CDL est faite avec une fenêtre de taille fixe pour trois tailles : 5×5 , 10×10 et 25×25 et la fenêtre localement adaptative :

fig. 3.8(c) taille 5×5 Cette CDL permet de bien distinguer les détails fins, notamment les contours, mais par contre la taille de la fenêtre limite à 5 les valeurs possibles dans la carte. L'information de dissimilarité n'est pas entièrement disponible.

fig. 3.8(d) taille 10×10 L'information sur les dissimilarités est plus importante, mais c'est au détriment de la précision sur les dissimilarités fines.

fig. 3.8(e) taille 25×25 Ici, la taille de la fenêtre permet d'accéder aux plus grandes valeurs possibles dans la CDL (car la taille de la fenêtre est le double de la DH globale) par contre, comme la fenêtre est partout de grande taille, elle ne mesure pas que des dissimilarités locales, ce qui rend la carte très peu lisible : les valeurs importantes sont présentes, mais pas localisées et par exemple, le contour du lapin et celui du requin ne sont pas visibles.

fig. 3.8(f) taille adaptative Cette DL permet d'accéder à la fois à toute l'échelle des valeurs, et au détail des dissimilarités (contours et localisation des pointes de valeurs..)

3.8.2 Forme de la fenêtre

Comme l'exemple sur le jeu des dix erreurs 3.4 (p. 85) le montre, l'erreur portant sur la bande le long de la jambe du pantalon n'a pas été détectée car, même si elle porte sur une longueur importante, elle est fine. Mais dans le calcul de la

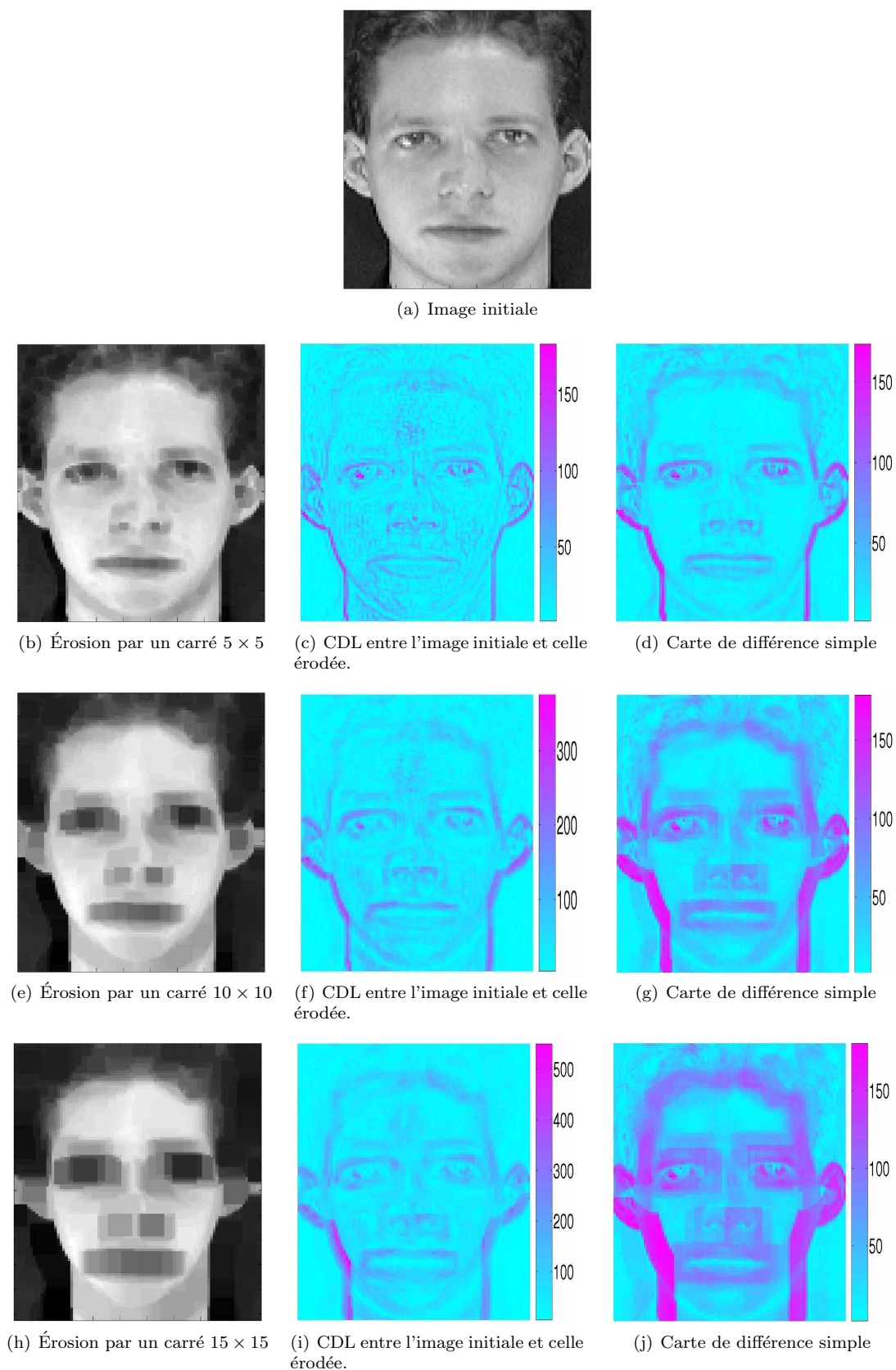


FIG. 3.7 – Un image de visage $x \times y$ (provenant de la base ATT) qui a été érodée par un ES carré de taille variable et les CDL correspondantes formées à partir du théorème 3.7.2. La comparaison est faite avec la carte de différence simple entre l'image initiale et les images filtrées. Pour un élément structurant assez gros, les bords du visage et les yeux sont bien érodés. Par contre, la CDL dont la valeur augmente avec la taille de l'élément structurant n'est pas autant affecté par la dégradation de l'image et elle ne fait pas apparaître les carrés résultant de l'érosion, ce qui est notable.

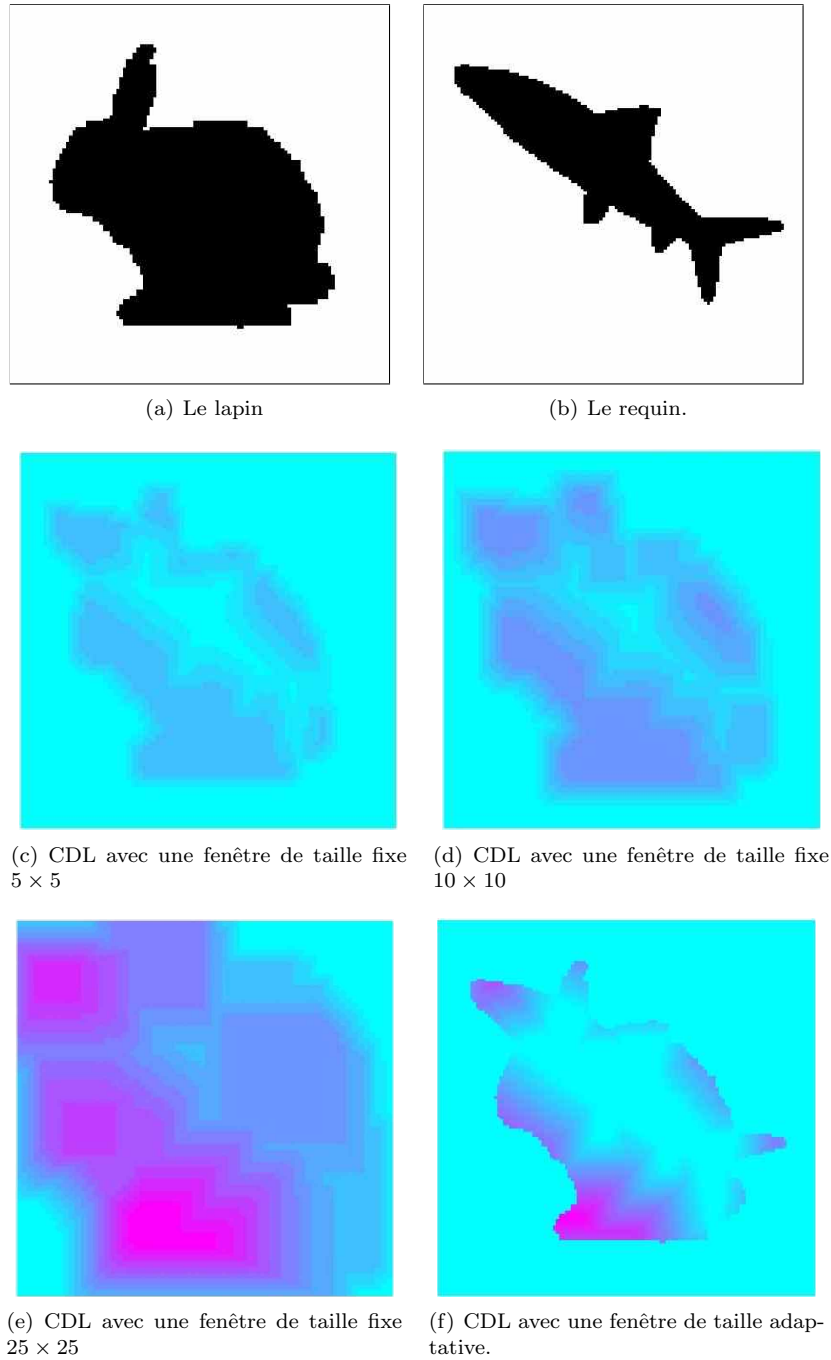


FIG. 3.8 – Deux images binaires dont la DH globale est de 25 et leurs quatre CDL. Les 3 premières sont faites avec une fenêtre de taille fixe 5×5 , 10×10 , 25×25 (DH globale=12) et la dernière avec une fenêtre localement adaptative.

CDL, les fenêtres sont carrées et isotropes. Ce qui fait que les traits les plus proches entrent en premier dans la mesure. Il est possible de jouer sur la forme de la fenêtre pour détecter les dissimilarités anisotropes, il suffit pour cela de prendre une fenêtre allongée dans une direction pour favoriser la détection dans cette direction. Comme la mesure de la distance entre les points entre en compte il est nécessaire que la fenêtre soit convexe. Nous présentons dans la figure 3.9 des exemples avec

des fenêtres orientées dans les directions horizontales et verticales. Comme la fenêtre augmente plus vite dans une direction, les dissimilarités dans cette direction sont mesurées même si les points sont plus éloignés que dans l'autre direction (dans la proportion du rapport hauteur largeur). Ce qui permet de favoriser la mesure dans une direction. Cette mesure, qui est facile à mettre en œuvre, ne nous a pas servi dans la pratique car nous avons eu affaire à des images globalement isotropes. Dans le cadre d'images ou de dissimilarités anisotropes, cette mesure peut être intéressante, même si les axes de la mesure anisotropique de la CDL ne peuvent pas avoir toutes les directions (principalement directions horizontale et verticale). Une autre possibilité est de faire la CDL entre les images de détails d'une analyse multirésolution des images. Une difficulté dans ce cas vient du fait que ce qui est représenté dans les images de détail n'est pas bien contrôlé.

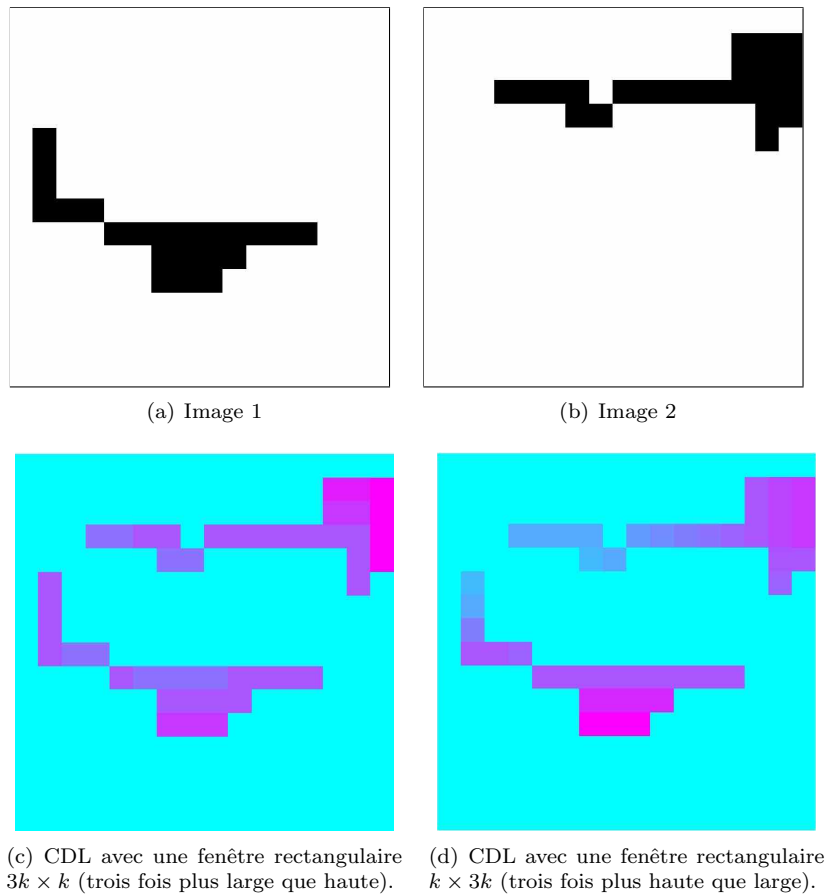


FIG. 3.9 – Deux images binaires et leur CDL avec des fenêtres de mesure locale asymétriques. Le rapport largeur hauteur est de 3 ou de $1/3$. Cela permet de valoriser les dissimilarités verticales ou horizontales.

3.9 Bilan

Dans cette section, la CDL a été définie de manière générale à partir d'une mesure locale de dissimilarité. Un algorithme a ensuite été développé pour le calcul de la CDL dans le cas général. Compte tenu de cet aspect général, la complexité a été majorée par l'ordre de grandeur $O(m^4)$. Dans le cas où la mesure locale repose sur la DH, la CDL a une expression simple reposant sur la transformation en distance. Sa complexité est alors un $O(m^2)$.

Des résultats qualitatifs ont été présentés sur des images simples pour mettre en avant les propriétés de la CDL concernant entre autre la distribution spatiale des mesures de dissimilarités. Une comparaison qualitative a été faite entre la CDL et les DH modifiées présentant l'apport d'information de la CDL.

Ensuite, deux extensions de la CDL ont été présentées, la première concernant la généralisation aux images en niveaux de gris et la deuxième s'intéressant à la forme de la fenêtre de mesure locale.

Ces définitions permettent maintenant d'aborder l'utilisation de la CDL dans les applications. Les applications en classification d'images montrent que l'apport d'information de la CDL permet d'obtenir de meilleurs résultats.

Troisième partie

Application à la classification

La CDL permet de mesurer et de représenter les dissimilarités entre deux images binaires. Nous présentons dans cette partie comment la CDL peut être exploitée. Tout d'abord dans la reconnaissance d'images par le contenu dans la section 0.1 puis en classification d'impressions anciennes (chap 1) et d'autres bases (chap. 3).

0.1 La Carte des Dissimilarités Locales au sein d'un processus global de comparaison

0.1.1 Généralités

La CDL représente les dissimilarités entre deux images binaires, elle peut donc être utilisée pour savoir si deux images sont similaires. Cependant, comme elle se présente sous la forme d'une image, la méthode pour la classer doit être assez élaborée.

Il y a deux optiques en classification : soit on suppose qu'il y a un exemple représentatif dans chaque classe appelé prototype, soit la classe est définie par un ensemble de règles ou une grammaire. Nous nous plaçons dans la première optique et nous supposons donc que chaque classe peut être représentée par un prototype.

Dans ce cadre, le premier point à souligner est que déterminer la similarité entre les images revient à classer les CDL. En effet, lorsqu'une image est présentée pour être classée, elle est comparée aux autres images (ou seulement aux prototypes s'il y a ambiguïté entre les classes) et le résultat de cette comparaison ne peut être que « similaire » ou « dissimilaire ». Les CDL peuvent donc être classées de la manière suivante : celles comparant des images similaires qui composent la classe C_{sim} et celles comparant des images dissimilaires qui forment la classe C_{dissim} . Lorsque deux images sont comparées, pour savoir si elles sont similaires, il suffit alors de savoir classer leur CDL dans C_{sim} ou C_{dissim} .

Le processus de calcul des CDL et de leur classification peut être assez long. Il n'est donc pas utile de mesurer la similarité des images à un niveau de finesse de détail qui n'entrera pas en compte dans le critère de similarité. En effet, seule l'information concernant les structures importantes rentre en jeu dans l'évaluation. L'échelle à laquelle la mesure de similarité s'effectue est déterminée par l'expert lors de la phase d'apprentissage. Pour choisir une échelle, il faut décomposer l'information contenue dans l'image en plusieurs échelles. Une AMR permet de le faire sans redondance. Le fait que les images traitées soient des images binaires oblige à faire une étude approfondie pour la formalisation et le choix de l'AMR (*cf* partie 1). Le résultat de cette étude est que l'AMR choisie est une AMR non-linéaire basée sur le filtre morphologique de la médiane qui permet une bonne conservation des structures principales de l'image .

Ainsi, il est possible de comparer les images à plusieurs échelles (au moyen de l'approximation ou des détails) et l'échelle retenue sera la plus grossière conservant une classification de qualité.

La sous-section suivante présente le schéma du processus de classification partant des images pour arriver à la décision sur leur similarité.

0.1.2 Schéma synoptique

Une fois les images prétraitées, elles sont binaires et recalées. Elles sont alors décomposées dans une AMR morphologique basée sur la médiane. Puis le calcul de la CDL est fait à partir de l'AMR. Il peut être fait à plusieurs échelles, sur les approximations ou sur les détails. Ensuite, la ou les CDL produites doivent être classées dans C_{sim} ou dans C_{dissim} . Pour cela, un module de classification supervisée est mis en place et un apprentissage est fait sur des échantillons de CDL (aux bonnes échelles) auparavant classées par un expert. La classification indique alors l'appartenance à l'un des deux groupes et de ce fait la similarité ou non des deux images. Le schéma de ce processus est donné fig. 0.10.

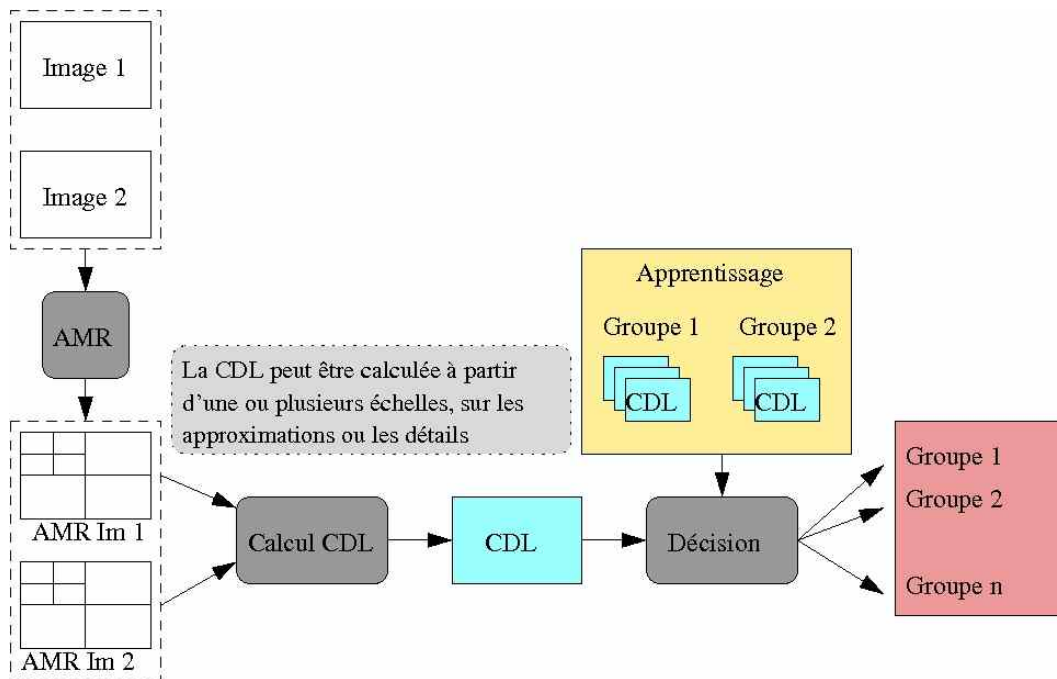


FIG. 0.10 – Schéma du processus global utilisant l'AMR, la CDL et le module de décision.

Le processus de classification a été appliqué à une base d'impressions anciennes numérisées. Cette application fait l'objet du chapitre suivant.

1

Classification d'impressions anciennes

1.1 Contexte

Les documents anciens sont en général des livres imprimés sur du papier qui se dégrade et dont la conservation est délicate. La numérisation permet le stockage sous format numérique des ouvrages anciens. Cela offre un support supplémentaire aux documents anciens. En premier lieu, ce format ne subit pas le même vieillissement que le papier, il conserve une qualité égale pour peu que le support informatique soit renouvelé lorsqu'il devient obsolète. Ce qui est très intéressant pour la conservation du patrimoine culturel. En deuxième lieu, il permet la consultation à distance. Les livres anciens sont généralement trop fragiles pour être transportés ou consultés en libre service. Alors qu'il fallait se rendre dans la bibliothèque dépositaire de l'œuvre imprimée, sa version numérique est consultable à distance. En troisième point, le format numérique autorise le stockage sous forme de base de données et offre alors la possibilité d'effectuer des recherches automatiques.

Ce dernier point intéresse particulièrement les chercheurs en histoire du livre du CEPLECA (Le Centre d'Étude du Patrimoine Linguistique et Ethnologique de Champagne-Ardenne). En effet, les illustrations contenues dans les livres anciens étaient imprimées avec des tampons en bois gravés. Ces tampons, après une période d'usage, étaient refaits (à la main) ou pouvaient être revendus entre maisons d'édition. Ils étaient alors réutilisés parfois pour illustrer des scènes complètement différentes. Retrouver ces différentes illustrations permet de mieux comprendre les relations tissées dans le monde de l'édition d'alors [Leclerc, 2000]. Cette piste n'a pas encore été complètement exploitée et est assez laborieuse à explorer. En effet, il n'est pas toujours possible de se référer au contexte des impressions puisqu'elles ont pu être détournées de leur usage initial. Dans ce cadre, la recherche automatique dans une base composée des impressions anciennes numérisées constitue une avancée certaine. Elle permet de rechercher à partir d'une impression celles qui dans la base sont proches.

1.2 Projet ANITA

Cependant le programme de numérisation des documents anciens est actuellement en cours. Les problèmes de traitement d'image spécifiques à ces bases de données naissantes demandent des solutions adaptées. C'est dans ce contexte qu'a été développé le projet ANITA (ANalyse d'Images et de Tampons Anciens) [Nicolier *et al.*, 2003a] : il est le fruit de la collaboration entre

- la médiathèque de l'agglomération troyenne (MAT) qui participe au programme de numérisation des documents anciens (miniatures, livrets de colportage de la Bibliothèque Bleue, microfilm de manuscrits médiévaux),

- Les chercheurs du CEPLECA qui sont intéressés par l'exploitation de ce nouveau support,
- l'équipe image du laboratoire CReSTIC qui étudie les problèmes soulevés et apporte des solutions.

Les thématiques de recherche développées dans cette collaboration vont de l'impression virtuelle à partir de tampons anciens trop fragiles pour être ré-encrés [Millon *et al.*, 2002; Baudrier *et al.*, 2003] à l'évaluation par le contenu du degré de similarité entre images binaires [Baudrier *et al.*, 2004a] en passant par la réalisation d'images de profondeur à moindre coût [Thomas *et al.*, 2005]. Le schéma du processus d'évaluation par le contenu du degré de similarité entre images binaires est présenté dans le paragraphe 0.1.1, figure 0.10. Avant d'arriver au stade de comparaison des images binarisées, plusieurs étapes ont lieu : l'acquisition, le recalage et la binarisation. Nous allons les détailler dans la section qui suit.

