



HAL
open science

Le signal monogène couleur : théorie et applications

Guillaume Demarcq

► **To cite this version:**

Guillaume Demarcq. Le signal monogène couleur : théorie et applications. Mathématiques générales [math.GM]. Université de La Rochelle, 2010. Français. NNT : 2010LAROS319 . tel-00583117

HAL Id: tel-00583117

<https://theses.hal.science/tel-00583117>

Submitted on 4 Apr 2011

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 LA ROCHELLE
ÉCOLE DOCTORALE S2I
SCIENCES ET INGÉNIERIE POUR L'INFORMATION

THÈSE

pour obtenir le titre de

Docteur

de l'Université de La Rochelle

Spécialité : MATHÉMATIQUES ET APPLICATIONS

Présentée et soutenue par

Guillaume DEMARCQ

Le signal monogène couleur : théorie et applications

Thèse co-dirigée par Michel BERTHIER et Laurent MASCARILLA

préparée au Laboratoire Mathématiques, Image et Applications - EA 3165

soutenue le 10 décembre 2010

Jury :

Rapporteurs :

Eduardo BAYRO-CORROCHANO - Professeur, Cinvestav, Unidad Guadalajara, Mexique

Yannick BERTHOUMIEU - Professeur, Institut Polytechnique de Bordeaux

Directeurs :

Michel BERTHIER - Professeur, Université de La Rochelle

Laurent MASCARILLA - Maître de conférences, HDR, Université de La Rochelle

Examineurs :

Lionel MOISAN - Professeur, Université Paris-Descartes

Frédéric BARBARESCO - Senior Scientist & Advanced Studies Manager,
Thales Air Systems



Remerciements

Et bien voilà, nous y sommes...Le temps des merci pour les personnes, qui m'ont permis d'en arriver jusqu'ici.

C'est quand même drôle une thèse, on est à l'aise pour écrire des pages et des pages d'hypothèses, mais quand vient le moment d'écrire les bons mots, on devient lent et la verve s'en ressent.

Allez, je me lance...En premier lieu, je souhaiterais remercier Michel Berthier et Laurent Mascarilla pour leur bienveillance. Merci de m'avoir accueilli et permis d'effectuer cette thèse au sein du laboratoire MIA. Je les remercie également pour leur soutien durant ces trois années de travail, leurs avis ainsi que leur aide dans les moments les plus délicats. Un énorme merci de m'avoir laissé libre dans mes choix et de m'avoir toujours soutenu lorsque j'en avais besoin.

Je voudrais remercier également et très sincèrement les membres de mon jury : merci à Yannick Berthoumieu et Eduardo Bayro-Corrochano de s'être intéressés à mon travail et d'avoir accepté de rapporter mon manuscrit de thèse, merci à Frédéric Barbaresco d'avoir pris le temps d'examiner mon manuscrit et de me faire l'honneur de participer à mon jury et enfin merci à Lionel Moisan de s'être intéressé à mon travail et de m'avoir fait l'honneur de présider mon jury.

Il ne serait pas concevable de ne pas remercier les gens qui, au quotidien, m'ont permis de passer de très bons moments durant ces trois années au sein du labo. Je pense notamment à la présence de certaines personnes que je souhaite vivement remercier à travers ces mots ! Alors merci à Benjamin Laquerrière et Thomas Bataud pour les pause-café et les discussions qui n'ont strictement rien à voir avec la thèse, merci à mon collègue de bureau Hoel Le Capitaine avec qui les échanges sont vraiment des plus agréables et merci à Renaud Pétéri, Christophe Saint-Jean et Carl Frélicot pour les discussions fructueuses à l'Annexe. Je souhaiterais également remercier notre super-secrétaire Christine Sicard pour le travail pas toujours évident qu'elle accomplit. Enfin, merci à tous les gens que j'ai pu côtoyer durant cette thèse, je pense notamment aux autres doctorants du laboratoire (Agathe, Sloven, Mathéo, Charles, Romain, José, Nadeem et tous les autres) ainsi qu'aux copains moniteurs (Andréa, Aurélie, Thomas, Arnaud, Guillaume, Manu, Matthieu P., Antoine, Nazir).

D'un point de vue plus personnel, je tiens à remercier du fond du coeur mes parents Catherine et Jean-Louis sans qui je ne serais pas là aujourd'hui. Merci pour votre éducation, votre amour et votre soutien tout au long de ma vie et de mes études.

Merci à mon frère Benoît, à sa femme Sophie et leur fille Anouk d'être là quand il faut, tout simplement.

Merci à ma tante Anny pour les dînettes ensemble et à ma grand-mère Jacqueline pour son vif soutien.

Merci également à toute la Jolly family de m'avoir toujours encouragé dans ce que je fais.

Merci à mon meilleur ami Kévin ainsi qu'à tous mes ami(e)s en dehors de la fac, je sais que vous vous reconnaitrez !

Enfin, le meilleur pour la fin ! Un très grand MERCI à ma femme Aude, ma source d'inspiration, pour son amour, sa patience et son indéfectible soutien durant ces trois années de travail. Cette thèse, je te la dédie car sans toi, je ne sais pas si elle aurait abouti...

Table des matières

1	Introduction	5
1.1	Contexte de la thèse	5
1.2	Contenu de la thèse	6
1.3	Remarques	7
2	Algèbres de Clifford et représentations des signaux	9
2.1	Rappels sur les algèbres de Clifford	9
2.1.1	Algèbres de Clifford : définitions et propriétés.	9
2.1.2	Exemples et compléments.	12
2.1.3	Les groupes de spineurs	13
2.1.4	L'opérateur de Dirac	16
2.2	Le signal analytique	17
2.2.1	Construction dans l'algèbre de Clifford $\mathbb{R}_{2,0}$	18
2.2.2	Représentation polaire d'un signal monodimensionnel	21
2.2.3	Analyse de signaux monodimensionnels	22
2.3	Le signal monogène	22
2.3.1	Construction dans l'algèbre de Clifford $\mathbb{R}_{3,0}$	24
2.3.2	Représentation polaire d'une image en niveau de gris	27
2.3.3	Analyse de signaux bidimensionnels	28
2.4	Le signal monogène couleur	30
2.4.1	Système colorimétrique	31
2.4.2	Construction dans l'algèbre de Clifford $\mathbb{R}_{5,0}$	31
2.4.3	Représentation polaire d'une image couleur	35
2.4.4	Fonction d'angle φ_V et représentation polaire χ	41
3	Applications	47
3.1	Flot optique couleur	48
3.1.1	Introduction	48
3.1.2	Méthode CLG (Combined Local-Global)	49
3.1.3	Méthode CLG modifiée	50
3.1.4	Résultats	50
3.1.5	Nouvelle contrainte	51
3.1.6	Résultats	53
3.2	Segmentation multi-échelle d'images couleur	55
3.2.1	Détection de contours métrique	55
3.2.2	Méthode de segmentation	56
3.2.3	Résultats	57
3.2.4	Détection de contours métrique et multi-échelle	57
3.2.5	Méthode de segmentation	60
3.2.6	Segmentation multi-échelle	61

3.2.7	Résultats	63
3.3	Suivi d'objets couleur	65
3.3.1	Un algorithme naïf de block-matching	65
3.3.2	Filtrage particulière