1.3 Prétraitement

1.3.1 Acquisition des échantillons

Numérisation

Le programme de numérisation est en cours et nous n'avons pas encore accès à la base d'impressions anciennes numérisées. Il a donc fallu effectuer notre propre acquisition afin de pouvoir développer nos méthodes. Nous nous sommes basés sur les livres et recueils d'illustrations fournis par Alain Robert (CEPLECA) qui sont les suivants : [Varlot, 1859; bib, 1735; bib, fin XVIIIe; aym, debut XIXeb; aym, 1784; aym, milieu XIXe; aym, debut XVIIIe; aym, 1631; aym, debut XIXea]

La numérisation a ensuite été faite grâce au matériel suivant : une caméra CCD et une carte d'acquisition standard sous des conditions d'éclairage relativement uniforme réalisées à l'aide d'un tube fluo. Les impressions sous format numérique sont codées en 256 niveaux de gris. Les impressions ne sont alors ni centrées ni cadrées. Il est donc souhaitable de faire un recalage pour pouvoir les comparer localement.



(a) Acquisition initiale d'une impression provenant de l'histoire des quatre frères Aymons.



(b) Acquisition initiale de l'impression provenant de l'histoire des quatre frères Aymons.

FIG. 1.1 – L'acquisition initiale des impressions n'est pas recalée ni en taille, ni affinement, même si les rotations et translations ne sont pas importantes. Les déformations non affines (les bords supérieurs et inférieurs ne sont pas parallèles dans les images) sont plus difficiles à corriger d'autant qu'elles peuvent être dues à la forme du tampon.

Recalage

Nous ne connaissons pas la manière dont sera effectuée l'acquisition des impressions. Or l'étape de recalage ne peut se faire qu'en connaissant les conditions initiales. En effet, le recalage ne s'effectuera pas de la même manière selon que les impressions sont cadrées, de la même taille ou calées par rapport à un des coins d'une image de taille fixe ou encore centrées. Nous avons fait l'hypothèse que les impressions étaient cadrées, or elles sont souvent munies d'un cadre noir, ce qui fait qu'elles sont alors recalées. Les tests ont été réalisés sur des images recadrées, ce qui permet d'évaluer uniquement l'efficacité du processus de classification. D'autre part, dans le cas où les impressions ne seraient pas recalées, nous avons choisi un procédé de recalage affine basé sur l'ellipse équivalente qui est définie à partir des moments géométriques de l'image.

Choix des groupes de similarités

Il y a deux possibilités pour le choix des groupes de similarité pour la base des impressions anciennes.

- La première est de classer comme similaires les images représentant la même scène (classe C_{sim}) et comme dissimilaires celles illustrant des scènes différentes (classe C_{dissim}). Ce que nous entendons par représenter la même scène est d'avoir les mêmes éléments aux mêmes endroits avec une manière de les représenter qui peut varier. Le classement en deux classes est relativement aisé dans le sens où il y a peu d'ambiguïté.
- La deuxième possibilité est de distinguer au sein de la classe des images similaires celle qui proviennent du même tampon (un exemple est donné fig. 1.2) de celles qui représentent la même scène mais qui proviennent de deux tampons différents (fig. 1.1). Ce classement est plus délicat du fait de la dégradation de certains tampons. En effet, dans ce cas, il peut y avoir des impressions provenant du même tampon qui ont vieilli de manière tellement différente qu'elles se ressemblent moins que deux impressions provenant de deux tampons distincts mais dont l'un est une reproduction fidèle de l'autre.

Dans les deux cas, la classification a été faite en collaboration avec l'expert Alain Robert.

1.3.2 Binarisation

Les impressions anciennes sont obtenues en imprimant des tampons encrés, elles sont donc initialement en noir et blanc (binaires), cependant la qualité de l'encre et du papier, leur vieillissement qui se traduit souvent par une dégradation, en font des images en niveaux de gris fortement contrastées. Le vieillissement est variable et peut assombrir ou éclaircir les impressions. La première étape consiste à effectuer une binarisation permettant de s'affranchir en partie de ces variations, tout en ne perdant que peu d'information puisque les images initiales sont fortement contrastées.

La binarisation consiste à obtenir une image binaire (constituée de 0 et de 1 ou de 0 et 255 selon les cas) à partir d'une image en niveaux de gris. Le principe commun à toutes les méthodes est de calculer un seuil τ et d'affecter 1 si la valeur du pixel est supérieur à τ et 0 sinon.

Pour $\mathbf{b} : x \in X \mapsto \{0, \dots, 255\}$ une image en niveaux de gris, l'image binarisée \mathbf{b}_{bin} est donnée par

$$\mathbf{b}_{bin}(x) \begin{cases} 1 & \text{si } x > \tau, \\ 0 & \text{si } x \leq \tau \end{cases} \quad (1.1)$$

Le seuil τ peut être global (le même pour toute l'image), ou local (dont la valeur est calculée localement pour chaque pixel). Une étude comparative est présentée dans [Øivind Due Trier and Taxt, 1995]. Outre la méthode d'Otsu [Otsu, 1979] qui est classée comme la meilleure méthode globale (et qui est programmée dans Matlab), la méthode de Niblack [Niblack, 1986]



(a) Première impression du tampon.



(b) Deuxième impression du tampon.

FIG. 1.2 – Deux illustrations imprimées à partir du **même** tampon. Le vieillissement qui n'a pas été identique dans les deux cas donne deux images différentes. Sur l'image 1.2(b), l'encre s'est dégradée par endroit alors que l'ensemble est plus foncé que dans l'image 1.2(a).

obtient les meilleurs résultats sur l'ensemble des critères lorsqu'elle est suivie d'un post-traitement. La version de Sauvola [Sauvola *et al.*, 1997] et son amélioration par Wolf [Wolf *et al.*, 2002] permet une bonne binarisation sans post-traitement. Nous présentons ici ces différentes méthodes puis leurs résultats dans les tableaux 1.1 et 1.2.

Méthode d'Otsu

La méthode d'Otsu est une méthode de binarisation globale. Elle est basée sur l'histogramme de l'image : la séparation (par le seuillage) entre les deux classes est fait pour maximiser la variance inter-classes de l'histogramme. Soit $h[i], i \in [a, b]$ l'histogramme normalisé de l'image, où $[a, b]$ est l'ensemble des niveaux de gris ($a < b$) (exemple, $a=0$ et $b=255$) et $\sum_{i=a}^b h[i] = 1$. Un seuil $\tau, \tau \in [a, b-1]$, scinde l'ensembles des niveaux de gris en deux parties $[a, \tau]$ and $[\tau+1, b]$, représentant respectivement le fond et le premier plan. La séparabilité des classes associée à τ est définie par :

$$S(\tau) = \omega_0(\mu_0 - \mu)^2 + \omega_1(\mu_1 - \mu)^2,$$

avec les moments d'ordre 1 et 2, ω et μ , qui sont donnés par

$$\omega_0 = \sum_{i=a}^{\tau} h[i], \quad \omega_1 = \sum_{i=\tau+1}^b h[i], \quad (1.2)$$

$$\mu_0 = \sum_{i=a}^{\tau} ih[i], \quad \mu_1 = \sum_{i=\tau+1}^b ih[i], \quad \mu = \sum_{i=a}^b ih[i]. \quad (1.3)$$

La méthode d'Otsu trouve le seuil τ qui maximise $S(\tau)$.

Méthode de Niblack

L'idée de cette méthode est de faire varier le seuil de binarisation en fonction de la moyenne locale et de l'écart-type local. Le seuil τ au pixel (x, y) est calculé de la manière suivante :

$$\tau(x, y) = \text{poids} \times \sigma(x, y) + \mu(x, y) \quad (1.4)$$

où $\sigma(x, y)$ et $\mu(x, y)$ désignent l'écart-type et la moyenne locale estimés dans un voisinage local de (x, y) de taille $k \times k$. Le poids sert à préciser quelle proportion des contours des objets est intégrée dans l'objet. La valeur recommandée pour ce poids est de $-0,2$ si le fond est plus clair que les objets (ce qui est notre cas) et $0,2$ si le fond est plus foncé. Le problème de la méthode de Niblack est que dans des zones de fond faiblement perturbées, des motifs peuvent apparaître à la binarisation. En effet, dans ces zones, la variance locale est très faible et en conséquence le seuil aussi. De faibles valeurs peuvent alors se retrouver au dessus du seuil et faire apparaître de l'information là où il n'y en a pas (une illustration en est donnée avec la binarisation de l'impression 1, tableau 1.1)

Méthode de Sauvola, amélioration de Wolf

Pour résoudre ce problème, [Sauvola *et al.*, 1997] ont proposé une version améliorée du seuil :

$$\tau(x, y) = \mu(x, y) - \text{poids} \mu(x, y) \left(1 - \frac{\sigma(x, y)}{R}\right) \quad (1.5)$$

où R désigne la gamme dynamique de l'écart-type ($R = 128$ pour des images codées sur 256 niveaux de gris), la valeur recommandée pour le poids est de $0,5$. Cependant cette version souffre de deux défauts qui sont la source d'une version améliorée [Wolf *et al.*, 2002] :

- pour les images faiblement contrastées, la gamme dynamique de l'écart-type peut être notablement réduite par rapport à la gamme théorique, il est donc préférable de la mesurer directement sur l'image.
- La prise en compte de l'histogramme peut être améliorée en le normalisant par rapport à la valeur la plus faible de l'image.

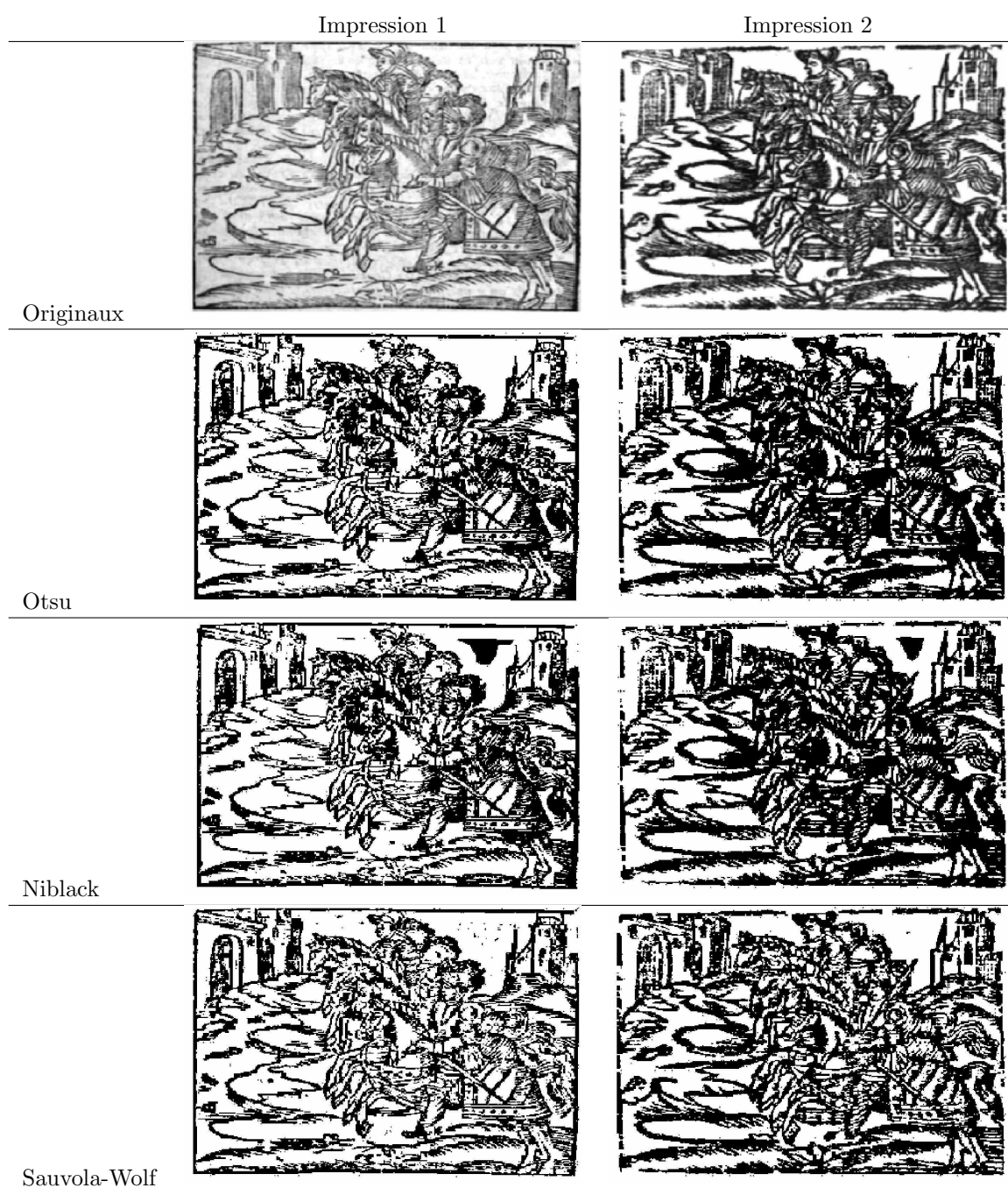
On obtient alors la formule suivante :

$$\tau(x, y) = \mu(x, y) - \text{poids}(\mu(x, y) - M) \left(1 - \frac{\sigma(x, y)}{R}\right) \quad (1.6)$$

avec, en récapitulant, $\mu(x, y)$ la moyenne locale, $\sigma(x, y)$ l'écart-type local, R la gamme dynamique de l'écart-type et M le minimum de l'histogramme.

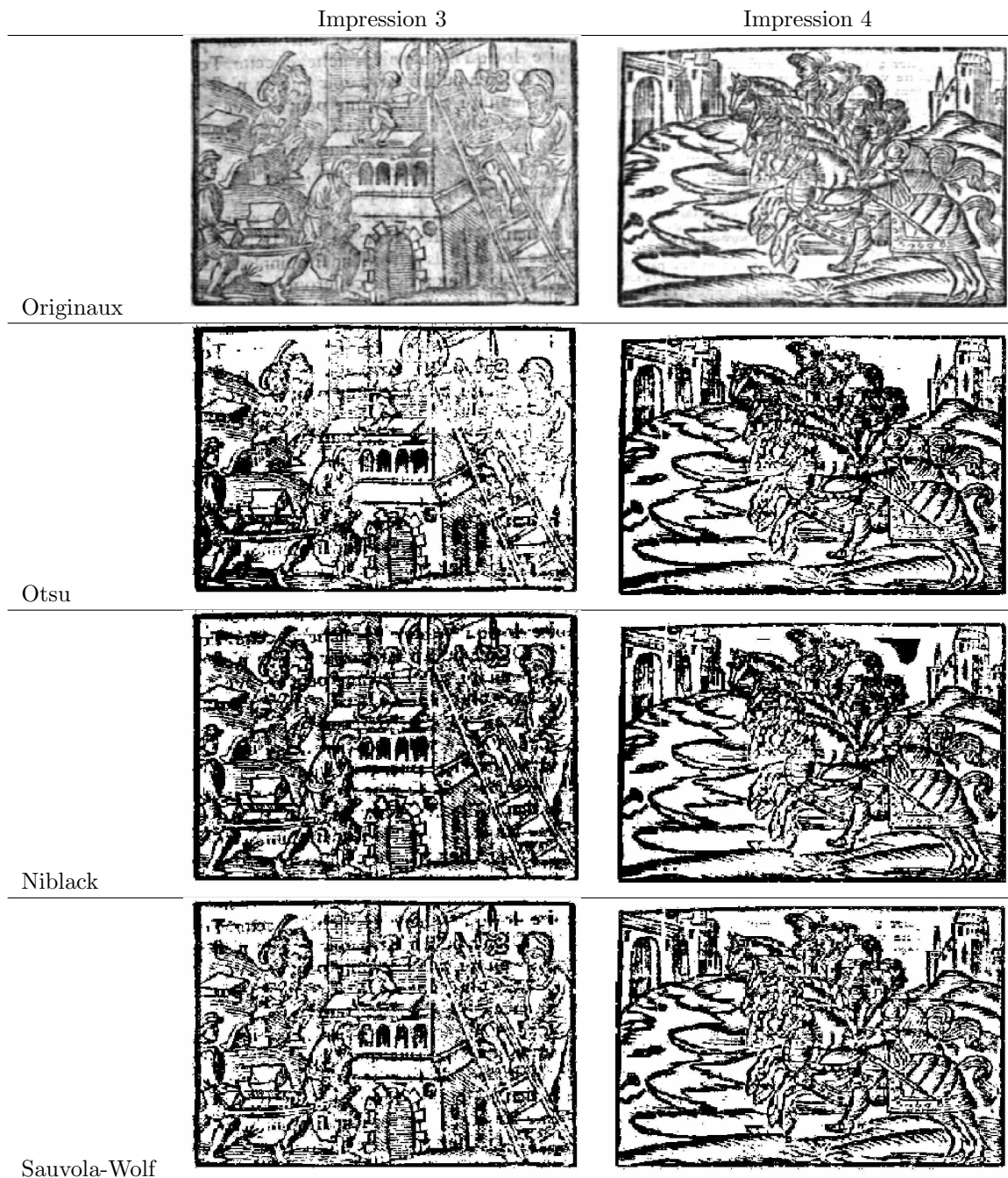
Résultats

Dans le processus de classification, la binarisation est une étape importante car elle peut être à l'origine d'erreurs de classification. Sur l'impression 3, qui est la plus délicate car elle est faiblement contrastée, on voit que la méthode d'Otsu est insuffisante dans les zones de faible contraste (au niveau du bâtiment et de l'échelle par exemple). D'autre part, la méthode de Niblack est trop sensible dans les zones faiblement contrastées comme dans l'image 3 ou dans les images 1, 2 et 4 dans la zone du ciel. Enfin, en comparant les binarisations des images 1,2 et 4, on voit que la méthode qui rend le plus fidèlement le niveau de gris moyen est la méthode de Sauvola-Wolf.



TAB. 1.1 – La binarisation des impressions 1 et 2 montrent que la méthode de Wolf est mieux adaptée que celles d'Otsu et de Niblack.

Ainsi, nous avons présenté des méthodes de binarisation. La méthode avec un seuillage global s'est révélée insuffisante. L'étude des méthodes basées sur un seuillage local a abouti au choix de la méthode de Sauvola améliorée par Wolf.



TAB. 1.2 – L'exemple de l'impression 3 montre dans la binarisation de Wolf des détails imperceptibles à l'œil dans l'impression en niveaux de gris. L'aspect dégradé de l'impression 4 fait que la qualité de toutes les binarisations testées est médiocre.

1.4 Bilan

Nous avons présenté dans cette partie le schéma du processus global de classification basé sur la CDL pour la comparaison d'images. Nous avons ensuite présenté une application de ce processus qui est l'évaluation de la similarité dans une base d'impression anciennes. Cette base a été constituée par nos soins en collaboration avec la MAT et A. Robert. Le prétraitement

des images obtenues a alors été détaillé et il comporte un recalage que nous avons fait manuellement et le choix de la méthode de binarisation. Différentes méthodes de binarisation ont été testées et notre choix s'est porté sur la méthode de Sauvola améliorée par Wolf.

Dans la partie suivante, le module de classification à proprement parler est abordé. Les performances du processus global sont évaluées sur cette base d'impressions anciennes. Puis le processus global est appliqué à d'autres bases : une base de formes et une de visages (sans que l'étude soit approfondie dans ce dernier cas).

2

Classification basée sur la Carte des Dissimilarités Locales

Lorsque la CDL est calculée, il faut prendre une décision visant à déterminer si les images sont similaires ou non. Cela revient de fait à classer la CDL dans l'une des deux classes C_{sim} ou C_{dissim} . Dans ce but, nous avons testé plusieurs méthodes qui extraient de l'information de la CDL afin de prendre la décision.

Le descripteur le plus simple à considérer est un réel comme le maximum de la CDL ou sa moyenne. Prendre le maximum sur la CDL revient à calculer la distance de Hausdorff globale, ce qui peut être fait sans le recours à la CDL. De même, calculer la moyenne sur la CDL revient à calculer la MHD. En fait, pour tirer parti de la CDL, il faut prendre plus qu'un seul élément. Par ordre croissant de complexité nous avons considéré l'histogramme de la CDL, les courbes granulométriques et enfin la CDL elle-même.

Le principe dans tous les cas reste le même : un apprentissage statistique est effectué sur un échantillon de descripteurs dans les deux classes. Il en résulte un modèle pour chaque classe. Ensuite, lorsqu'une CDL est présentée, une comparaison est effectuée avec les modèles pour évaluer son degré d'appartenance à chaque classe. Dans tous les cas de l'histogramme ou de la granulométrie la méthode qui a donné les meilleurs résultats est la méthode des séparateurs à vaste marge (SVM) [Vapnik, 1998]. Dans le cas où c'est la carte elle-même qui est donnée en entrée, la grande dimension des données conduit aussi au choix des SVM comme méthode de classification. Une section présente le cadre général de la classification et les principales méthodes de classification. Puis ces différentes méthodes de décision utilisant l'information de la CDL sont présentées ici, ainsi que leurs résultats. L'aspect théorique de la classification est détaillée dans [Ragot, 2003].

2.1 Classification : généralités

2.1.1 Apprentissage

La tâche d'un module d'apprentissage est d'apprendre un modèle au travers d'exemples. Un module d'apprentissage tire ses caractéristiques de la classe de fonctions F qu'elle peut mettre en œuvre. Dans notre cas, les fonctions considérées sont des fonctions de décision. Cependant la définition d'un module d'apprentissage n'indique pas comment obtenir un classificateur adapté à la tâche considérée.

2.1.2 Formalisation de la fonction de classification

Un classifieur peut être vu comme une fonction de décision f qui associe à une entrée e décrite dans un espace E une sortie s d'un espace S

$$f : e \in E \rightarrow s = f(e) \in S \quad (2.1)$$

Les entrées du système représentent les objets à classer et les sorties les classes ou catégories auxquelles ils appartiennent. Le rôle du classificateur est d'identifier chaque objet afin de retrouver la classe à laquelle il appartient parmi ceux de l'espace de sortie. Cela revient à déterminer dans l'espace des objets les régions associées à chaque catégorie de S . Ces régions sont appelées régions de décision et la frontière entre deux catégories est appelée frontière de décision. Pour faire cette identification, le classificateur utilise la fonction de décision f qui confronte l'entrée à une modélisation M du problème. Pour marquer le lien entre la modélisation et la fonction de décision nous noterons cette dernière f_M par la suite.

On distingue trois étapes dans la mise en œuvre d'un système de reconnaissance. La première étape correspond à la conception du classificateur c'est-à-dire à l'élaboration de la modélisation M et à son association avec la fonction de décision f_M . Cette étape peut être faite par un expert ou par un apprentissage automatique à partir d'un ensemble de données du problème appelé ensemble ou base d'apprentissage. C'est la phase d'apprentissage. La deuxième étape consiste à évaluer les performances du système à partir d'un ensemble de données, la base de test, pour déterminer ses capacités de généralisation. Elle est couramment appelée phase de généralisation. Finalement, si le système est intégré dans un cadre applicatif réel, il rentre dans sa phase d'exploitation.

Dans les paragraphes suivants nous décrivons plus précisément chacune des trois entités caractérisant le système de classification : les entrées, les sorties et la modélisation.

Espace de représentation des entrées

D'une façon générale l'entrée d'un classifieur est une forme qui correspond à une observation de l'environnement extérieur au système. Cependant la quantité de données à traiter en entrée peut vite devenir très importante (matrice de pixels par exemple) et dans celle-ci l'information utile pour l'identification des classes n'est pas toujours directement accessible. C'est donc le classificateur qui doit compenser cette faiblesse. Une possibilité très souvent utilisée en pratique, bien que nécessitant des traitements supplémentaires, consiste à extraire du signal un certain nombre de caractéristiques ou attributs décrivant l'entrée. Celles-ci sont très souvent numériques (quantité de pixels noirs, centre de gravité, position de points remarquables, etc.) mais peuvent aussi dans certains cas être symboliques. L'objectif est d'obtenir une description à la fois réduite et la plus pertinente possible, c'est-à-dire contenant le maximum d'information pour pouvoir différencier les classes et donc faciliter le travail du classificateur. À ce niveau, une sélection et une hiérarchisation des caractéristiques peuvent être faites soit automatiquement, soit *a priori* afin de n'employer que les plus pertinentes dans un contexte donné. D'autres connaissances *a priori* peuvent aussi être introduites telles que les relations entre les différentes parties des entrées du classificateur pour représenter une information importante. On parlera dans ce dernier cas de caractéristiques de haut niveau.

Nous présentons ici deux types de représentations : les représentations vectorielles (de faible dimension) telles que l'histogramme ou la courbe granulométrique (sec. 2.3) et une représentation en deux dimensions qui est la CDL elle-même.

Dans la suite, nous ne considérerons que le cas le plus général sur le mode de représentation d'une entrée e lorsqu'elle est décrite par un vecteur (si l'entrée est une image, alors elle sera traitée sous la forme d'une seule colonne). Il est à noter que dans cette représentation, l'information spatiale entre les pixels de l'image n'est pas perdue dans la mesure où l'image peut être reconstruite à partir du vecteur-colonne.

Les représentations de faible dimension d'une image conduisent à une perte d'information qui comme on va le voir engendre

de moins bonnes performances. Dans l'autre cas, un problème apparaît aussi car la plupart des algorithmes de classification gère mal les vecteurs de grande dimension. Or la représentation vectorielle des images donnent des vecteurs de très grande dimension. Heureusement, bien que la plupart des méthodes d'apprentissage ont de grandes difficultés à traiter les données de grande dimension, les SVM (sec. 2.2.4) s'accommodent bien de ce type de données.

Espace de sortie

Les classes de S correspondent à l'identité des objets par rapport au problème de classification considéré. Ils sont représentés par leurs étiquettes (ou labels). Les classes sont généralement connues *a priori* par expertise du problème. Cependant il arrive que cet espace de sortie ne soit pas initialement connu. Il devra alors être déterminé automatiquement par apprentissage.

Modélisation et apprentissage automatique à partir de données

Une des principales difficultés en classification est de choisir et de concevoir la fonction de décision f et la modélisation M sur laquelle elle repose. Parfois des connaissances *a priori* sur le problème de classification sont disponibles suite à expertise. Elles peuvent alors permettre de choisir le type de classificateur (et donc de modélisation) le plus adapté à la situation. Cependant la plupart du temps, très peu de connaissances directement exploitables sont disponibles pour l'élaboration du système. Seul un ensemble d'observations est recueilli. Ces données sont souvent difficiles et complexes à analyser pour en extraire les informations nécessaires à l'élaboration du système.

Apprentissage automatique d'un classificateur

Le mécanisme d'apprentissage a pour objectif d'extraire d'une base d'apprentissage l'ensemble des connaissances nécessaires à la modélisation du problème et à les structurer. Les données de la base sont appelées observations, échantillons, individus ou encore exemples suivant les auteurs et les contextes. Quand elles ont été expertisées au préalable pour leur attribuer l'étiquette d'une des classes du problème, ces informations peuvent être utilisées lors de l'apprentissage qui est alors qualifié de *supervisé* (ou avec professeur). Dans le cas d'un apprentissage *non supervisé* (sans professeur), l'information concernant les classes n'est pas disponible *a priori* et l'apprentissage consiste à identifier une structure dans les données en regroupant celles qui possèdent des propriétés similaires, ce qui est appelé induction. Une des limites de l'induction et des méthodes d'apprentissage à partir de données réside en général dans leur principe même : comme la modélisation est établie à partir d'un ensemble d'observations, celles-ci doivent être suffisamment représentatives pour que les capacités de généralisation du modèle soient bonnes. Par contre, si la modélisation s'appuie trop sur les détails de la base d'apprentissage le classifieur deviendra trop spécialisé et n'aura pas de bonnes capacités de généralisation. Ce phénomène est appelé sur-apprentissage.

Évaluation des performances d'un classificateur

Nous décrivons dans ce paragraphe les critères les plus souvent employés pour évaluer la performance des systèmes de reconnaissance en phase de généralisation. L'objectif est d'obtenir une estimation la plus fidèle possible du comportement du système dans des conditions réelles d'utilisation. Pour cela des critères classiques comme les taux de reconnaissance et taux d'erreurs sont presque systématiquement utilisés.

Taux de reconnaissance et taux d'erreur Les taux de reconnaissance et d'erreur permettent d'évaluer la qualité du classificateur f_M par rapport au problème pour lequel il a été conçu. Ces taux sont évalués grâce à une base de test qui contient des formes décrites dans le même espace de représentation E que celles utilisées pour l'apprentissage. Elles sont aussi étiquetées par leur classe réelle d'appartenance afin de pouvoir vérifier les réponses du classificateur. Pour que l'estimation du taux de reconnaissance soit la plus fiable possible il est primordial que la base de test n'ait aucun individu en commun avec la base d'apprentissage et les éventuelles bases de validation. De plus cette base de test doit être suffisamment représentative du problème de classification. En général quand les échantillons étiquetés à disposition sont suffisamment nombreux ils sont séparés en deux parties disjointes et en respectant les proportions par classe de la base initiale. Une partie sert pour former la base d'apprentissage et l'autre pour former la base de test. Les performances en terme de taux de reconnaissance sont alors déterminées en présentant au classificateur chacun des exemples de la base de test et en comparant la classe donnée en résultat à la vraie classe. En considérant que la base de test contient N individus et que sur ceux-ci $N_{corrects}$ sont biens classés par le système, le taux de reconnaissance est simplement défini par :

$$\tau_{rec} = 100 \frac{N_{corrects}}{N}. \quad (2.2)$$

Le taux d'erreur τ_{err} est défini à partir du nombre d'individus N_{err} mal classés, c'est-à-dire les individus pour lesquels la classe affectée par f diffère de celle donnée par l'expert :

$$\tau_{err} = 100 \frac{N_{err}}{N}. \quad (2.3)$$

Si les taux de reconnaissance et d'erreur peuvent fournir une estimation des capacités de généralisation du classificateur, ils ne permettent pas d'évaluer la stabilité de la méthode d'apprentissage par rapport à des variations sur le contenu de la base d'apprentissage. Le mécanisme de validation croisée permet de répondre aux deux problèmes précédents. Il consiste à diviser la base contenant l'ensemble des échantillons en k parties disjointes et de tailles équivalentes. k étapes d'apprentissage/validation sont ensuite faites. L'étape d'apprentissage est effectuée sur l'union des parties sauf une et le test sur cette dernière. Le taux de reconnaissance est alors égal à la moyenne des taux obtenus à chaque étape. La stabilité de la méthode d'apprentissage peut être évaluée en calculant l'écart-type du taux de reconnaissance sur ces k étapes. Citons par exemple le cas où $k = N_{tot}$ avec N_{tot} le nombre total d'éléments disponibles. Cette méthode de validation est appelée *le test circulaire*. C'est elle qui donne en général la meilleure approximation du taux de reconnaissance réel du classificateur. Son inconvénient majeur est qu'elle demande beaucoup de temps de calcul.

2.2 Différentes approches de classification

Dans ce paragraphe nous passons en revue quelques exemples de systèmes et de méthodes de classification couramment utilisés. Nous nous sommes concentrés ici sur les méthodes reposant sur un apprentissage supervisé.

2.2.1 Méthodes probabilistes

Les méthodes probabilistes modélisent les entrées comme des variables aléatoires suivant des lois différentes selon leur classes d'appartenance. Notons $\omega_1, \dots, \omega_n$ les classes de S . L'objectif de la modélisation est d'établir les probabilités *a posteriori* des classes c'est-à-dire la probabilité d'obtenir la classe ω_j sachant que l'entrée e a été observée. Cette probabilité

est déterminée à partir de la règle de Bayes :

$$P(\omega_j|e) = \frac{p(e|\omega_j)P(\omega_j)}{p(e)}, \quad (2.4)$$

où $P(e|\omega_j)$ est la densité de probabilité conditionnelle de e , $P(\omega_j)$ est la probabilité *a priori* de la classe ω_j et $p(e)$ est la densité de probabilité de e définie par

$$p(e) = \sum_{i=1}^n p(e|\omega_i)P(\omega_i). \quad (2.5)$$

La difficulté de cette approche réside essentiellement dans la détermination des densités $p(e|\omega_i)$. Elles peuvent être estimées de façon paramétriques ou non-paramétriques selon les données du problème. Les estimations les plus connues sont la méthode bayésienne et la méthode du maximum de vraisemblance. Pour plus d'information sur ces méthodes, le lecteur peut se référer à [Barbe and Ledoux, 1998].

2.2.2 k plus proches voisins

L'approche par les k plus proches voisins (k-PPV) [Cover and Hart,] possède au moins deux avantages : elle est très simple et elle ne nécessite pas d'apprentissage. La modélisation utilisée pour la classification est simplement la base d'apprentissage elle-même qui constitue l'ensemble des connaissances du système. Il n'y a donc pas d'architecture à proprement parler. La règle de décision utilisée consiste à attribuer à une entrée e la classe représentée majoritairement dans son voisinage. Ce voisinage est défini par les k plus proches individus de e parmi ceux se trouvant dans la base d'apprentissage. Les seuls paramètres à déterminer sont donc k ainsi que la métrique utilisée pour comparer les individus et trouver les plus proches voisins. Dans plusieurs cadres applicatifs, cette approche a donné des résultats intéressants. Cependant elle possède plusieurs inconvénients. Son efficacité dépend directement de la pertinence de la base d'apprentissage et notamment de sa densité dans les différentes régions de l'espace des données. La présence de bruit (données mal étiquetées) est aussi fortement handicapante. Autre inconvénient : la base d'apprentissage entière doit être stockée, ce qui nécessite en général une place mémoire importante. Enfin la recherche des plus proches voisins est coûteuse et cela d'autant plus que la métrique utilisée est complexe.

2.2.3 Réseaux de neurones de type perceptron

Les classificateurs de type réseaux de neurones proviennent d'une modélisation des neurones biologiques [McCulloch and Pitts, 1943]. Du point de vue de la classification, ils reposent sur une modélisation discriminante [Bishop, 1995]. Un neurone permet de définir une fonction discriminante linéaire g dans l'espace de représentation E des entrées. Cette fonction réalise une combinaison linéaire du vecteur de caractéristiques de l'entrée e

$$g(e) = \omega^t e + \omega_0, \quad (2.6)$$

où ω est un vecteur de poids de la combinaison linéaire et ω_0 est le biais. Ainsi $g(e) = 0$ définit un hyperplan permettant de séparer E en deux régions de décision. En déterminant les bonnes valeurs de ω et ω_0 , celles-ci peuvent être associées à deux classes pour faire leur discrimination. Pour des problèmes à n classes, n fonctions discriminantes sont établies. L'extension à des réseaux de type perceptron multi-couches (PMC) permet d'obtenir des frontières de décision de complexité quelconques. L'apprentissage des poids est généralement fait par des méthodes comme la rétro-propagation du gradient de l'erreur.

Les PMC ont été très largement utilisés en classification et en reconnaissance de formes notamment pour leur bonnes performances et leur simplicité. Ainsi, l'algorithme de rétro-propagation du gradient de l'erreur risque de converger vers un

minimum local donnant une solution sous-optimale. La détermination de l'architecture et des paramètres du réseau est aussi un inconvénient majeur. En effet avant d'entreprendre l'apprentissage il faut déterminer le nombre de couches cachées du classificateur ainsi que le nombre de neurones de chacune de ses couches. Aussi sans connaissances *a priori* il faut rechercher l'architecture optimale par essais successifs.

2.2.4 Séparateurs à vaste marge

Cette sous-section décrit brièvement le principe des Séparateurs à Vaste Marge (SVM, en Anglais Support Vector Machine). Les SVM sont une méthode de classification et de régression introduite par [Boser *et al.*, 1992], pour une description détaillée, voir [Vapnik, 1998]. Les SVM sont particulièrement efficaces pour la classification supervisée car ils peuvent traiter des problèmes dépendant d'un grand nombre de descripteurs et qu'ils ont été utilisés avec succès sur des problèmes réels de haute dimension. Par exemple, en reconnaissance de formes, les SVM ont servi à la reconnaissance de caractères écrits isolés [Cortes and Vapnik, 1995], de reconnaissance d'objets [Blanz *et al.*, 1996], dans la détection de visages dans les images [Osuna *et al.*, 1997] et dans la catégorisation de texte [Joachims, 1998].

L'étude des différentes SVM pour la classification des SVM a été réalisée lors du stage de master 2 par [Aït Aouït, 2004] et les résultats présentés ici en sont une version approfondie.

Cas linéairement séparable Soit un ensemble d'apprentissage $\{x_i, y_i\}_{i=1..l}$ où $x_i \in \mathbb{R}^d$ et $y_i \in \{-1, 1\}$, y_i définissant la classe d'un exemple donné. L'objectif des SVM est de trouver un hyperplan séparant l'ensemble d'apprentissage de sorte que tous les points d'une même classe se retrouvent du même côté de l'hyperplan. Ce qui revient (après une éventuelle normalisation) à trouver un hyperplan $h(x) = 0$ défini par $\omega \in \mathbb{R}^d$ et $b \in \mathbb{R}$ tel que :

$$h(x) = \langle \omega, x \rangle + b, \quad (2.7)$$

$$y_i(\langle \omega, x_i \rangle + b) \geq 1 \quad i = 1, \dots, l \quad (2.8)$$

Parmi l'ensemble des hyperplans satisfaisant ces conditions, les SVM cherchent celui qui maximise la distance entre l'hyperplan et les points les plus proches de chaque classe, cette distance valant $\frac{2}{\|\omega\|}$. Ainsi le problème global équivaut à un problème de minimisation sous contrainte :

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \quad (2.9)$$

$$\text{tel que } y_i f(x_i) \geq 1 \quad i = 1, \dots, l \quad (2.10)$$

Dans la mesure où $\|\omega\|^2$ est un critère convexe et l'ensemble des contraintes sont linéaires, ce problème peut se résoudre en utilisant les multiplicateurs de Lagrange α et le dual lagrangien de ce problème. Ce qui permet d'écrire le vecteur ω sous la forme

$$\omega = \sum \alpha_i^* y_i x_i \quad (2.11)$$

où les α_i^* sont la solution du problème de minimisation :

$$\min_{\alpha} \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle - \sum_{i=1}^l \alpha_i \quad (2.12)$$

$$\text{avec } \begin{cases} \sum_{i=1}^l \alpha_i y_i = 0 \\ \forall i, \alpha_i \geq 0 \end{cases} \quad (2.13)$$

Ce qui est un problème d'optimisation quadratique standard qui se résout avec des outils courants. Finalement, l'hyperplan de séparation est le noyau de la fonction :

$$h(x) = \sum_{i=1}^l \alpha_i y_i \langle x_i, x \rangle + b, \quad (2.14)$$

où b est le multiplicateur de Lagrange associé à la contrainte $\sum_{i=1}^l \alpha_i y_i = 0$ du problème dual. La classe du point x est alors donnée par le signe de h .

Cas non-linéaire Dans le cas où les données ne sont pas linéairement séparables, c'est-à-dire dans beaucoup de problèmes réels, l'espace de données est transformé en un espace H de dimension supérieure appelé espace de redescription grâce à une fonction Φ . Dans l'espace H , le problème a des chances d'être linéairement séparable. Il suffit alors de revenir dans l'espace initiale pour avoir une classification. Le problème de séparation dans l'espace H s'obtient simplement puisque d'après les équations (2.12) et (2.15), les points de l'espace original x_i n'interviennent dans la solution qu'au travers de leurs produits scalaires. Ainsi le cas non-linéaire s'obtient aisément en remplaçant le produit scalaire $\langle x_i, x_j \rangle$ par $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$. Et la fonction de décision est alors :

$$h(x) = \sum_{i=1}^l \alpha_i y_i k(x_i, x) + b, \quad (2.15)$$

En pratique, les noyaux les plus utilisés sont les noyaux gaussiens $k(x, y) = \exp(-\frac{\|x-y\|^2}{2\sigma^2})$ (d'écart-type σ), les noyaux polynômiaux d'ordre p $k(x, y) = (\langle x, y \rangle + 1)^p$. On retrouve le cas linéaire lorsque le noyau est linéaire : $k(x, y) = \langle x, y \rangle$

Cas non-séparable Dans le cas où l'ensemble d'apprentissage n'est pas linéairement séparable, (*i.e.* il n'existe pas d'hyperplan pouvant séparer les classes), il est nécessaire d'introduire des variables de relâchement dans l'équation des contraintes. Le problème 2.9 devient alors :

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i^2 \quad (2.16)$$

$$\text{tel que} \quad y_i h(x_i) \geq 1 - \xi_i \quad i = 1, \dots, l \quad (2.17)$$

où ξ_i est la variable de relâchement d'une contrainte et C le coefficient de pénalisation du relâchement. On peut montrer que ce problème revient au cas non-linéaire en remplaçant $k(x_i, x_j)$ par $k(x_i, x_j) + \frac{1}{C} \delta_{i,j}$ où $\delta_{i,j}$ est le symbole de Kronecker [Cristianini and Shawe-Taylor, 2000].

Ainsi, pour prendre une décision en utilisant les SVM, plusieurs choix doivent être faits : le choix du type de SVM, le choix du noyau et le choix du coefficient de pénalisation C . Il y a de plus le choix de la taille des échantillons d'apprentissage et de test. Ces choix dépendent de l'application et nous les faisons *a posteriori* en fonction des performances obtenues. Ils sont présentés dans la section 2.4.

2.3 Représentation vectorielle de la Carte des Dissimilarités Locales

Maintenant que nous avons présenté le cadre de la classification, nous abordons les données permettant de classer les CDL. L'histogramme et la granulométrie fournissent une suite de valeurs réelles sur laquelle il faut fonder une décision statistique. Nous présentons ici les méthodes de décision statistiques que nous employons ainsi que leurs résultats. Ces méthodes exploitent la CDL. Or il y a plusieurs distances avec laquelle calculer la CDL. Celle qui donne les meilleures

mesures locales est la distance de Hausdorff. C'est donc la CDL basée sur la DH (CDL_{DH}) qui a servi de support pour tester les méthodes de décisions statistiques. Nous la notons CDL lorsque cela n'est pas ambigu.

2.3.1 Histogramme de la Carte des Dissimilarités Locales

L'histogramme est un descripteur de l'image largement utilisé en traitement de l'image, par exemple en recherche d'image par le contenu [Rui *et al.*, 1999] ou en segmentation [Philipp-Foliguet *et al.*, 2002]. Il s'agit en général de caractériser tout ou partie de l'image avec son histogramme (et éventuellement d'autres descripteurs) puis de l'utiliser pour évaluer la similarité entre deux images ou pour segmenter l'image. Notre optique est différente : l'histogramme ne sert pas ici à décrire les images à comparer mais à caractériser la CDL issue de la comparaison.

Une CDL regroupe les valeurs des mesures locales de dissimilarité. Intuitivement, une CDL provenant de deux images dissimilaires comporte plus de valeurs hautes qu'une CDL provenant de deux images similaires. En conséquence, les histogrammes correspondant ne doivent pas avoir la même allure. C'est cette intuition qui pousse à utiliser l'histogramme comme descripteurs. Deux méthodes de classification des histogrammes ont été testées :

- la première est explicite, un histogramme de référence est choisi dans chaque classe puis l'histogramme à classer est comparé aux histogrammes de référence de chaque classe grâce à une distance euclidienne. L'histogramme de référence est l'histogramme moyen.
- La deuxième repose sur les SVM.

L'ensemble d'apprentissage est constitué de 50 éléments pour la classe C_{sim} et de 50 éléments pour C_{dissim} . L'ensemble de test de 75 éléments dans C_{sim} et de 200 dans C_{dissim} . Les résultats présentés ont été obtenus à l'échelle 3, ce qui correspond à des images de taille 128×128 . Les autres échelles donnent de moins bons résultats pour la première méthode. Les résultats sont présentés dans le tableau 2.1. Ils présentent le taux de réussite pour la méthode de la moyenne et les meilleurs taux pour les SVM. Pour trouver le meilleur taux de réussite, plusieurs noyaux ont été testés (linéaire, polynômiaux de degré 1 à 3 et gaussiens) ainsi que le paramètre C entre 1 et 10000. Les résultats sont présentés sous forme de graphes dans la figure 2.1. Les noyaux gaussiens n'y sont pas présentés car leur performances sont médiocres (inférieures à 50%).

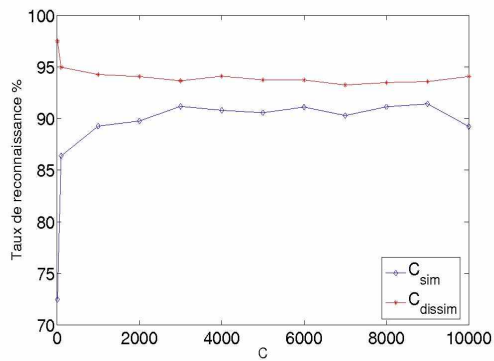
méthode	Histogramme moyen	SVM
provenant de C_{sim}	84%	91%
provenant de C_{dissim}	95%	94%

TAB. 2.1 – Histogramme de la CDL : taux de reconnaissance à la résolution 128×128 .

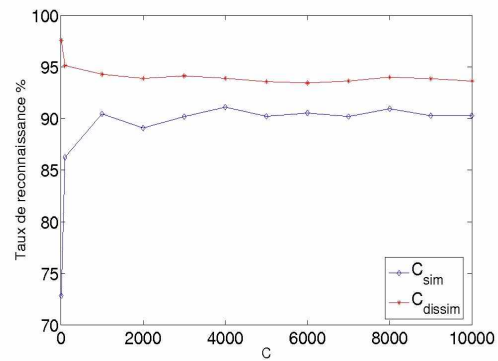
Commentaires Les résultats sont meilleurs qu'avec une mesure globale. Cependant, l'histogramme perd complètement l'information spatiale contenue dans la CDL. Et cette information a une importance certaine pour déterminer la classe comme la figure 2.2 l'illustre. La répartition des distances dans la carte n'est pas la même mais les histogrammes sont eux par contre très proches. Les résultats moyens de la décision basée sur l'histogramme proviennent de la carence de l'histogramme en information spatiale contenue dans la CDL_{DH} . Il est possible d'extraire de l'information sur la distribution spatiale de la CDL_{DH} grâce à la granulométrie.

2.3.2 Granulométrie

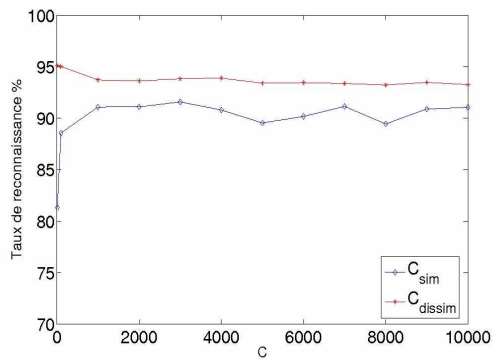
La granulométrie présentée dans la sous-section 2.4.1 (p. 31) est applicable aux images en niveaux de gris et donc aux CDL. Son intérêt est de pouvoir caractériser le regroupement des valeurs dans l'image : par exemple la courbe granulométrique



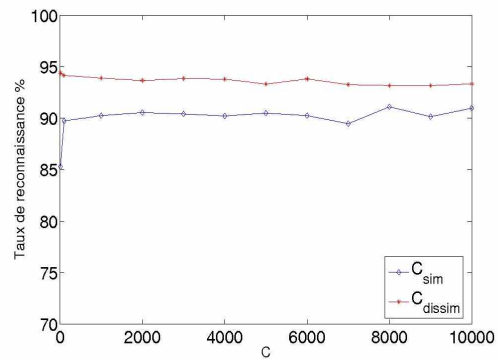
(a) Noyau linéaire



(b) Noyau polynômial de degré 1



(c) Noyau polynômial de degré 2



(d) Noyau polynômial de degré 3

FIG. 2.1 – Taux de reconnaissance de la classification des histogrammes basée sur les SVM en testant les noyaux linéaires et polynômiaux et pour C variant entre 1 et 10000. Les meilleurs résultats sont de 91% pour C_{sim} et 94% pour C_{dissim} pour tous les noyaux. Le moins gourmand en temps de calcul reste le noyau linéaire qui obtient les meilleurs taux pour $C=4000$.

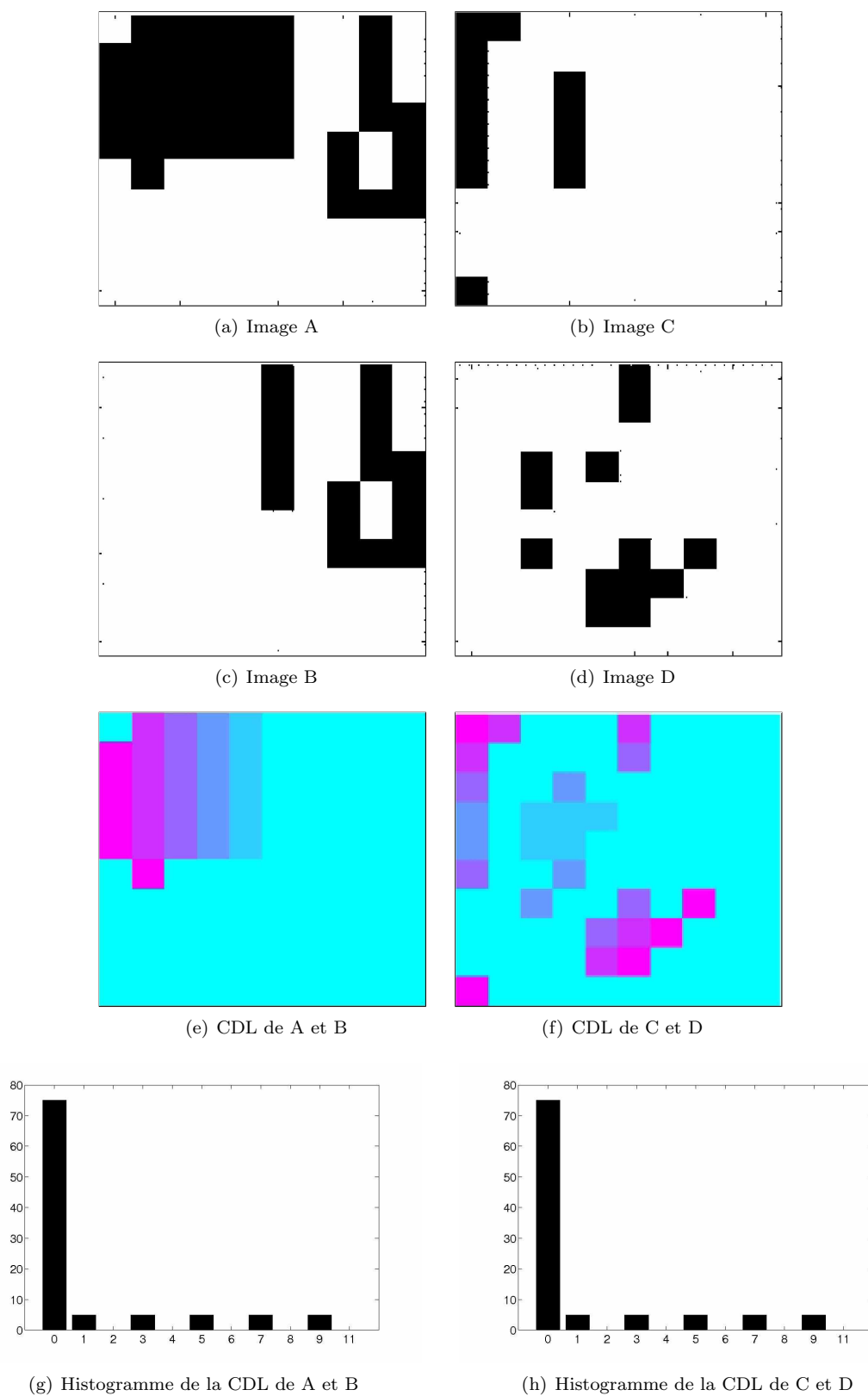


FIG. 2.2 – La CDL des images A et B et celle des images C et D sont spatialement très différentes. Cependant leurs histogrammes sont identiques. Une décision basée sur l'histogramme ne peut pas distinguer ces deux CDL.

permet de savoir si les hautes valeurs de l'image agglomérées en gros amas ou ponctuelles. Cette information ne caractérise pas complètement la distribution spatiale de la CDL mais apporte un renseignement dessus.

En pratique

Il y a deux possibilités pour la granulométrie (cf sous-section 2.4.1) : celle par ouverture ou celle par fermeture. La granulométrie par ouverture permet de caractériser les structures qui ont des plus hautes valeurs que leur voisinage, et pour la granulométrie par fermeture, c'est le contraire. Dans la CDL, c'est la granulométrie par ouverture qui est intéressante pour les images binaires.

Pour pouvoir donner une granulométrie, la famille d'éléments structurants doit vérifier la propriété d'absorption. Cependant la famille d'éléments structurant avec laquelle les résultats sont les meilleurs est une famille de carrés qui ne vérifie pas cette propriété d'absorption. L'information concernant la taille des structures présentes reste valable, mais si on applique deux ouvertures successives, on n'obtient pas le résultat égal à celui de l'ouverture la plus grande, ce qui doit être le cas si on veut conserver le parallèle avec le tamisage. Dans notre cas, cette propriété n'est pas essentielle.

La « courbe granulométrique » est représentée par un vecteur en dimension 30, ce qui est assez important. Il est donc naturel d'utiliser les SVM pour classer les courbes granulométriques. De plus, cela permet de comparer les différentes représentations combinées avec les mêmes classificateurs. Pour trouver le meilleur taux de réussite, plusieurs noyaux ont été testés (linéaire, polynômiaux de degré 1 à 3 et gaussiens) ainsi que le paramètre C entre 1 et 10000. Les résultats sont présentés sous forme de graphes dans la figure 2.3. Les noyaux gaussiens n'y sont pas présentés car leur performances sont médiocres (inférieures à 50%). Sachant que l'on souhaite retrouver les images similaires, on doit avoir un bon taux pour C_{sim} tout en ayant une bonne réussite globale, ce qui conduit au choix $C=4000$.

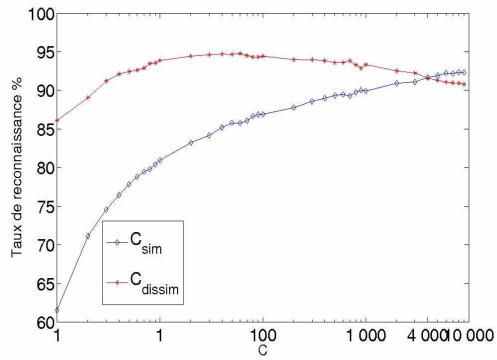
Commentaires La décision basée sur la courbe granulométrique donne de moins bons résultats que celle basée sur l'histogramme. Ce résultat semble en contradiction avec le fait que la granulométrie apporte une information spatiale qui n'est pas présente dans le cas de l'histogramme. Cependant, la courbe granulométrique ne contient pas toute l'information apportée dans l'histogramme, en particulier sur la distribution des valeurs. Les résultats montrent que cette information est plus importante que l'information extraite par la granulométrie.

2.3.3 Bilan

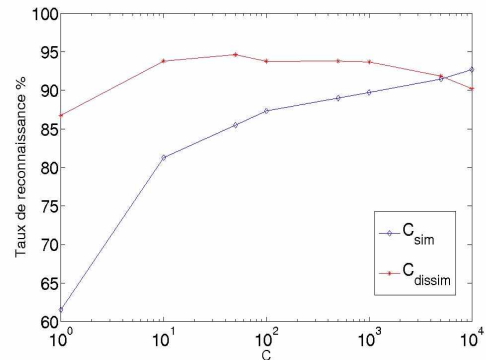
Pour avoir une information plus précise que celle fournie par une mesure globale, les décisions basées sur l'histogramme et sur les courbes granulométriques ont été étudiées. La méthode de classification qui offre les meilleurs résultats est la méthode des SVM. Une étude dans chaque cas permet de fixer le type de noyau et le paramètre C de sorte que l'efficacité soit la meilleure. La décision basée sur la courbe granulométrique est proposée pour améliorer celle basée sur l'histogramme en apportant de l'information sur la distribution spatiale des valeurs. Cependant les résultats montrent que l'information apportée par la granulométrie ne contient pas celle de l'histogramme. Pour améliorer les résultats, il faut prendre en compte toute l'information présente dans la CDL.

2.4 Classification directe sur la Carte des Dissimilarités Locales

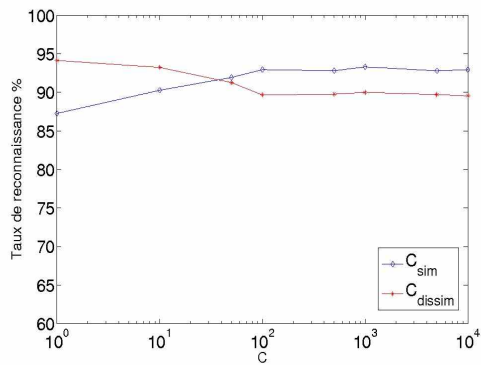
Trois types de cartes ont été testés : la différence simple entre les images $CDS(A, B) = |B - A|$, la CDL basée sur la différence simple (CDL_{DS}) et la CDL basée sur la distance de Hausdorff (CDL_{DH}).



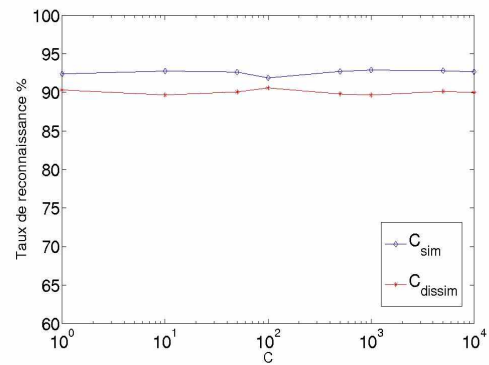
(a) Noyau linéaire



(b) Noyau polynômial de degré 1



(c) Noyau polynômial de degré 2



(d) Noyau polynômial de degré 3

FIG. 2.3 – Efficacité de la classification des courbes granulométriques basée sur les SVM en testant les noyaux linéaires et polynômiaux et pour C variant entre 1 et 10000 (en échelle semi-logarithmique). Les meilleurs résultats sont de 91% pour les deux classes pour tous les noyaux. Le moins gourmand en temps de calcul est le noyau linéaire qui obtient les meilleurs taux pour $C=4000$.

2.4.1 Méthodologie d'évaluation

Nous avons choisi deux méthodes d'évaluation :

- méthode circulaire aussi dénommée *leave-one-out* où les ensembles d'apprentissage sont les classes C_{sim} et C_{dissim} auxquelles a été retiré dans chacune un élément, le test est alors fait sur le couple d'éléments restant. Ce test est réalisé de manière circulaire pour tous les couples d'éléments possibles et le résultat est constitué de la moyenne sur l'ensemble des résultats.
- Méthode sur échantillons aléatoires : l'apprentissage se fait sur un échantillon aléatoire pour chaque classe et le test aussi (parmi les éléments qui n'ont pas servis à l'apprentissage). Pour limiter l'effet du hasard, ce test est répété plusieurs fois et une moyenne est alors faite sur les résultats obtenus.

L'intérêt de la méthode circulaire est d'être exhaustive, le processus de décision est évalué sur l'ensemble des éléments de test. Cependant, le test est effectué sur tous les éléments sauf un par classe, ce qui ne reflète pas le cas réel où l'ensemble des données n'est en général pas disponible lors de l'apprentissage. La méthode des échantillons aléatoires correspond mieux à une situation réelle. Par contre, elle a le défaut de ne pas être exhaustive et d'avoir un résultat qui peut varier avec les choix des échantillons d'apprentissage et de test. Même si une moyenne est faite sur plusieurs essais il est intéressant d'estimer l'écart-type de cette moyenne pour fixer judicieusement le nombre d'éléments sur lequel elle doit être faite.

Les deux méthodes de tests ne sont pas utilisées de la même manière : toutes les méthodes sont testées avec la méthode des échantillons aléatoires car c'est une méthode qui reflète la réalité et qui est rapide. La méthode du test circulaire nécessite plus de temps. Elle est donc utilisée pour affiner l'évaluation des performances de la CDL_{DH} qui est la méthode la plus efficace au sens de la méthode des échantillons aléatoires.

La classification utilisant la comparaison d'images basée sur la CDL_{DH} est la plus intéressante et, à ce titre, elle bénéficie d'une étude détaillée. Elle est présentée en premier afin que les détails de cette étude puissent servir pour les autres classifications.

2.4.2 Carte des Dissimilarités Locales basée sur la distance de Hausdorff (CDL_{DH})

Choix de la taille des échantillons et du nombre d'itérations

Les tailles des échantillons d'apprentissage et de test doivent être fixées en fonctions de plusieurs contraintes. D'une part le nombre de données disponibles dans chaque classe. D'autre part, l'ensemble d'apprentissage doit être aussi petit que possible tout en conservant les performances de classification et l'ensemble de test doit être assez grand pour fournir des résultats fiables et doit être limité pour permettre un temps de calcul raisonnable.

D'un point de vue pratique, la classe C_{sim} comporte 125 CDL et la classe C_{dissim} en comporte 2153. Comme c'est indiqué dans le paragraphe 2.1.2, les ensembles d'apprentissage et de test doivent être disjoints. Pour pouvoir faire le test sur un nombre supérieur de CDL à celui de l'apprentissage, nous avons choisi les ensembles d'apprentissage de chaque classe à 50 éléments. Les performances des SVM ont été testées pour des ensembles d'apprentissage de taille allant de 10 à 60 éléments pour C_{sim} et de 20 à 120 pour C_{dissim} . La différence entre les tailles pour les deux classes provient simplement de la différence de taille entre les deux classes (125 pour C_{sim} et 2153 pour C_{dissim}). Pour tester la taille des ensembles de test, nous avons estimé la variance du taux de réussite (par la méthode des échantillons aléatoires) sur plusieurs essais. Les résultats de ces tests sont présentés fig.2.4. Ils montrent que les résultats sont mauvais pour des échantillons d'apprentissage trop petits (10 à 20 pour C_{sim} et 20 à 40 pour C_{dissim}) et que l'écart-type est assez important (surtout pour C_{sim} : $\pm 1,8\%$). Par contre, à partir de 30 éléments et en particulier pour 50 pour C_{sim} (le double pour C_{dissim}), les résultats sont bons et avec un faible écart-type. De plus, ils s'améliorent peu pour 60 éléments par rapport à 50 éléments. Le nombre d'éléments dans

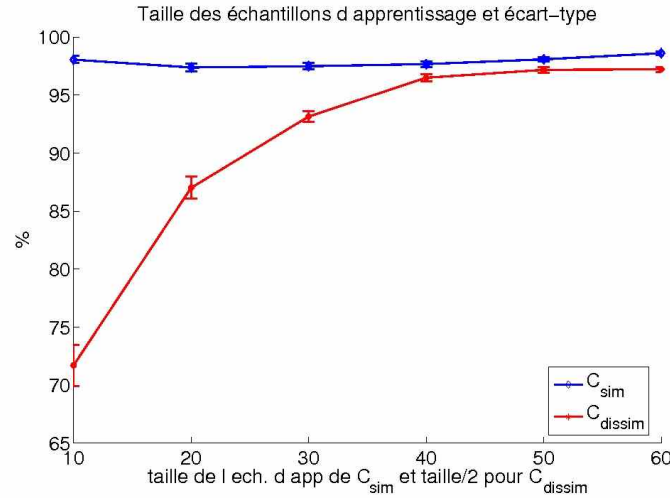


FIG. 2.4 – Taux de réussite dans les deux classes C_{sim} et C_{dissim} en fonction de la taille des échantillons d'apprentissage : entre 10 et 60 éléments pour C_{sim} et entre 20 et 120 pour C_{dissim} . L'écart-type est calculé sur 100 itérations.

C_{sim} ne permet pas de tester un nombre plus grand, sinon l'échantillon de test aurait une taille trop faible. Les échantillons comportant 50 éléments pour C_{sim} et 100 éléments pour C_{dissim} sont donc pleinement justifiés.

Influence des paramètres des SVM

Le coefficient C règle le compromis entre la marge possible entre les exemples et le nombre d'erreurs admissibles. Il est fixé par l'utilisateur.

Nous avons donc testé une série de valeurs pour C afin de déterminer celle qui donne les meilleurs résultats. Les résultats dépendent d'un autre paramètre qui est le type de noyau. Nous avons donc combiné le test de C avec celui du type de noyau. Nous avons testé trois types de noyaux classiques déjà présentés ci-dessus : les noyaux linéaires, polynômiaux et gaussiens. Les résultats sont présentés dans le tableau 2.2. Les résultats montrent que le noyau linéaire est un peu plus efficace que le

noyau	$C = 1$	$C = 10$	$C = 100$	$C = 1000$
linéaire	96,3%	97%	96%	99,6%
degré 1	96%	95,5%	94,3%	97%
degré 2	96,5%	96%	95,7%	99,2%
degré 3	96,6%	94%	95,3%	97%
degré 4	91,7%	94,1%	91,2%	95%
gaussien, $\sigma = 2$	80,6%	80%	80%	79,4%
gaussien, $\sigma = 1$	81,2%	82,5%	81,8%	81,8%
gaussien, $\sigma = 0,5$	81,2%	81,2%	90,6%	80,6%
gaussien, $\sigma = 0,01$	95,5%	94,3%	94,3%	94,5%

TAB. 2.2 – CDL_{DH} : taux de reconnaissance lorsque le coefficient C et le type de noyau varient.

noyau polynômial et que le noyau gaussien. De plus, c'est pour ce noyau que le temps de calcul est le plus court. La valeur de C pour laquelle l'efficacité est la plus grande est $C = 4000$.

résolution	512×512	256×256	128×128	64×64	32×32	16×16	8×8
pour C_{sim}	99,6%	99,6%	99,4%	98,6%	93,8%	86%	82%
pour C_{dissim}	94%	94%	93,7%	91,1%	86,5%	80,5%	67,8%

TAB. 2.3 – CDL_{DH} : taux de reconnaissance lorsque les CDL sont calculées après une AMR. Le fait que les résultats restent bons jusqu'à la taille 32×32 est notable car les images sont alors 16 fois plus petites que les images initiales.

2.4.3 Résultats en combinant la Carte des Dissimilarités Locales avec l'analyse multirésolution



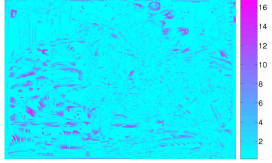
L'utilisation de l'AMR a ici deux buts : d'une part offrir un gain de temps en permettant de réaliser les calculs sur des images de taille réduites et d'autre part d'offrir une décomposition contrôlée des structures présentes dans l'image. Concernant le choix de l'AMR morphologique avec le filtre de la médiane, nous renvoyons le lecteur à la section 2.5 (p. 36). Ce choix permet de conserver les principales structures de l'image pour des échelles assez grossières, ce qui permet le calcul de la CDL pour des résolutions grossières. Deux évaluations ont été effectuées ici, la première concerne la classification basée sur la CDL calculée à partir des approximations à plusieurs résolutions. La deuxième a pour objet la classification basée sur les images de détails.

Carte des Dissimilarités Locales calculée sur les approximations



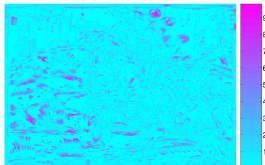


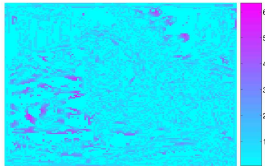


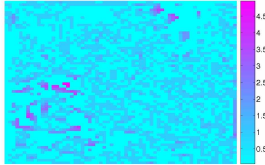


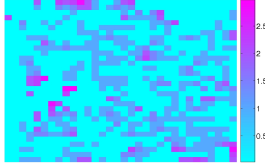


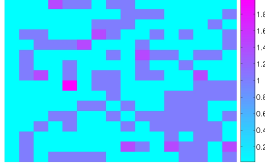
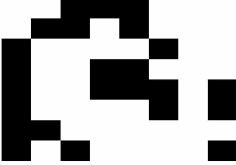
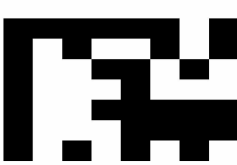
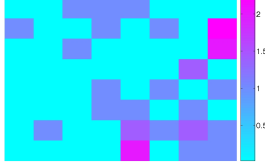
L'évaluation porte sur l'influence de la résolution des images sur l'efficacité de la classification : une AMR est effectuée sur les images binaires et les CDL sont calculées aux différentes échelles entre les approximations. Les résultats sont reportés dans le tableau 2.3. Nous constatons que les résultats sont très bons de la résolution 512×512 à la résolution 128×128 puis décroissent tout en restant corrects jusqu'à la résolution 32×32 .

Un compromis doit être fait entre l'efficacité et le temps de calcul : avoir une bonne efficacité sur des images aussi petites que possible. On peut considérer que la similarité se base sur les structures principales des images. En conséquence, les résultats restent bons tant que les structures discriminantes sont présentes dans l'approximation. Dans le cas des impressions anciennes, les structures discriminantes sont parfaitement conservées jusqu'à la résolution 128×128 . C'est cette résolution qui est choisie pour effectuer les tests. Le tableau 2.4 permet de visualiser la conservation des structures au travers des résolutions sur deux exemples. Les scènes sont bien interprétables à la résolution 128×128 . Par contre, elles ne le sont pratiquement plus aux échelles 5 et 6 (16×16 et 8×8) et les résultats le reflètent.

TAB. 2.4: Exemple d'approximations pour 2 impressions par l'AMR de la médiane morphologique et leur CDL aux différentes résolutions.

	Impression 1	Impression 2	CDL-HD
échelle 0 : 512×512			
suite à la page suivante			

TAB. 2.4: suite

1 : 256×256			
2 : 128×128			
3 : 64×64			
4 : 32×32			
5 : 16×16			
6 : 8×8			

TAB. 2.4: L'AMR de la médiane morphologique appliquée aux images 1 et 2 montre que les structures sont conservées aux différentes échelles dans les approximations jusqu'à un certain point.

Carte des Dissimilarités Locales à partir des images de détails

Il y a trois images de détails à chaque résolution. Une possibilité pour effectuer une classification basée sur ces images est de calculer trois CDL en comparant les trois images de détails de chaque image. Cependant, il faut alors prendre une décision à partir de trois CDL, ce qui n'est pas trivial. Une autre possibilité est de les recombinaison mais *sans* l'approximation.

Ce qui revient à faire une synthèse à partir des trois images de détails et une approximation nulle.

Les détails de trois échelles ont été recomposés afin d'avoir une image moins morcellée. Autrement dit, les détails de trois échelles successives ont été recombinaés successivement avec une approximation mise à zéro. L'image finale est de taille 128×128 . Deux exemples sont fournis dans la figure 2.5 Les résultats sont reportés dans le tableau 2.5. Ces résultats sont



FIG. 2.5 – Exemple d'images de détail recomposés avec une approximation nulle : les détails de trois échelles successives ont été recomposés.

retrouvé dans	C_{sim}	C_{dissim}
provenant de C_{sim}	93%	7%
provenant de C_{dissim}	11%	89%

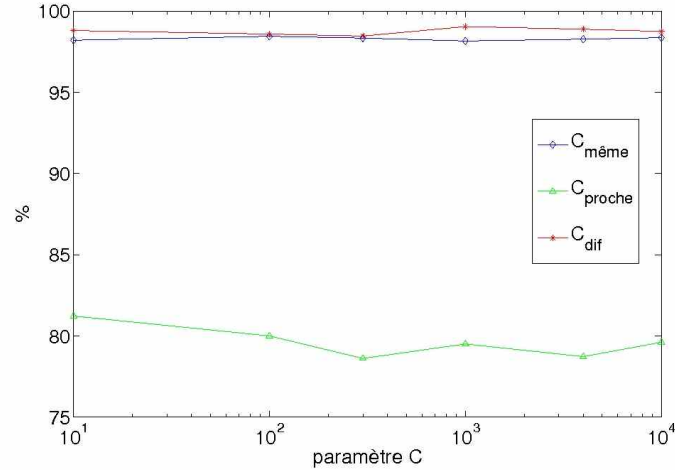
TAB. 2.5 – CDL_{DH} : taux de reconnaissance de la classification à partir des images de détails recomposés de taille 128×128 .

moins bons que dans le cas où la carte est calculée à partir des approximations de 5% environ. Une raison est que la similarité pour cette base repose sur les traits principaux qui sont en partie absents des images de détails. De plus, les images de détail sont composées de petites structures et les distances entre de petites structures dans la CDL ont un peu moins de sens.

2.4.4 Classification en trois groupes

Nous avons donc testé la classification en trois classes avec cette méthode. Les trois classes (déjà présentées dans la sous-section 1.3.1) sont les suivantes : C_{sim} est scindée en deux classes qui sont la classe des impressions provenant du même tampon $C_{même}$ et celle des impressions illustrant la même scène mais provenant de tampons différents C_{proche} . C_{dissim} reste identique. Les résultats de la CDL_{DH} sont les meilleurs pour la classification en deux classes par rapport à ceux des autres méthodes. Comme la classification en trois classes n'est en fait qu'une séparation en deux sous-classes de la classe C_{sim} , il est logique de s'intéresser aux résultats de la CDL_{DH} plutôt qu'aux autres. Le test est effectué à la résolution 128×128 et les CDL_{DH} sont calculées à partir des approximations. La classification est effectuée selon la méthode des échantillons aléatoires. Le résultat est présenté dans le tableau 2.6. Le détail des résultats pour les différentes valeurs du paramètre C du SVM est présenté dans le graphe 2.6. Les résultats concernant la classification en trois classes est bon. C'est d'autant plus remarquable qu'il est souvent délicat de distinguer à l'œil nu les impressions provenant du même tampon de celles illustrant la même scène mais ne provenant pas du même tampon.

	$C_{\text{retrouvé dans même}}$	retrouvé dans C_{proche}	retrouvé dans C_{dif}
$C_{\text{même}}$	98,3%	0,3%	1,4%
C_{proche}	16,9%	80,5%	2,6%
C_{dif}	1,3%	0%	98,7%

TAB. 2.6 – CDL_{DH} : résultat du test de la classification en trois classes à la résolution 128×128 .FIG. 2.6 – Taux de classification de la CDL_{DH} pour les 3 classes en fonction du paramètre C du SVM. Les meilleurs taux sont trouvés pour $C=100$.

2.4.5 Carte de Différence Simple (CDS)

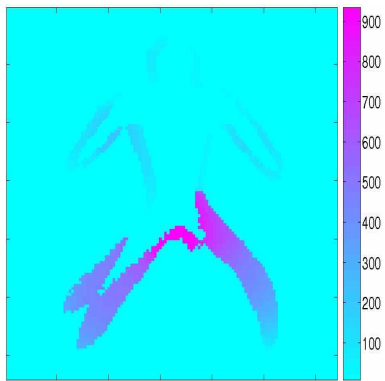
La carte de différence simple est obtenue en calculant la différence absolue pixel à pixel des deux images binaires $CDS(A, B) = |B - A|$. Comme on prend la valeur absolue de la différence, la CDS est symétrique et n'est nulle que si les deux images sont égales. Par contre, la CDS est binaire et ne permet donc pas de mesurer une grandeur de dissimilarité, mais simplement d'en indiquer spatialement la présence. La comparaison des performances de la CDS avec celles de la CDL permet d'évaluer l'apport de la mesure locale des dissimilarités par rapport à la simple mesure pixel à pixel.

2.4.6 Carte des Dissimilarités Locales basée sur la différence simple $CDL_{DifSimple}$

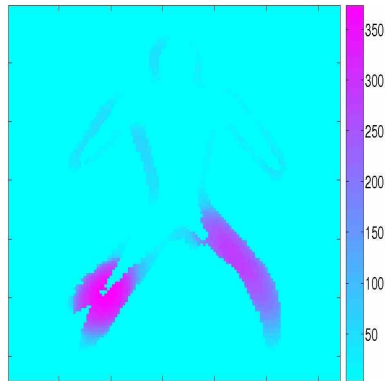
La CDL_{DS} est la CDL obtenue en utilisant la mesure basée sur la différence simple. Le principe de la CDL générale tel qu'il est décrit dans la sous-section 2.4.3 (p. 2.4.3). La $CDL_{difSimple}$ est une CDL paramétrique dont la mesure locale $\Phi_{difSimple}$ est détaillée dans le paragraphe exemple applicatif de cette sous-section. Le principe est rappelé ici : la fenêtre croît tant que la proportion de pixel ayant une valeur différente dans les deux images est supérieur au critère. Ainsi, plus le paramètre est grand, plus le critère est strict quand à la croissance de la fenêtre. Quand il est égale à 1, il faut que l'ensemble des points de la fenêtre soient différents dans les deux images pour que la fenêtre croisse. Une illustration de la différence entre deux images avec différentes valeurs pour le paramètre est donné fig. 2.7.



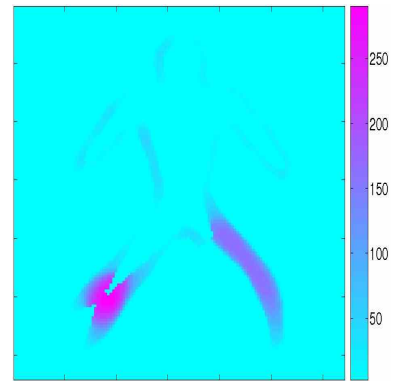
(a) $CDL_{difSimple}$ pour le paramètre = 0,8. (b) $CDL_{difSimple}$ pour le paramètre = 1.



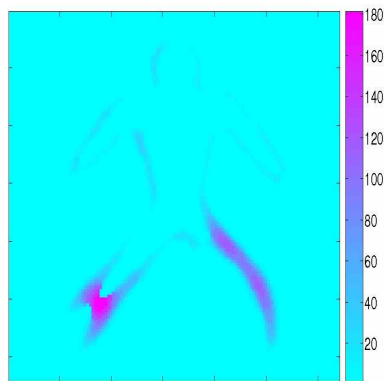
(c) $CDL_{difSimple}$ pour le paramètre = 0,2.



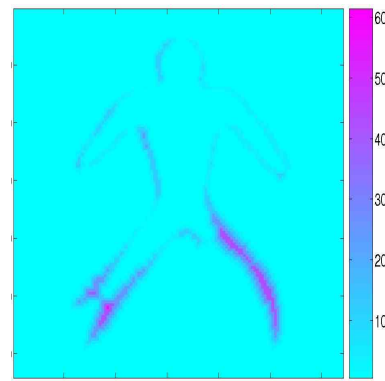
(d) $CDL_{difSimple}$ pour le paramètre = 0,4.



(e) $CDL_{difSimple}$ pour le paramètre = 0,6.



(f) $CDL_{difSimple}$ pour le paramètre = 0,8.



(g) $CDL_{difSimple}$ pour le paramètre = 1.

FIG. 2.7 – Représentation de la $CDL_{difSimple}$ pour différentes valeurs du paramètre d'arrêt. Différentes dissimilarités sont valorisées selon la valeur du paramètre. L'échelle des valeurs est la même que pour la $CDL_{difSimple}$: le bleu représente les faibles valeurs de mesure de dissimilarité et le rose les hautes valeurs.

Résultats

Les tests sont réalisés sur les approximations 128×128 en faisant varier le paramètre entre 0 et 1 pour le calcul de la $CDL_{difSimple}$ puis la classification est faite avec un SVM avec la méthode des échantillons aléatoire. Les résultats sont présentés dans le graphe 2.8 Les résultats sont corrects, sauf pour la valeur 0,5 du paramètre où le taux de reconnaissance

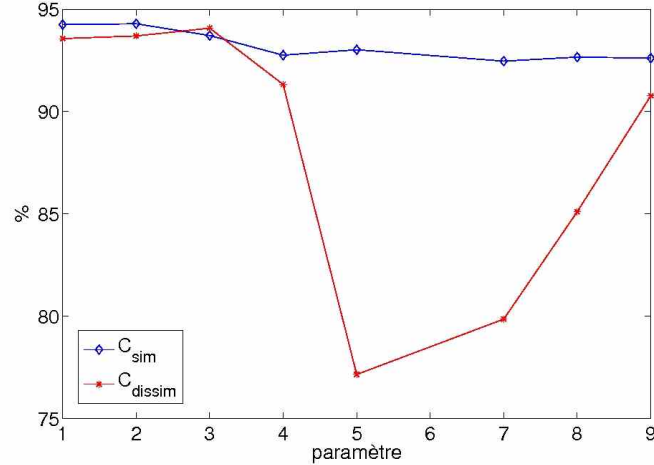


FIG. 2.8 – Taux de réussite pour la classification à partir de la $CDL_{difSimple}$ dans les deux classes C_{sim} et C_{dissim} en faisant varier le paramètre d'arrêt entre 0 et 1.

pour C_{dissim} est proche de 75%. Les meilleurs taux sont d'environ 94%, ce qui les place mieux que les résultats des mesures globales et moins bien que la CDL basée sur la distance de Hausdorff. Il est notable que le taux de reconnaissance reste bon pour C_{sim} pour toutes les valeurs du paramètre, alors que celui de C_{dissim} varie. Une explication pour cela est dans le fait que la différence simple rend mieux compte des faibles dissimilarités que des fortes.

2.4.7 Comparatif

Le but premier de la mise en œuvre de la CDL était d'offrir une alternative plus précise aux mesures de similarité globales pour les images binaires parfois complexes. Pour évaluer le gain d'efficacité qu'apporte le processus de classification reposant sur la CDL, nous faisons un comparatif ici entre les résultats obtenus sur la base d'impressions anciennes avec la CDL basée sur la DH et ceux obtenus avec :

- la DH globale et ses variantes la DH partielle, la DH modifiée,
- la différence simple image à image $|B - A|$ et la CDL basée sur la différence simple (où la mesure locale n'est pas la DH mais la différence simple) notée $CDL_{difSimple}$,
- le Vecteur des Valeurs Classées de la CDL noté VVC : il contient l'ensemble des valeurs de la CDL classées par ordre croissant (pour tester l'importance de l'information spatiale).

Méthode de classification pour les distances globales

Cette méthode repose aussi sur une décision avec apprentissage et la méthode des échantillons aléatoires (cf sous section 2.4.2) est utilisée ici. Par contre, l'apprentissage ne se fait pas avec les SVM mais en calculant les distributions empiriques de chaque classe : à partir des échantillons d'apprentissage pour les deux classes sont évaluées les deux distributions empiriques

Taux de réussite	CDL_{DH}	$CDL_{difSimple}$	Carte de Dif Simple	VVC	DH	DH partielle	DH modifiée
pour C_{sim}	98%	92%	66%	94%	60%	83%	77%
pour C_{dissim}	97%	94%	78%	88%	75%	81%	83%

TAB. 2.7 – Comparaison des différentes méthodes avec celles basées sur la CDL. Les méthodes sont la CDL basée sur la DH (CDL_{DH}), la CDL basée sur la différence simple ($CDL_{difSimple}$), la DS entre les images (CDS), le Vecteur des Valeurs de la CDL Classées (VVC), la DH globale, la DH partielle et la DH modifiée. Les résultats montrent d’une part que l’information apportée dans la CDL fournit une meilleure efficacité que les mesures globales et d’autre part que la mesure de dissimilarité est mieux faite dans la CDL avec la DH qu’avec la différence simple.

(en utilisant une fenêtre de Parzen [Parzen, 1962]). Puis la classification de l’échantillon de test est faite par la méthode du maximum de vraisemblance sur les deux distributions empiriques obtenues.

La base de comparaison est la base d’impressions anciennes numérisées et la résolution choisie pour le test est 128×128 . La taille des échantillons d’apprentissage et de test est la même que celle employée auparavant :

- 50 éléments pour chaque classe en apprentissage ,
- 75 éléments pour C_{sim} et 200 pour C_{dissim} en test.

Les résultats sont moyennés sur 100 itérations. Ils sont présentés dans le tableau 2.7.

La DH partielle dépend d’un paramètre $p = \frac{K}{N}$ qui représente la proportion de valeurs prises en compte dans le calcul du maximum (cf 1.3.3, p. 57). Le choix de ce paramètre a été fait en testant l’efficacité de la classification pour une série de valeurs comprise entre $p = 0,5$ et $p = 1$. Les résultats pour cette série de valeurs sont représentés fig. 2.9.

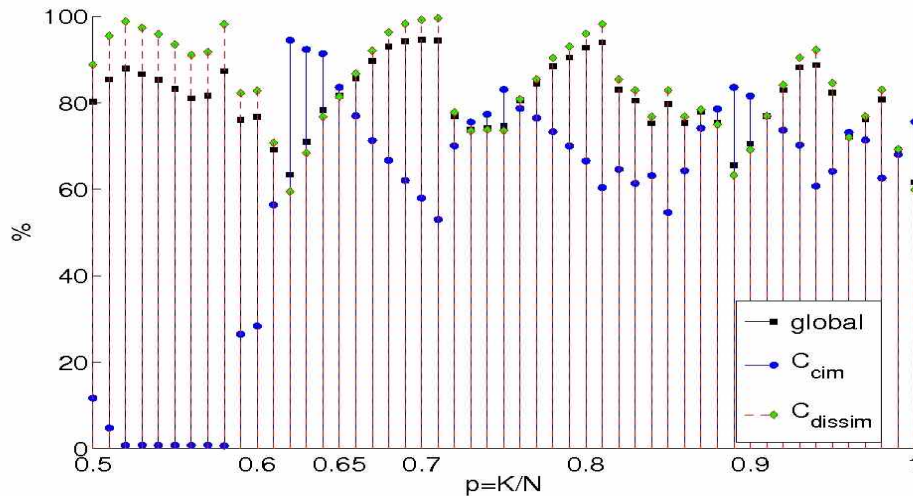


FIG. 2.9 – DH partielle : efficacité de la classification globale et pour les deux classes avec un paramètre p variant entre 0,5 et 1. Les meilleurs taux de classification global qui soit aussi bon pour C_{sim} est trouvé pour $p = 0,65$. Les valeurs correspondantes : 83% pour C_{sim} et 81% pour C_{dissim} sont reportées dans le tableau de comparaison 2.7.

Processus de classification pour les cartes et du vecteur de valeurs classées

Le processus est le même que pour la CDL : l’apprentissage et le test pour les SVM se fait directement à partir des cartes selon la méthode des échantillons aléatoires. Les résultats sont présentés dans le tableau 2.7.

Analyse

Plusieurs points peuvent être déduits de la comparaison. Nous les développons ci-dessous.

Concernant les mesures globales Les résultats de la CDL sont meilleurs que ceux obtenus avec les mesures globales, ce qui était le but recherché. Ceci a deux raisons :

- la première est que la CDL donne accès à toutes les valeurs locales de dissimilarité alors que la DH ne donne que le maximum, la DH partielle une valeur intermédiaire ; l'apprentissage permet de choisir la plus pertinente,
- la deuxième est que ces valeurs sont organisées les unes par rapport aux autres et que leur répartition spatiale apporte là aussi de l'information. D'ailleurs ceci est confirmé par les bons résultats de la CDL DH lorsqu'elle est combinée à l'analyse multirésolution : sur les approximations 16×16 on obtient encore des résultats meilleurs qu'avec les mesures globales (cf tab. 2.3) malgré la dégradation inévitable de la qualité des images à cette échelle.

Cependant, pour mieux mesurer l'influence des paramètres tels que la répartition spatiale de l'information ou la qualité de la mesure de la dissimilarité locale, la comparaison aux cartes et aux valeurs classées est plus appropriée.

Concernant les cartes et le vecteur de valeurs classées La comparaison entre la CDL_{DH} et la $CDL_{difSimple}$ montre que la DH est plus à même de mesurer la dissimilarité locale que la différence simple. Le fait de mesurer une distance entre ensemble de points plutôt que point à point apporte une information discriminante. Ce point est illustré fig. 2.10 où sur un exemple, la $CDL_{difSimple}$ rend moins bien compte des dissimilarités que la CDL_{DH} . L'apport de l'information spatiale est évaluée par la comparaison entre la CDL_{DH} et le VVC. En effet, la $CDL_{difSimple}$ et le vecteur de valeurs classées contiennent exactement les mêmes valeurs, mais les relations spatiales entre les points ont été perdues puisqu'elles sont classées par ordre croissant dans le VVC. Les résultats montrent que l'efficacité de la décision se dégrade pour les deux classes C_{sim} (-4%) et C_{dissim} (-9%) et que cette dégradation affecte particulièrement C_{dissim} , autrement dit, l'information apportée par la distribution spatiale des mesures de dissimilarités est particulièrement importante pour caractériser les images dissimilaires.

Test circulaire

Le test circulaire est coûteux en temps de calcul, il a donc servi à tester la méthode donnant les meilleurs résultats, la CDL_{DH} . La méthode consiste à retirer un élément par classe, de faire l'apprentissage sur l'ensemble des éléments qui restent et le test sur le couple qui n'a pas servi dans l'apprentissage. Le test est effectué à la résolution 128×128 et les CDL_{DH} sont calculées à partir des approximations. Le résultat est présenté dans le tableau 2.8. Les résultats sont très bons, ce qui

taux de reconnaissance moyen sur le couple test	2/2	1/2	0/2
	99,2%	0,8%	0%

TAB. 2.8 – CDL_{DH} : résultat du test circulaire à la résolution 128×128 .

provient sans doute du fait que l'apprentissage est fait sur un nombre suffisant de cartes pour représenter les classes.

2.4.8 Robustesse

Nous avons testé la robustesse à deux types de bruit qui correspondent à une réalité pour les impressions anciennes et qui sont la tache d'encre (une partie de l'impression est noire) et l'effacement (une partie de l'impression est blanche). La procédure a été la même dans les deux cas : l'apprentissage a été fait sur la base d'impression ancienne sans bruit, à l'échelle

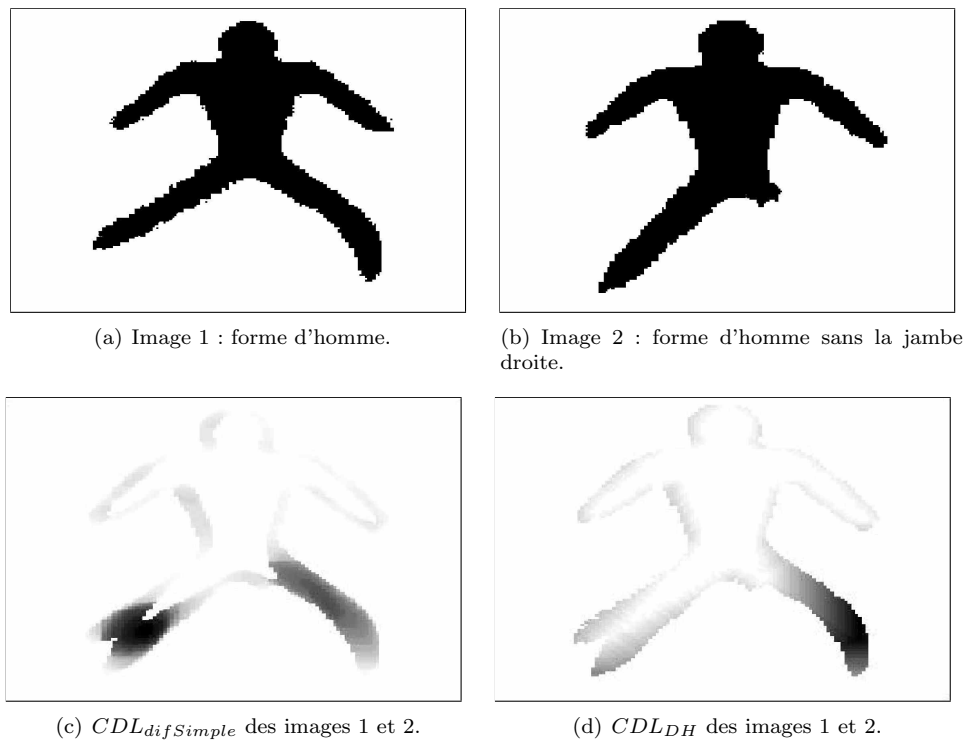
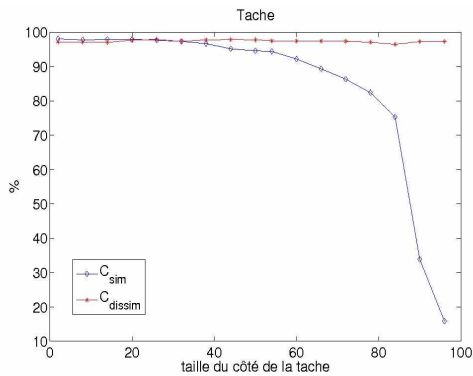


FIG. 2.10 – Illustration de la capacité à mesurer les dissimilarités locales. Fig. 2.10(c) : bien que les deux jambes gauches soient proches, leur différence de position font que la $CDL_{difSimple}$ de ces deux images obtient sa valeur maximale à cet endroit. De plus, les valeurs mesurées dans la $CDL_{difSimple}$ au niveau de la jambe droite sont plus basses bien que la dissimilarité soit plus importante. Fig. 2.10(d) : la CDL_{DH} est plus proche de l'intuition avec une valeur plus grande au niveau de la jambe droite.

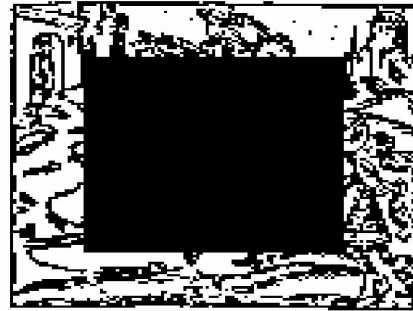
2 (résolution 128×128), et le test a été fait sur des CDL résultant de la comparaison d'une image sans bruit et d'une image bruitée. Dans ce cas, ce qui est testé, c'est le pouvoir du classificateur à reconnaître comme similaires (ou dissimilaires) deux images qui le sont initialement mais dont une est tachée ou en partie effacée.

Les taches ou effacements ont été fait de la même manière : une partie centrale de l'image carrée de côté compris entre 5 et 80 est noircie ou blanchie.

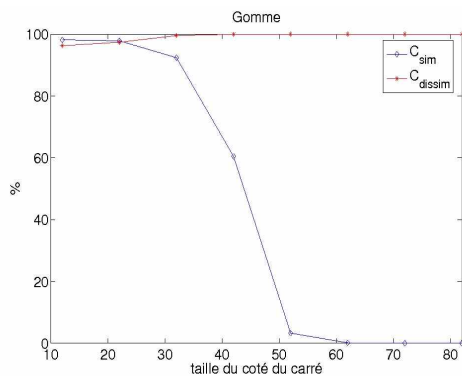
Les résultats sont présentés dans la figure 2.11 sous formes de graphiques, ainsi que des illustrations correspondant aux tailles critiques des zones bruitées. Les résultats montrent que le processus est très robuste aux taches noires et relativement



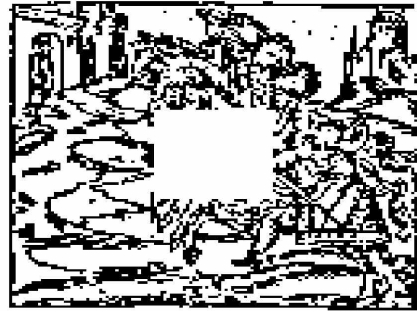
(a) Efficacité de la classification selon la taille de la tache.



(b) Exemple de tache (taille 80×80) pour laquelle l'efficacité est encore au dessus de 80% pour les deux classes : 82% pour C_{sim} et 97% pour C_{dissim} . Au-delà, l'efficacité se dégrade.



(c) Efficacité de la classification selon la taille de la partie effacée.



(d) Image gommée pour une tache de taille 32×32 . Le taux de réussite est convenable pour cette taille : 92% pour C_{sim} et 99% pour C_{dissim} . Au-delà, l'efficacité se dégrade.

FIG. 2.11 – Robustesse aux tâches et aux effacements : la méthode est plus robuste aux taches qu'aux effacements. Ceci étant, la robustesse de la méthode dans le deux cas est bonne.

robuste aux taches blanches. Une raison pour cela est que le noir est de l'information, autrement dit, ce qui est comparé dans la CDL, alors que le blanc est le fond. Une tache noire est traitée comme de l'information supplémentaire alors qu'une tache blanche l'est comme du fond. Si la zone de fond est trop importante, les images sont considérées comme dissimilaires.

Récapitulatif

Nous présentons ici un tableau récapitulatif des résultats pour toutes les méthodes de comparaison testées.

	C_{sim}	C_{dissim}
Méthodes globales		
DH	60%	75%
DH partielle	83%	81%
DH modifiée	77%	83%
Méthodes vectorielles		
Histogramme :		
histogramme moyen	84%	95%
SVM	91%	94%
Granulométrie	91%	91%
Vecteur des Valeurs de la CDL Classées	94%	88%
Méthodes 2D		
Différence Simple	66%	78%
$CDL_{diffSimple}$	92%	94%
CDL_{DH} :		
512 × 512	99,6%	94%
256 × 256	99,6%	94%
128 × 128	99,4%	93,7%
64 × 64	98,6%	91,1%
32 × 32	93,8%	86,5%
16 × 16	86%	80,5%
8 × 8	82%	67,8%
3 classes	98%	80%
Détails reconstitués	93%	89%
<i>Robustesse</i>		
Tache 80 × 80	82%	97%
Effacement 32 × 32	92%	99%

TAB. 2.9 – Récapitulatif des résultats pour toutes les méthodes de comparaison testées.

Ce tableau montre que le taux de reconnaissance est relié la dimension des mesures de comparaison sur cette base d'impressions anciennes : les mesures globale testées ont une taux de reconnaissance au environ de 80%, les mesures vectorielles ont leurs meilleurs taux entre 90% et 94% et les mesures sous forme de cartes ont leurs meilleurs taux entre 94% et 99%. Cela va dans le sens de notre projet initial : mesurer de manière plus détaillée les dissimilarités entre les images et mesurer aussi leur répartition spatiale permet de mieux les classer. Il montre aussi que la mesure locale joue un rôle important : la CDL basée sur la distance de Hausdorff offre de bons résultats alors que la CDL basée sur la différence simple a des résultats à peine meilleurs que ceux des mesures vectorielles. La mesure locale dérivée de la DH permet donc de bien rendre compte des dissimilarités locales. Enfin la combinaison de l'analyse multirésolution et de la classification permet de cerner quelles type structures interviennent dans la détermination de la similarité, en l'occurrence, ici, les traits principaux des images.

Autres applications

3.1 Classification sur une base de formes

Le processus de comparaison et de décision a été testé sur une base de formes comportant 99 formes provenant de [Sebastian *et al.*, 2001]. Un exemple de formes de cette base est présentée figure 3.1. Pour l'évaluation, nous avons effectué

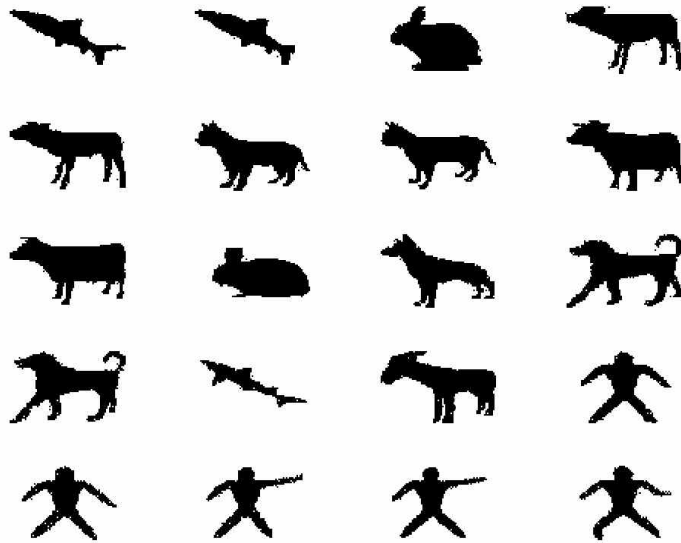


FIG. 3.1 – Quelques exemples de formes de la base.

nous mêmes le choix des classes. Comme nous n'avons pas intégré de recalage dans le processus, et que la comparaison avec la CDL est locale, nous avons mis dans la même classe des formes similaires et à peu près recalées. La méthode appliquée a été la même que pour la base d'impressions anciennes : un apprentissage de 50 éléments dans chaque classe et un test sur 50 et 100 éléments dans C_{sim} et C_{dissim} respectivement. Une moyenne a été faite sur cent itérations. Les résultats du test sont résumés dans le tableau 3.1. La simplicité des formes permet d'employer des méthodes plus spécialisées. Ainsi dans [Sebastian *et al.*, 2001], les auteurs proposent un méthode adaptée à ce type d'objets reposant sur le calcul de ESG (Editing Shock Graph, qui est une sorte de squelette). Les résultats obtenus sont proches de 100% (entre 98% et 100%). Bien que

taux de réussite	retrouvé dans C_{sim}	retrouvé dans C_{dissim}
provenant de C_{sim}	97%	3%
provenant de C_{dissim}	7%	93%

TAB. 3.1 – Résultat de la méthode sur la bases de formes

moins bons, les résultats de la CDL sont acceptables (un peu moins de 95% en moyenne), ce qui montre que la méthode peut être adaptée à différents types d’images.

L’application suivante porte sur la reconnaissance de visage.

3.2 Classification sur une base de visages

La reconnaissance de visages est un sujet très étudié actuellement. Comme les images de visages ne sont pas des images binaires mais des images en niveaux de gris, elles ne rentrent pas exactement dans le cadre de l’application de la CDL. Cependant, il est possible d’en extraire les contours, l’image résultante est alors une image binaire sur laquelle on peut utiliser la méthode CDL SVM.

3.2.1 Description pratique

La base sur laquelle a été testé la méthode est la base d’ATT qui comporte les photos de 40 personnes dont le visages a été photographié de 10 manières différentes. L’extraction de contours a été faite selon la méthode de Canny. Les images résultantes sont de taille 112×92 . Une illustration comportant des visages de la bases, les images de contours et les CDL de ces images de contours est donnée fig. 3.2

Les échantillons d’apprentissages sont constitués de 300 CDL (pour chaque classes) choisies aléatoirement parmi 900 CDL pour C_{sim} et 19000 pour C_{dissim} concernant 20 personnes uniquement et le test sur autant de CDL provenant des 20 autres personnes. Les résultats de la classification sont présentés dans le tableau 3.2.

taux de réussite	retrouvé dans C_{sim}	retrouvé dans C_{dissim}
provenant de C_{sim}	88%	12%
provenant de C_{dissim}	30%	70%

TAB. 3.2 – Résultat de la méthode sur la bases de visages

3.2.2 Analyse

Les résultats ne sont pas aussi bons sur cette base de visages que sur les précédentes. Par ailleurs, les méthodes spécialisées donnent des résultats sur cette base compris entre 90% et 98% [Liu and Wechsler, 2003]. Il y a deux raisons possibles à cela : d’une part l’extraction de contours par la méthode de Canny n’est peut-être pas la plus adaptée (on peut penser par exemple à une binarisation ne reposant pas sur les contours), d’autre part un prétraitement est peut-être nécessaire pour recalibrer les images (par rapport à la position des yeux et de la bouche par exemple). Compte tenu du fait que ces opérations n’ont pas été faites car cela nécessiterait une étude approfondie en soi, les résultats, avec une moyenne de 80% de réussite, sont encourageants.

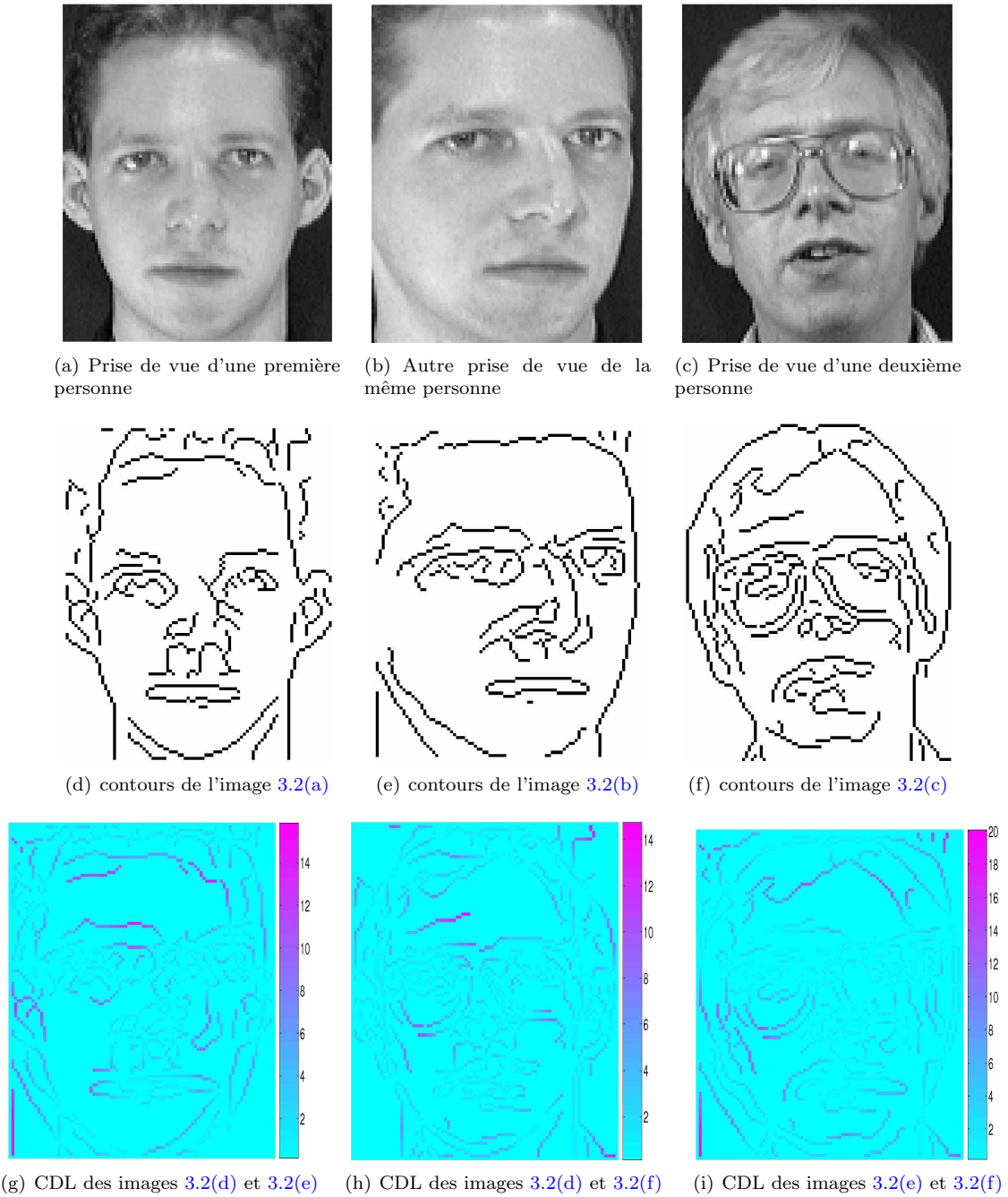


FIG. 3.2 – Exemples de visages de la base ATT, de leur contours par extraction de contours de Canny et de leurs CDL. Les CDL ne sont pas très lisibles car l'extraction de contours n'est pas très bonne et que les visages ne sont pas recalés.

3.3 Mesure globale tirant parti de la Carte des dissimilarités Locales dans le cas de la DH

Nous présentons dans cette section une mesure globale dont le calcul repose sur la CDL dans le cas de la DH. En effet dans ce cas, le calcul de la CDL est très rapide, ce qui permet de l'utiliser pour calculer une valeur globale.

La DH globale entre deux images peut être obtenue en calculant le maximum sur leur CDL (cf prop 3.3.1, p. 79). Elle peut aussi être calculée directement sans utiliser la CDL. Cependant, la CDL permet de visualiser un des défauts de la DH globale : ce peut être des valeurs isolées qui donnent la valeur à la mesure globale. Ce qui fait l'efficacité de la CDL réside en partie dans la connaissance de la répartition spatiale des valeurs de mesure de dissimilarité. Notre idée ici est d'exploiter cette répartition spatiale en appliquant un filtrage à la CDL avant d'en prendre le maximum (pour retirer les valeurs isolées dans la CDL par exemple). Cela revient à mesurer une DH filtrée (abrégié en DHF). Le filtrage dépend bien entendu du type d'image traité. Pour retirer les valeurs isolées, il est possible d'effectuer le filtrage avec un filtre de la médiane 3×3 . Si les valeurs hautes (susceptible d'être maximales) sont regroupées dans la CDL, le filtre de la médiane ne va pas en modifier la valeur maximale. Par contre, si elles sont ponctuelles (2 ou 3 sur une CDL grande) alors elles vont être lissées. Une illustration en est donnée fig. 3.3.

D'autres filtres sont envisageables : érosion, médiane, moyenne etc... L'exploitation de l'information spatiale de la CDL offre un intérêt théorique à la DH Filtrée par rapport aux variantes de la DH classiques :

- Alors que la DH partielle classe toutes les valeurs et en prend une proportion fixe $p = K/N$, la DH Filtrée retire du maximum une partie des valeurs qui dépend de l'information spatiale de la CDL. Cette proportion n'est donc pas fixe mais adaptée à chaque CDL.
- La DH Modifiée prend une moyenne globale sur les valeurs des distances aux ensembles. Mais cette moyenne est faite sur de nombreuses valeurs faibles, qui peuvent « noyer » l'information. Avec la DH Filtrée, en filtrant avec le filtre de la moyenne, il est possible de faire une moyenne locale et d'en prendre le maximum.
- Dans le cas de la DH Pondérée, les valeurs des distances aux ensembles sont pondérées en fonction de leur localisation. Il est possible d'appliquer un masque de pondération plus fin en localisant dans la CDL les zones d'intérêt.

Ainsi, cette section a présenté une mesure de dissimilarité globale reposant sur la CDL dans le cas de la DH. La rapidité du calcul de la CDL dans ce cas et la possibilité de calculer la DH globale à partir de la CDL sont les deux principales raisons pour la mise en œuvre de cette famille de mesure DH Filtrée qui tirent partie de l'information spatiale contenue dans la CDL.

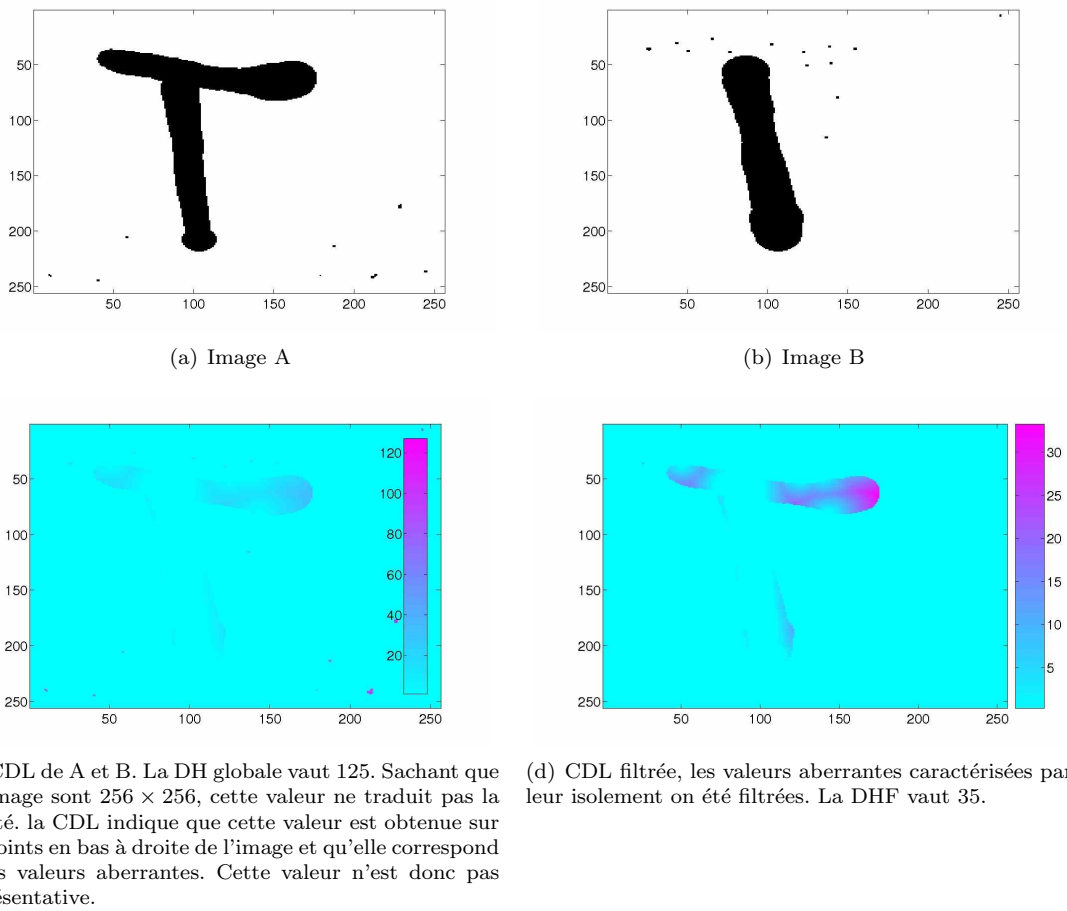


FIG. 3.3 – Illustration pour la DH Filtrée : la CDL des images 3.3(a) et 3.3(b) présentée en 3.3(c) montre que la valeur de la DH globale est due à des valeurs aberrantes. Le calcul du maximum à partir de la CDL filtrée (en 3.3(d)) permet de s'affranchir des valeurs aberrantes et d'obtenir une valeur plus fiable.

4

Bilan

La mise en œuvre d'une décision après le calcul de la CDL a permis de mettre en place un processus global de décision qui, grâce à la classification des CDL en deux classes, permet de décider si deux images sont similaires ou non. La base de support principale sur laquelle se fait les tests est une base d'impressions anciennes constituée dans le cadre du projet ANITA à partir de livres anciens issus de la MAT de Troyes.

La décision peut se faire à partir des descripteurs de faible dimension comme l'histogramme de la CDL ou sa courbe granulométrique. Les meilleurs résultats de classification sont obtenus avec les Séparateurs à Vaste Marge (SVM) avec un noyau linéaire. La granulométrie apporte de l'information sur la répartition spatiale de la CDL, mais ne permet pas d'obtenir de meilleurs résultats que ceux fournis par l'histogramme. Pour améliorer les résultats, il faut donner en entrée directement la CDL entière. Les données sont alors de haute dimension, problème que les SVM supportent mieux que la plupart des méthodes de classification. Plusieurs types de CDL sont alors testés pour déterminer le plus efficace. Il s'agit de la $CDL_{diffSimple}$ basée sur la différence simple, la CDL_{DH} , basée sur la distance de Hausdorff et la carte de différence simple.

Pour la CDL_{DH} qui donne les meilleurs résultats, l'étude détaillée du choix des paramètres des SVM est présentée. Sa combinaison avec une Analyse Multirésolution morphologique de la médiane permet d'obtenir des résultats à peine moins bons en réduisant la taille des images traitées par 8. Ceci est possible car le filtre médian morphologique conserve bien les structures dans les images binaires.

Ensuite, un tableau comparatif montre que les résultats obtenus avec les CDL sont meilleurs que ceux des mesures globales ou que ceux obtenus avec l'histogramme et la courbe granulométrique. D'autre part, la comparaison des résultats obtenus avec les différentes cartes permet de constater que ce n'est pas seulement l'information spatiale qui joue mais que cette information concerne aussi les dissimilarités locales (et pas simplement la différence pixel à pixel). Enfin, cette mesure de dissimilarité est mieux faite avec la DH qu'avec la Différence Simple.

Le processus global de classification a été testé sur deux autres bases, une base de formes et une base de visages. Les résultats obtenus sur la base de formes sont satisfaisants tandis que ceux obtenus sur la base de visages sont moins bons, notamment parce que le prétraitement peut-être amélioré.

La CDL peut être aussi utilisée sans le recours à une méthode de décision pour sa capacité de visualisation des dissimilarités locales. Dans cette optique, elle a été utilisée afin d'étudier les caractéristiques des filtres morphologiques.

Enfin, l'objet initial était de fournir une mesure de dissimilarité plus précise que les mesures globales proposées pour les images binaires. Il est possible de tirer parti de l'information contenue dans la CDL (distribution spatiale des dissimilarités) pour définir une mesure globale plus précise : la DH Filtrée, mesure définie comme le maximum sur la CDL filtrée.

Quatrième partie

Conclusion générale

Bilan

La comparaison d'images est un domaine de recherche très actif en traitement de l'image, en particulier en indexation et en recherche d'images par le contenu. Une méthode générale de comparaison d'images se fait en extrayant des descripteurs de l'image qui peuvent provenir de la texture, de la couleur ou des formes. Les descripteurs sont ensuite comparés à l'aide d'une mesure de dissimilarité.

Dans le cas des images binaires, la pauvreté de ces descripteurs fait que cette méthode ne donne en général pas de résultat satisfaisant surtout si l'image n'est pas composée d'une simple forme. Une comparaison directe des images est alors préférable. Elle se base en général sur une mesure de dissimilarité entre les images. Cette mesure peut être basée sur la différence pixel à pixel, mais elle est alors pauvre en information sur les formes et n'est pas en accord avec l'intuition. Elle peut aussi reposer sur une mesure entre ensembles de points, notamment la distance de Hausdorff. Cependant ces mesures sont globales et ne permettent pas d'accéder à l'information de dissimilarité qui peut être différente selon les zones des images comparées.

Pour répondre à ce manque, nous proposons la Carte de Dissimilarité Locale, qui est une méthode permettant de mesurer localement les dissimilarités entre les images. La mesure locale est faite à travers une fenêtre dont la taille est ajustée automatiquement à la taille de la dissimilarité. Dans le cas où la mesure dans la fenêtre est la distance de Hausdorff (DH), la formule de la CDL est simple et très vite calculée. Mais elle peut être une variante de la DH ou basée sur une mesure qui doit respecter quelques hypothèses afin d'assurer la cohérence de la CDL. La CDL apporte non seulement l'ensemble des valeurs de dissimilarités locales mais aussi leur distribution spatiale. La similarité ne concerne pas toutes les structures de l'image, il n'est donc pas nécessaire de comparer les images dans le détail. Pour une raison évidente de gain de temps de calcul, il est aussi intéressant de comparer les images à une résolution plus grossière. Le contrôle de la décomposition de l'image conduit à la structure d'analyse multirésolution. Les analyses multirésolutions adaptées aux images binaires sont non-linéaires et celles associées aux filtres morphologiques conviennent bien pour conserver les structures principales de l'image. Dans le cas de l'analyse multirésolution morphologique, le filtre associé doit être auto-dual et un filtre qui conserve assez bien les structures de l'image est le filtre médian.

La CDL entre deux images dépend de plusieurs choix :

- la mesure de dissimilarité (distance de Hausdorff, différence simple, ...),
- la distance spatiale entre les pixels (issue de la norme L_1 , L_2 , L_∞ , ...) si la mesure de dissimilarité repose sur une distance spatiale,
- la résolution à laquelle est faite la comparaison,
- ainsi que le choix entre l'image d'approximation ou celles des détails.

Ces choix peuvent dépendre des applications.

Dans le cas de la distance de Hausdorff, le calcul de la CDL se base sur la transformation en distance, ce qui constitue un pont pour étendre la CDL aux images en niveaux de gris. En effet, il existe des généralisations de la transformation en distance aux images en niveaux de gris. Cette généralisation est intéressante pour comparer des images proches, par exemple pour évaluer l'effet d'un filtre.

Un autre prolongement de la CDL est la DH Filtrée, en effet, il est possible de tirer partie de la CDL pour donner une mesure globale de dissimilarité. La mesure donnée dépend alors du type de dissimilarité entre les images et donne une solution au problème de la mesure due à des points non représentatifs des images.

La CDL donne une évaluation des similarités locales entre deux images et à ce titre ne permet pas de donner directement une décision sur la dissimilarité entre les deux images. Dans le but de prendre une décision sur la similarité des images, il est alors nécessaire de classer la CDL. Bien que cela ne semble que déplacer le problème (plutôt que de classer les images, il faut classer les CDL), il y a deux avantages :

- Autant il peut y avoir une multitude de classes d’images, autant il n’y a que deux types de CDL (celles provenant de la comparaison de deux images similaires et celles provenant de la comparaison de deux images dissimilaires),
- chacune des classes a une distribution en valeurs et une répartition spatiale des valeurs distinctes, ce qui permet de distinguer les deux classes.

Le fait qu’il y ait seulement deux classes de CDL permet de surcroît d’avoir de la souplesse concernant le nombre de classes d’images : il n’a pas besoin d’être fixé à l’avance et peut évoluer avec de nouvelles entrées.

La CDL a été utilisée dans une application sur une base d’impressions anciennes numérisées. Cette base constituée dans le cadre du projet ANITA en collaboration avec la MAT et le CEPLECA comporte 68 impressions anciennes numérisées fortement contrastées, regroupées en une vingtaine de classes. Leur fort contraste permet de les binariser sans perdre trop d’information et aussi de s’affranchir des variations des niveaux de gris initiaux (teinte du papier, dégradation de l’encre et du papier...). La comparaison par la CDL étant une comparaison locale d’images, il est nécessaire que les images soient recalées auparavant. Le recalage proposé est celui basé sur les deux premiers moments géométriques de l’image. Cependant pour évaluer précisément la CDL, cette étape a été faite manuellement sur la base d’impressions anciennes.

La classification de la CDL peut se faire à partir de descripteurs, nous avons testé comme descripteur l’histogramme et la courbe granulométrique. La méthode de classification qui donne les meilleurs résultats repose sur les séparateurs à vaste marge (SVM). Cependant, c’est avec la CDL entière comme donnée d’entrée du classificateur que les résultats sont les meilleurs. La CDL est une donnée de haute dimension, une méthode de classification qui le supporte bien est celle des SVM. Parmi les CDL, celle basée sur la DH, la CDL_{DH} , offre les meilleurs résultats avec plus de 96% de bonne classification. La comparaison avec les mesures globales montre la supériorité du processus global. Cela provient d’une part de la capacité de la DH locale à appréhender les dissimilarités locales et d’autre part au fait que la CDL entière permet l’accès à la distribution spatiale des mesures de dissimilarités. Le taux de classification est testé à plusieurs résolutions fournies par l’analyse multirésolution morphologique et les résultats restent bons pour des images de taille huit fois inférieure à celle de l’image initiale.

L’application à une base de formes avec de bons résultats montre la souplesse de la méthode. L’application à la base d’images de visages est moins concluante mais cela peut provenir du prétraitement.

Enfin, la CDL permet de visualiser les dissimilarités entre les images, ce que nous avons illustré en montrant l’effet de plusieurs filtres non-linéaires sur des images binaires.

Perspectives

Cette étude offre plusieurs perspectives.

Amélioration du processus global

Bien qu’il donne de très bons résultats dans l’application, l’apprentissage est supervisé et se fait de manière automatique avec les SVM, nous n’avons donc pas explicitement les critères pertinents de décision. Or ces critères peuvent être intéressants pour deux raisons :

- la première est de réduire le nombre d’entrées et d’accélérer ainsi le processus de décision,
- la deuxième est d’avoir un nombre réduit de paramètres afin de pouvoir faire un apprentissage non supervisé.

Les études menées ont données des indications : la distribution spatiale dans la carte joue un rôle important. L’histogramme et la granulométrie à eux seuls ne suffisent pas à capter l’ensemble de cette information. Cependant, nous n’avons pas de critère explicite, c’est donc un point à creuser. Un élément d’information se trouve peut être dans les vecteurs de support qui servent aux SVM pour prendre leur décision.

Un deuxième point à approfondir est la prise en compte du recalage des images pour avoir une chaîne complète de traitement. En effet, si dans le cadre de cette thèse, le sujet d'étude était le développement de la méthode, elle repose sur une comparaison locale et de ce fait doit s'appliquer à des images recalées. Elle peut d'ailleurs, à l'instar de ce qui a été fait par Huttenlocher [Huttenlocher and Rucklidge, 1993], servir pour optimiser le recalage.

Enfin, une piste à poursuivre est d'extraire de l'information de plus haut niveau et de l'intégrer dans la comparaison.

Carte des Dissimilarités Locales et développement

La CDL est une méthode assez générale de comparaison locale d'images qui repose sur une mesure locale. Cette mesure locale se trouve avoir une expression très simple dans le cas où la mesure initiale est la distance de Hausdorff. Il serait intéressant de développer la CDL avec d'autres mesures, par exemple celles dérivées des variantes de la distance de Hausdorff ou encore une mesure dérivant du psnr. Par ailleurs, une piste à explorer est la possibilité d'une CDL entre des surfaces en trois dimensions. Enfin, l'exploitation des propriétés de la CDL pour fournir une mesure globale a été explorée et mérite d'être approfondie, notamment pour le choix du filtrage effectué sur la CDL.

Analyse multirésolution morphologique

Le choix de l'AMR morphologique s'est révélé ardu car les possibilités de filtres dans le cas des images binaires ne sont pas très étendues ni facilement maîtrisables du fait de la non-linéarité des filtres. Les méthodes de caractérisation sont pratiquement inexistantes et la CDL a peut-être un rôle à y jouer. De même, la piste d'un histogramme permettant d'affiner les résultats de la granulométrie mérite d'être approfondie.

Visualisation

Un aspect intéressant de la CDL est qu'elle propose une image qui permet de visualiser immédiatement les dissimilarités. L'exploitation de cette propriété pour visualiser des transformations d'images (y compris en niveaux de gris) est une piste ouverte. Les applications peuvent se trouver dans le filtrage, la compression, la mesure de la qualité ou dans la quantification des dissimilarités entre deux images.

Les travaux effectués dans le cadre de la thèse ont donné lieu à deux exposés en conférence nationale [Nicolier *et al.*, 2003a; Baudrier *et al.*, 2003], quatre en conférences internationales [Nicolier *et al.*, 2003b; Baudrier *et al.*, 2004b; Baudrier *et al.*, 2004a; Baudrier *et al.*, 2005], dont une [Baudrier *et al.*, 2004b] a conduit à la sélection pour publication en revue internationale (en cours) et enfin, un article est en cours de relecture dans une revue internationale.

A

Preuves

A.1 Preuve de la proposition 2.3.1 (identité)

Démonstration. (\Leftarrow) L'hypothèse est $A \cap W = B \cap W$,

Si ces ensembles sont vides, par l'expression (2.4), les deux distance directes sont nulles et par l'expression (2.5), la DH locale est nulle.

Sinon, puisqu'ils sont égaux, l'un *et* l'autre des deux ensembles sont non-vides. Par conséquent

$$\forall a \in A \cap W, \min_{b \in B \cap W} (d(a, b)) = 0, \quad (\text{A.1})$$

et donc

$$\forall a \in A \cap W, \min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) = 0. \quad (\text{A.2})$$

Ainsi,

$$\max_{a \in A \cap W} \left[\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right] = 0. \quad (\text{A.3})$$

i.e., par l'expression (2.4),

$$h_W(A, B) = 0, \quad (\text{A.4})$$

de la même manière,

$$h_W(B, A) = 0 \quad (\text{A.5})$$

et donc, par l'expression (2.5),

$$HD_W(A, B) = 0. \quad (\text{A.6})$$

(\Rightarrow) L'hypothèse est $HD_W(A, B) = 0$,

alors les deux distances directes sont égales à 0. Traitons le cas de $h_W(A, B)$, deux cas se présentent :

(Cas où $A \cap W = \emptyset$) alors, il nous faut montrer que $B \cap W = \emptyset$. Supposons $B \cap W \neq \emptyset$, par l'expression (2.5),

$$h_W(B, A) = \max_{b \in B \cap W} \left[\min_{w \in Fr(W)} d(b, w) \right] \quad (\text{A.7})$$

or,

$$\forall b \in B \cap W, \forall w \in W, d(b, w) > 0 \quad (\text{A.8})$$

les bornes dans le minimum et dans le maximum étant atteintes, nous avons donc

$$h_W(B, A) > 0, \quad (\text{A.9})$$

Ce qui est contraire à l'hypothèse, donc

$$B \cap W = \emptyset. \quad (\text{A.10})$$

(Cas où $A \cap W \neq \emptyset$) alors $B \cap W$ n'est pas vide. En effet, si $B \cap W$ était vide, alors par le raisonnement précédent, nous aurions $A \cap W$ qui serait aussi vide. $A \cap W$ et $B \cap W$ ne sont donc pas vide, l'expression pour $h_W(A, B)$ est donc

$$h_W(A, B) = \max_{a \in A \cap W} \left[\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right] \quad (\text{A.11})$$

Nous avons donc, par hypothèse,

$$\max_{a \in A \cap W} \left[\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right] = 0, \quad (\text{A.12})$$

or

$$\forall a \in A, \forall w \in W, d(a, w) > 0, \quad (\text{A.13})$$

donc l'expression (A.12) devient

$$\max_{a \in A \cap W} \left[\min_{b \in B \cap W} d(a, b) \right] = 0, \quad (\text{A.14})$$

i.e.

$$h(A \cap W, B \cap W) = 0, \quad (\text{A.15})$$

En raisonnant de la même manière, nous obtenons aussi

$$h(B \cap W, A \cap W) = 0, \quad (\text{A.16})$$

et donc

$$DH(B \cap W, A \cap W) = 0. \quad (\text{A.17})$$

Or la DH est une distance mathématique qui vérifie la propriété de l'identité et (A.17) implique

$$B \cap W = A \cap W \quad (\text{A.18})$$

Ce qui finit de démontrer la réciproque. □

A.2 Preuve de la proposition 2.3.2 (majoration)

Démonstration. Supposons par l'absurde que $HD_W(A, B) > HD(A, B)$. Sans perte de généralité, on peut supposer que $HD_W(A, B) = h_W(A, B)$. Par hypothèse, $h_W(A, B) > 0$ donc, d'après l'expression (2.4), deux cas sont possibles :

1. (Si $A \neq \emptyset, B \neq \emptyset$) alors par (2.4),

$$h_W(A, B) = \max_{a \in A \cap W} \left(\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right)$$

donc

$$\exists a_0 \in A \cap W / h_W(A, B) = \min \left(\min_{b \in B \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right). \quad (\text{A.19})$$

Deux cas se présentent alors :

(a) Le minimum est atteint par le premier membre entre parenthèses :

$$h_W(A, B) = \min_{b \in B \cap W} d(a_0, b)$$

alors

$$\exists b_0 \in B \cap W / \min_{b \in B \cap W} d(a_0, b) = d(a_0, b_0),$$

i.e.

$$\exists b_0 \in B \cap W / h_W(A, B) = d(a_0, b_0). \quad (\text{A.20})$$

Or, pour le point a_0 , deux informations sont disponibles :

– La distance de ce point au bord de W est supérieure à $d(a_0, b_0)$.

En effet, si ce n'était pas le cas, nous aurions

$$\min \left(\min_{b \in B \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right) = \min_{w \in Fr(W)} d(a_0, w),$$

ce qui est contraire à l'hypothèse.

– On rappelle que W est une boule, d'après le point précédent, on a donc

$$B(a_0, d(a_0, b_0)) \subset W \quad (\text{A.21})$$

et par définition de b_0 , il n'y a pas d'autre point de B dans $B(a_0, d(a_0, b_0))$ que b_0 .

Avec cette dernière donnée, il est facile de calculer que pour a_0 ,

$$\min_{b \in B} d(a_0, b) = d(a_0, b_0)$$

i.e. $d(a_0, b_0)$ est l'un des minima, donc

$$d(a_0, b_0) \leq \max_{a \in A} \left(\min_{b \in B} d(a, b) \right).$$

Maintenant, on peut déduire dans ce cas (a) que

$$HD_W(A, B) = d(a_0, b_0),$$

donc

$$HD_W(A, B) \leq HD(A, B)$$

ce qui est en contradiction avec l'hypothèse.

- (b) Voyons maintenant le cas où le minimum de l'équation (A.19) est atteint par le deuxième membre entre parenthèses :

$$h_W(A, B) = \min_{w \in Fr(W)} d(a_0, w),$$

notons $r_0 = \min_{w \in Fr(W)} d(a_0, w)$, alors $B(a_0, r_0) \subset W$. D'autre part, comme nous sommes dans le deuxième cas, le minimum (de l'équation (A.19)) n'est pas atteint par un point de B , donc $B \cap B(a_0, r_0) = \emptyset$. D'où

$$r_0 \leq \min_{b \in B} d(a_0, b)$$

et donc

$$r_0 \leq \max_{a \in A} \left(\min_{b \in B} d(a, b) \right), \quad (\text{A.22})$$

or par définition,

$$\begin{aligned} r_0 &= \min_{w \in Fr(W)} d(a_0, w) \\ &= h_W(A, B) \\ &= HD_W(A, B). \end{aligned}$$

En remplaçant dans l'équation (A.22), on obtient

$$HD_W(A, B) \leq \max_{a \in A} \left(\min_{b \in B} d(a, b) \right), \quad (\text{A.23})$$

d'autre part,

$$\begin{aligned} \max_{a \in A} \left(\min_{b \in B} d(a, b) \right) &= h_W(A, B) \\ &\leq HD(A, B). \end{aligned}$$

En remplaçant dans l'équation (A.23), on obtient

$$HD_W(A, B) \leq HD(A, B),$$

ce qui est en contradiction avec l'hypothèse.

2. (Si $A \neq \emptyset, B = \emptyset$) alors par (2.4),

$$h_W(A, B) = \max_{a \in A \cap W} \left(\min_{w \in Fr(W)} d(a, w) \right)$$

donc

$$\exists a_0 \in A \cap W / h_W(A, B) = \min_{w \in Fr(W)} d(a_0, w). \quad (\text{A.24})$$

la situation est alors la même que dans le cas (b) traité ci-dessus.

□

A.3 Preuve de la proposition 2.3.3 (croissance)

Démonstration. La différence avec la propriété précédente est la présence de la distance aux bords dans HD_W . Montrons d'abord un résultat préliminaire qui sera utile pour la démonstration.

Comme $V \subset W$, on a

$$\forall v \in V, d(v, Fr(V)) \leq d(v, Fr(W)), \quad (\text{A.25})$$

or, par définition,

$$h_V(A, B) = \max_{a \in A \cap V} \left[\min \left(\min_{b \in B \cap V} d(a, b), \min_{v \in Fr(V)} d(a, v) \right) \right]$$

grâce à l'inégalité (A.25), on peut majorer le minimum sur la frontière de V par celui sur celle de W

$$h_V(A, B) \leq \max_{a \in A \cap V} \left[\min \left(\min_{b \in B \cap V} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right]. \quad (\text{A.26})$$

Commençons maintenant la démonstration de la même manière que la précédente :

Supposons par l'absurde que

$$HD_V(A, B) > HD_W(A, B) \quad (\text{A.27})$$

alors, sans perte de généralité, nous pouvons supposer que $HD_V(A, B) = h_V(A, B)$. (A.27) indique que $h_V(A, B) > 0$, donc d'après (2.6)

$$\exists a_0 \in A \cap V / h_V(A, B) = \begin{cases} \min(\min_{b \in B \cap V} d(a_0, b), \min_{v \in Fr(V)} d(a_0, v)) & \text{si } A \cap V \neq \emptyset, B \cap V \neq \emptyset \\ \min_{v \in Fr(V)} d(a_0, v) & \text{si } A \cap V \neq \emptyset, B \cap V = \emptyset. \end{cases} \quad (\text{A.28})$$

Comme dans la démonstration précédente, nous allons traiter le premier cas et le deuxième va s'en déduire.

1. (Si $A \cap V \neq \emptyset, B \cap V \neq \emptyset$) Deux cas peuvent être distingués :

(a) Le minimum est atteint par le premier membre entre parenthèses :

$$h_V(A, B) = \min_{b \in B \cap V} d(a_0, b),$$

alors

$$\exists b_0 \in B \cap V / \min_{b \in B \cap V} d(a_0, b) = d(a_0, b_0), \quad (\text{A.29})$$

i.e.

$$\exists b_0 \in B \cap V / h_V(A, B) = d(a_0, b_0).$$

Or, deux informations concernant le point a_0 sont disponibles :

– Comme le minimum est atteint par le premier membre dans l'équation (A.28), cela signifie que la valeur du premier membre est inférieure à la valeur du second :

$$\min_{b \in B \cap V} d(a_0, b) \leq \min_{v \in Fr(V)} d(a_0, v),$$

en utilisant l'expression (A.29), on obtient

$$d(a_0, b_0) \leq \min_{v \in Fr(V)} d(a_0, v) \quad (\text{A.30})$$

– Nous obtenons comme conséquences de cette première information :

$$\begin{aligned} B(a_0, d(a_0, b_0)) &\subset V \\ \forall b \in (B(a_0, d(a_0, b_0)) \cap B), \quad d(a_0, b) &= d(a_0, b_0). \end{aligned} \quad (\text{A.31})$$

Avec ce dernier point, il est facile de calculer que $d(a_0, b_0)$ est l'un des minima pour $b \in B \cap W$, *i.e.*

$$\min_{b \in B \cap W} d(a_0, b) = d(a_0, b_0).$$

Prouvons maintenant que

$$d(a_0, b_0) \leq \min_{w \in Fr(W)} d(a_0, w).$$

Comme $a_0 \in V$ et $V \subset W$, nous avons

$$\min_{v \in Fr(V)} d(a_0, v) \leq \min_{w \in Fr(W)} d(a_0, w)$$

D'après l'équation (A.30), nous avons $d(a_0, b_0) \leq \min_{v \in Fr(V)} d(a_0, v)$ donc

$$d(a_0, b_0) \leq \min_{w \in Fr(W)} d(a_0, w). \quad (\text{A.32})$$

D'autre part, d'après l'équation (A.31), nous avons

$$\forall b \in B \cap B(a_0, d(a_0, b_0)), \quad d(a_0, b) = d(a_0, b_0),$$

donc

$$\forall b \in B \cap W, \quad d(a_0, b) \geq d(a_0, b_0)$$

i.e.

$$d(a_0, b_0) \leq \min_{b \in B \cap W} d(a_0, b) \quad (\text{A.33})$$

D'après les équations (A.31) et (A.33), nous avons

$$d(a_0, b_0) \leq \min \left(\min_{b \in B \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right). \quad (\text{A.34})$$

Maintenant, par définition,

$$h_W(A, B) = \max_{a \in A \cap W} \left(\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right),$$

or le maximum sur l'ensemble des points de A est supérieur à la valeur au point a_0 :

$$\max_{a \in A \cap W} \left(\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right) \geq \min \left(\min_{b \in B \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right),$$

et donc, d'après (A.34)

$$\max_{a \in A \cap W} \left(\min \left(\min_{b \in B \cap W} d(a, b), \min_{w \in Fr(W)} d(a, w) \right) \right) \geq d(a_0, b_0)$$

donc

$$h_W(A, B) \geq d(a_0, b_0).$$

Ce qui est en contradiction avec l'hypothèse (A.27).

(b) Le minimum de l'équation (A.28) est atteint par le deuxième terme entre parenthèses :

$$h_V(A, B) = \min_{v \in Fr(V)} d(a_0, v).$$

Notons $r_0 = \min_{v \in Fr(V)} d(a_0, v)$, comme V est une boule par hypothèse, nous avons

$$B(a_0, r_0) \subset V$$

d'autre part, le minimum est atteint par le deuxième membre de (A.28) donc

$$\forall b \in B \cap B(a_0, r_0), d(a_0, b) \geq r_0,$$

donc

$$\forall b \in B, d(a_0, b) \geq r_0,$$

d'où

$$\forall b \in B \cap W, d(a_0, b) \geq r_0,$$

et donc

$$\min_{b \in B \cap W} d(a_0, b) \geq r_0. \tag{A.35}$$

Or, comme $a_0 \in V$ et $V \subset W$, alors

$$\forall v \in Fr(V), \forall w \in Fr(W), d(a_0, v) \leq d(a_0, w)$$

d'où

$$r_0 = \min_{v \in Fr(V)} d(a_0, v) \leq \min_{w \in Fr(W)} d(a_0, w) \tag{A.36}$$

en regroupant (A.35) et (A.36), on obtient :

$$\min \left(\min_{b \in B \cap W} d(a_0, b), \min_{w \in Fr(W)} d(a_0, w) \right) \geq r_0,$$

i.e.

$$h_W(A, B) \geq r_0,$$

or $r_0 = h_V(A, B)$ d'où

$$h_W(A, B) \geq h_V(A, B),$$

ce qui est en contradiction avec l'hypothèse (A.27).

2. (Si $A \cap V \neq \emptyset, B \cap V = \emptyset$) alors nous avons

$$h_V(A, B) = \min_{v \in Fr(V)} d(a_0, v) \tag{A.37}$$

Et cela nous ramène au cas (b), qui est démontré ci-dessus.

□

Glossaire

AMR : Analyse MultiRésolution

AMRM : Analyse MultiRésolution Morphologique

CDL : Carte des Dissimilarités Locales

CDL_{DH} : Carte des Dissimilarités Locales basée sur la distance de Hausdorff

CDL_{DifSimple} : Carte des Dissimilarités Locales basée sur la différence simple

CDS : Carte de Différence Simple

DH : Distance de Hausdorff

ES : Élément Structurant

TeD : Transformée en Distance

VVC : Vecteur des Valeurs Classées (valeurs contenues dans la CDL)

Bibliographie

- [Aït Aouït, 2004] Djédjiga Aït Aouït. Classification d'images par la méthode des support vector machines : étude et applications. Master's thesis, University of Reims-Champagne-Ardenne, Troyes, septembre 2004.
- [Angulo and Serra, 2003] Jesús Angulo and Jean Serra. Automatic analysis of DNA microarray images using mathematical morphology. *Bioinformatics*, 19(5) :553–562, 2003.
- [Antani *et al.*, 2002] Sameer Antani, Rangachar Kasturi, and Ramesh Jain. A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. *Pattern Recognition*, 35(4) :945–965, avril 2002.
- [Arlandis and Perez-Cortes, 2000] J. Arlandis and J.C. Perez-Cortes. The continuous distance transformation : A generalization of the distance transformation for continuous-valued images. In A.Sanfeliu M.I. Torres, editor, *Pattern Recognition and Applications*, volume 56 of *Frontiers in Artificial Intelligence and Applications*, pages 89–98. IOS Press, 2000.
- [aym, 1631] *L'Histoire des quatre fils Aymon*. Nicolas Oudot, Troyes, 1631.
- [aym, 1784] *L'Histoire des quatre fils d'Aimon*. Jacques Garrigan, Avignon, 1784.
- [aym, debut XIXea] *L'Histoire des quatre fils Aymon*. Citoyenne Garnier, Troyes, début XIX^e.
- [aym, debut XIXeb] *L'Histoire des quatre fils d'Aymon*. F. Chapoulaud, Limoges, début XIX^e.
- [aym, debut XVIIIe] *L'Histoire des quatre fils d'Aymon*. Pierre Garnier, Troyes, début XVIII^e.
- [aym, milieu XIXe] *L'Histoire des quatre fils Aymon*. Blocquel-Castiaux, Lille, milieu XIX^e.
- [Baddeley, 1992] A. J. Baddeley. An error metric for binary images. In W. Förstner and H. Ruwiedel, editors, *Robust Computer Vision : Quality of Vision Algorithms*, pages 59–78. Wichmann, Karlsruhe, 1992.
- [Bangham *et al.*, 1994] J. Andrew Bangham, T. George Campbell, and Richard V. Aldridge. Multiscale median and morphological filters for 2d pattern recognition. *Signal Processing*, 38(3) :387–415, 1994.
- [Barbe and Ledoux, 1998] P. Barbe and M. Ledoux. *Probabilité, de la licence à l'agrégation*. Belin, 1998.
- [Battle, 1987] G. Battle. A block spin construction of ondelettes. *Comm. Math. Phys.*, 110 :601–615, 1987.
- [Baudrier *et al.*, 2003] Étienne Baudrier, Gilles Millon, and Frédéric Nicolier. Impressions virtuelles de tampons en bois gravés anciens : binarisation et recalage d'images de profondeur. In *colloque Maquette Virtuelle et Patrimoine*, pages 25–27, Cluny (France), mars 2003.
- [Baudrier *et al.*, 2004a] Étienne Baudrier, Gilles Millon, Frédéric Nicolier, and Su Ruan. A new similarity measure using Hausdorff distance map. In *Proc of international conference on image processing (ICIP)*, pages 669–672, Singapour, 24-27 oct 2004. IEEE.
- [Baudrier *et al.*, 2004b] Étienne Baudrier, Gilles Millon, Frédéric Nicolier, Ralph Seulin, and Su Ruan. Hausdorff distance based multiresolution maps applied to an image similarity measure. In *Proc of the The Topical Meeting on Optical Sensing and Artificial Vision (OSAV'2004) (SPIE)*, St Petersburg, Russie, 18-21 oct 2004.

- [Baudrier *et al.*, 2005] Étienne Baudrier, Gilles Millon, and Frédéric Nicolier. A study of non-linear morphological filters. In *À paraître*, Boston, oct 2005. SPIE.
- [bib, 1735] *Figures de la Sainte Bible*. Jean Oudot, Troyes, 1735.
- [bib, fin XVIIIe] *Figures de la Sainte Bible*. Jean-Antoine Garnier, Troyes, fin XVIII^e.
- [Bijaoui *et al.*, 1994] Albert Bijaoui, Jean-Louis Starck, and Fionn Murtagh. Restauration des images multi-échelles par l'algorithme à trous. *Traitement du signal*, 11(3) :229–243, 1994.
- [Bishop, 1995] C. Bishop. *Neural networks for pattern recognition*. Oxford University Press, 1995.
- [Blanz *et al.*, 1996] Volker Blanz, Bernhard Schölkopf, Heinrich H. Buhmann, Chris Burges, Vladimir Vapnik, and Thomas Vetter. Comparison of view-based object recognition algorithms using realistic 3d models. In *ICANN 96 : Proceedings of the 1996 International Conference on Artificial Neural Networks*, pages 251–256, London, UK, 1996. Springer-Verlag.
- [Borgefors, 1986a] Gunilla Borgefors. Distance transformations in digital images. *Comput. Vision Graph. Image Process.*, 34(3) :344–371, 1986.
- [Borgefors, 1986b] Gunilla Borgefors. A new distance transformation approximating the euclidean distance. In *Proc. International Joint Conference on Pattern Recognition*, pages 336–338, 1986.
- [Boser *et al.*, 1992] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *COLT '92 : Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, New York, USA, 1992. ACM Press.
- [Breu *et al.*, 1995] H. Breu, J. Gil, D. Kirkpatrick, and M. Werman. Linear time euclidean distance transform algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5) :529–533, 1995.
- [Brown, 1994] R.L. Brown. The fringe distance measure. *IEEE Transactions on Systems Man and Cybernetics*, 24(1), 1994.
- [Brunelli and Mich, 2001] R. Brunelli and O. Mich. Histograms analysis for image retrieval. *Pattern Recognition*, 34(8) :1625–1637, 2001.
- [Chen *et al.*, 2003] Jingying Chen, Maylor K. Leung, and Yongsheng Gao. Noisy logo recognition using line segment Hausdorff distance. *Pattern Recognition*, 36(4) :943–955, avril 2003.
- [Chim *et al.*, 1999] Y. C. Chim, A. A. Kassima, and Y. Ibrahim. Character recognition using statistical moments. *Image and Vision Computing*, 17(3-4) :299–307, mars 1999.
- [C.I.E., 1978] C.I.E. Recommendations on uniform color spaces, color difference equations, psychometric color terms. *Supplement No.2 to CIE publication, Bureau Central de la CIE, Paris*, 15, 1978.
- [Ciocca and Schettini, 1999] G. Ciocca and R. Schettini. Using a relevance feedback mechanism to improve content-based image retrieval. In *Visual Information and Information Systems*, pages 107–114, 1999.
- [Cohen and Daubechies, 1993] A. Cohen and I. Daubechies. Non-separable bidimensional wavelet bases. *Rev. Mat. Iberoamericana*, 9(1) :51–137, 1993.
- [Cortes and Vapnik, 1995] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Mach. Learn.*, 20(3) :273–297, 1995.
- [Cover and Hart,] T.M. Cover and P.E. Hart. Nearest neighbor pattern classification.
- [Cristianini and Shawe-Taylor, 2000] N. Cristianini and J. Shawe-Taylor. *Introduction to Support Vector Machines*. Cambridge University Press, 2000.

- [Daubechies, 1992] Ingrid Daubechies. Ten lectures on wavelets. *Applied Mathematics*, 61, 1992.
- [Daubechies, 1993] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets II : variations on a theme. *SIAM J. Math. Anal.*, 24(2) :499–519, 1993.
- [Donoho, 1992] David L. Donoho. Interpolating wavelet transforms. Technical report, Preprint Department of Statistics, Stanford University, 1992.
- [Dubuisson and Jain, 1994] Marie-Pierre Dubuisson and A. K. Jain. A modified Hausdorff distance for object matching. In *Proc. 12th IAPR International Conference on Pattern Recognition*, pages 566–568, octobre 1994.
- [Gagalowicz, 1983] A. Gagalowicz. *Vers un modèle de texture*. PhD thesis, Université Pierre et Marie Curie, Paris VI, 1983.
- [Gao and Leung, 2002] Yongsheng Gao and Maylor K. H. Leung. Line segment hausdorff distance on face matching. *Pattern Recognition*, 35(2) :361–371, février 2002.
- [Goutsias and Heijmans, 2000] J. Goutsias and H.J.A.M. Heijmans. Multiresolution signal decomposition schemes. part 1 : Linear and morphological pyramids. *IEEE Trans. Image Processing*, 9(11) :1862–1876, novembre 2000.
- [Guo *et al.*, 2001] B. Guo, K. Lam, W. Siu, and S. Yang. Human face recognition using a spatially weighted Hausdorff distance. In *The 2001 IEEE International Symposium on Circuits and Systems*, pages 145–148, 2001.
- [Heijmans and Goutsias, 2000] H.J.A.M. Heijmans and J. Goutsias. Multiresolution signal decomposition schemes. part 2 : Linear and morphological pyramids. *IEEE Trans. Image Processing*, 9(11) :1897–1913, novembre 2000.
- [Herley and Vetterli, 1993] C. Herley and M. Vetterli. Wavelets and recursive filter banks. *IEEE Trans. Signal Process.*, 41(8) :2536–2556, 1993.
- [Holschneider *et al.*, 1989] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian. A real-time algorithm for signal analysis with the help of the wavelet transform. In *Wavelets, Time-Frequency Methods and Phase Space*, pages 289–297, Berlin, 1989. Springer-Verlag.
- [Huang and Mitchell, 1994] C.T. Huang and O.R. Mitchell. An euclidean distance transform using grayscale morphology decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(4) :443–448, 1994.
- [Huang *et al.*, 1997] Jin Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. Image indexing using color correlograms. In *Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 762–768, 1997.
- [Huttenlocher and Rucklidge, 1993] Daniel P. Huttenlocher and William J. Rucklidge. A multi-resolution technique for comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 705–706, 1993.
- [Huttenlocher *et al.*, 1993] Daniel P. Huttenlocher, D. Klanderman, and William J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9) :850–863, septembre 1993.
- [Øivind Due Trier and Taxt, 1995] Øivind Due Trier and Torfinn Taxt. Evaluation of binarization methods for document images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(3) :312–315, 1995.
- [Jesorsky *et al.*, 2001] Oliver Jesorsky, Klaus J. Kirchberg, and Robert W. Frischholz. Robust Face Detection Using the Hausdorff Distance. In Josef Bigun and Fabrizio Smeraldi, editors, *Audio- and Video-Based Person Authentication - AVBPA 2001*, volume 2091 of *Lecture Notes in Computer Science*, pages 90–95, Halmstad, Sweden, 2001. Springer.
- [Joachims, 1998] Thorsten Joachims. Text categorization with support vector machines : Learning with many relevant features. In *ECML '98 : Proceedings of the 10th European Conference on Machine Learning*, pages 137–142, London, UK, 1998. Springer-Verlag.

- [Kovács and Guerrieri, 1992] Zs. M. Kovács and R. Guerrieri. Computer recognition of handwritten characters using the distance transform. *Electronic Letters*, 28(19) :1825–1827, 1992.
- [Laÿ, 1987] B. Laÿ. Recursive algorithms in mathematical morphology. *Acta Stereologica*, 6(3) :691–696, September 1987.
- [Leclerc, 2000] Marie-Dominique Leclerc. À la recherche du livre perdu : Identification de quelques bois gravés dans la Bibliothèque Bleue de Troyes. In *La Bibliothèque Bleue et les littératures de colportage*, volume 7, pages 109 – 129. Étude réunies par T. Delcourt et E. Parinet, Études et rencontres de l'École des chartes, Paris - Troyes, École des chartes - La maison du Boulanger, 2000.
- [Lemarié, 1988] P.-G. Lemarié. Ondelettes a localisation exponentielle. *J. Math. Pures Appl.*, 67(3) :227–236, 1988.
- [Leonard, 1991] Desiree M. Leonard. Analysis of object in binary images. Technical report, NASA, 1991.
- [Lin *et al.*, 2003] Kwan-Ho Lin, Kin-Man Lam, and Wan-Chi Siu. Spatially eigen-weighted hausdorff distances for human face recognition. *Pattern Recognition*, 36(8) :1827–1834, août 2003.
- [Liu and Wechsler, 2003] Chengjun Liu and Harry Wechsler. Independent component analysis of Gabor features for faces recognition. *IEEE transaction on neural networks*, 14(4) :919–929, juillet 2003.
- [Lu and Tan, 2002] Y. Lu and C.L. Tan. Word spotting Chinese document images without layout analysis. In *Proceeding of 16th International Conference on Pattern Recognition*, volume 3, pages 57–60, août 2002.
- [Lu *et al.*, 2001] Y. Lu, C.L. Tan, W. Huang, and L. Fan. An approach to word image matching based on weighted Hausdorff distance. In *Proc. 6th Internat. Conf. on Document Anal. Recogn.*, pages 921–925, 10-13 septembre 2001.
- [Maddess *et al.*, 2004] T. Maddess, Y. Nagai, A.C. James, and A. Ankiewicz. Binary and ternary textures containing higher-order spatial correlations. *Vision research*, 44 :1093–1113, 2004.
- [Mallat, 1989] Stéphane Mallat. A theory for multiresolution signal decomposition : The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11 :674–693, juillet 1989.
- [Mallat, 2000] Stéphane Mallat. *Une exploration des signaux en ondelettes*. Editions de l'école Polytechnique, 2000.
- [Matheron, 1967] George Matheron. *Éléments pour une théorie des milieux poreux*. Masson, Paris, 1967.
- [Matheron, 1975] George Matheron. *Random sets and integral geometry*. John Wiley and sons, 1975.
- [McCulloch and Pitts, 1943] W.S. McCulloch and W. Pitts. A logical calculus of the ideas of immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5 :115–133, 1943.
- [Meyer, 1990] Yves Meyer. *Ondelettes et opérateurs*. Herman, Paris, 1990.
- [Millon *et al.*, 2002] Gilles Millon, Frédéric Nicolier, and Alain Robert. Études d'images médiévales. In *Colloque national sur la recherche universitaire dans les IUT CNR'IUT2002*, pages 66–70, Le Creusot, mai 2002.
- [Niblack, 1986] W. Niblack. *An Introduction to Digital Image Processing*, pages 115–116. Prentice Hall, 1986.
- [Nicolier *et al.*, 2003a] Frédéric Nicolier, Gilles Millon, and Étienne Baudrier. Le projet ANITA : ANalyse d'Images et de Tampons Anciens. In *colloque Maquette Virtuelle et Patrimoine*, pages 51–54, Cluny, mars 2003.
- [Nicolier *et al.*, 2003b] Frédéric Nicolier, Gilles Millon, and Étienne Baudrier. A wavelet filter criterion for an a-priori evaluation of wavelet coding and denoising performances. In *Int. Symp. on Signal Proc. and its Applications*, pages 475–478, Paris, juillet 2003. IEEE.
- [Osuna *et al.*, 1997] Edgar Osuna, Robert Freund, and Federico Girosi. Training support vector machines : an application to face detection. In *Conference on Computer Vision and Pattern Recognition*, pages 130–136, juin 1997.

- [Otsu, 1979] N. Otsu. A threshold selection method from grey scale histogram. *IEEE Trans. on Syst. Man and Cyber.*, 1 :62–66, 1979.
- [Parzen, 1962] E Parzen. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 1962.
- [Pass *et al.*, 1996] Greg Pass, Ramin Zabih, and Justin Miller. Comparing images using color coherence vectors. In *ACM International Multimedia Conference*, pages 65–73, 1996.
- [Paumard, 1997] José Paumard. Robust comparison of binary images. *Pattern Recognition Letters*, 18(10) :1057–1063, octobre 1997.
- [Philipp-Foliguet *et al.*, 2002] S. Philipp-Foliguet, M. B. Vieira, and M. Sanfourche. Fuzzy segmentation of color images and indexing of fuzzy regions. In *CGIV*, pages 507–512, Poitiers, 2002.
- [Pratt, 1977] W. K. Pratt. *Digital image processing*. John Wiley and Sons, 1977.
- [Ragot, 2003] Nicolas Ragot. *MÉLIDIS : Reconnaissance de formes par modélisation mixte intrinsèque/discriminante, base de systèmes d'inférence floue hiérarchisés*. PhD thesis, Université de Rennes 1, oct 2003.
- [Rosenfeld and Kak, 1982] A. Rosenfeld and A. C. Kak. *Digital picture processing*. Academic press, 1982.
- [Rosenfeld and Pfalz, 1966] A. Rosenfeld and J. L. Pfalz. Sequential operations in digital picture processing. *Journal of the ACM*, 13 :471–494, 1966.
- [Rui *et al.*, 1999] Y. Rui, T. Huang, and S. Chang. Image retrieval : current techniques, promising directions and open issues, 1999.
- [Sauvola *et al.*, 1997] Jaakko J. Sauvola, Tapio Seppänen, Sami Haapakoski, and Matti Pietikäinen. Adaptive document binarization. In *ICDAR '97 : Proceedings of the 4th International Conference on Document Analysis and Recognition*, pages 147–152, Washington, DC, USA, 1997. IEEE Computer Society.
- [Scassellati *et al.*, 1994] Brian Scassellati, Sophoclis Alexopoulos, and Myron Flickner. Retrieving images by 2D shape : a comparison of computation methods with human perceptual judgements. *SPIE*, 2185, 1994.
- [Sebastian *et al.*, 2001] Thomas B. Sebastian, Philip N. Klein, and Benjamin B. Kimia. Recognition of shapes by editing shock graphs. In *ICCV*, volume 1, pages 755–762, 2001.
- [Serra, 1982] Jean Serra. *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, 1982.
- [Serra, 1988] Jean Serra. *Image Analysis and Mathematical Morphology*, volume 2. Academic Press, 1988.
- [Shensa, 1992] Mark J Shensa. The discrete wavelet transform : Wedding the atrous and mallat algorithms. *IEEE Transactions on Signal Processing*, 40(10) :2464–2482, 1992.
- [Sim *et al.*, 1999] D.-G. Sim, O.-K. Kwon, and R.-H. Park. Object matching algorithms using robust Hausdorff distance measures. *IEEE Transaction on Image Processing*, 8(3) :425–429, 1999.
- [Smeulders *et al.*, 2000] A. W. M. Smeulders, M. Worring, M. Santini, S. Gupta, and Ramesh Jain. Content based image retrieval at the end of the early years. *IEEE Transcription on Pattern Analysis and Machine Intelligence*, 22 :1349–1380, 2000.
- [Smith and Chang, 1996] John R. Smith and Shih-Fu Chang. Visualseek : a fully automated contentbased image query system. In *ACM Multimedia Conference*, pages 87–98, Boston, USA, novembre 1996.
- [Soille, 2003] Pierre Soille. *Morphological Image Analysis*. Springer-Verlag, Heidelberg, 2nd edition, 2003.
- [Soille, 2005] Pierre Soille. Beyond self-duality in morphological image analysis. *Image and Vision Computing*, 23(2) :249–257, Février 2005.

- [Stricker and Orengo, 1995] Markus A. Stricker and M. Orengo. Similarity of color images. In Ramesh C. Jain Wayne Niblack, editor, *Proc. of Storage and Retrieval for Image Video Databases III*, volume 2420, pages 381–392. SPIE, mars 1995.
- [Strintzis, 1996] Michael G Strintzis. Optimal biorthogonal wavelet bases for signal decomposition. *IEEE Transactions on Signal Processing*, 44(6) :1406–1417, 1996.
- [Sweldens, 1995] Wim Sweldens. Lifting scheme : a new philosophy in biorthogonal wavelet constructions. volume 2569, pages 68–79. SPIE, 1995.
- [Sweldens, 1997] W. Sweldens. The lifting scheme : A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2) :511–546, 1997.
- [Takàcs, 1998] B. Takàcs. Comparing faces using the modified Hausdorff distance. *Pattern Recognition*, 31(12) :1873–1881, décembre 1998.
- [Thomas *et al.*, 2005] Édouard Thomas, David Fofi, Frédéric Nicolier, Gilles Millon, and Ralph Seulin. Low-cost system for ancient stamps range image acquisition. In *Applications in Industrial Inspection XIII, part of the IS&T/SPIE International Symposium on Electronic Imaging : Science and Technology*, San Jose, California USA, jan 2005.
- [Truchetet, 1998] Frédéric Truchetet. *Ondelettes pour le signal numérique*. Hermès puis Lavoisier, 1998.
- [Tversky, 1977] Amos Tversky. Features of similarity. *Psychological review*, 84 :327–352, 1977.
- [Unser, 1995] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. on Image Processing*, 4 :1549–1560, novembre 1995.
- [Vapnik, 1998] Vladimir Vapnik. *Statistical Learning Theory*. John Wiley, NY, 1998.
- [Varlot, 1859] Varlot. *Xylographie de l'imprimerie troyenne*. Varlot/Aubry, Troyes/Paris, 1859.
- [Veltkamp and Hagedoorn, 1999] R. Veltkamp and M. Hagedoorn. State-of-the-art in shape matching. Technical Report UU-CS-1999-27, Utrecht University, the Netherlands, 1999.
- [Veltkamp and Hagedoorn, 2000] Remco C. Veltkamp and Michiel Hagedoorn. Shape similarity measures, properties and constructions. Technical Report UU-CS-2000-37, Utrecht University, October 2000.
- [Wang *et al.*, 2004] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P Simoncelli. Image quality assessment : from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), janvier 2004.
- [Wolf *et al.*, 2002] Christian Wolf, Jean-Michel Jolion, and F. Chassaing. Extraction de texte dans des vidéos : le cas de la binarisation. In *13ème congrès francophone de reconnaissance des formes et intelligence artificielle*, pages 145–152, 2002.
- [Wong and Hall, 1978] R.Y. Wong and E.L. Hall. Scene matching with invariant moments. *Comput. Graphics Image Processing*, 8 :16–24, 1978.
- [Zhao *et al.*, 2004] Chunjian Zhao, Wenkang Shi, and Yong Deng. A new Hausdorff distance for image matching. *Pattern Recognition Letters*, 2004.

Résumé

Mots-clés : comparaison locale d'image, images binaires, analyse multirésolution, distance de Hausdorff, morphologie mathématique, classification, SVM, impressions anciennes

Cette thèse se situe dans le cadre de la comparaison d'image. Elle est consacrée au développement d'une méthode de comparaison locale d'images binaires. En s'appuyant sur un exemple de mesure - la distance de Hausdorff (DH) - une mesure locale (à travers une fenêtre) est définie, et ses propriétés en fonction de la taille de la fenêtre et de la mesure globale entre les deux images sont prouvées. Cela permet de définir un critère pour fixer une taille de fenêtre ajustée à celle de la dissimilarité locale. Cette méthode permet de définir une Carte de Dissimilarité Locales (CDL) lorsque la mesure locale est faite sur tous les points de l'image. Elle n'est pas valable uniquement avec la DH, cependant les propriétés de la DH font que le calcul de la CDL dans ce cas est très rapide. La CDL est à la fois un outil de visualisation des dissimilarités entre deux images et une base pour décider de la similarité des images.

Dans cette optique, une première étape est la mise en œuvre d'une analyse multirésolution adaptée aux images binaires reposant sur le filtre de la médiane morphologique qui offre la possibilité de traiter l'information à une résolution adaptée au degré de similarité recherché. La deuxième étape est l'utilisation de l'information de la CDL concernant les dissimilarités et leur distribution spatiale pour comparer les images. Plusieurs méthodes sont testées, et la plus efficace est basée sur les SVM auxquels on fournit en entrée les CDL entières. Les tests réalisés sur une base d'impressions anciennes numérisées et sur une base de formes montrent l'efficacité de la méthode.

Abstract

Keywords: local image comparison, binary images, multiresolution analysis, Hausdorff distance, mathematical morphology, classification, SVM, ancient impressions

This PhD deals with the comparison of binary images that are not composed of a single pattern. The proposed method is then extended to gray level images. Using a measure example - the Hausdorff distance (HD) - a local measure is defined through a window, and its properties depending on the window size and the global HD measure are proved. Thanks to them, a window-size criterion is defined so as the window to be adjusted to the local dissimilarity. The local dissimilarity map (LDM) is then defined when the window slides over all the compared images. The LDM can be defined with other measure than the HD using the same algorithm, nevertheless, the DH properties leads to a LDM fast computation. The LDM can be used as image dissimilarity visualization method or a tool to decide on image similarity.

For this last point, a first step is a binary-image adapted multiresolution analysis which is base on the median morphological filter. This allows to have an resolution adapted to the researched similarity degree. A second step consists in using LDM information concerning the dissimilarities and their spatial distribution to compare the images. Several comparison methods are tested, the most efficient one is based on the SVM with the whole LDM as input data. The method efficiency is successfully tested on an ancient-impressions database and on a face database.

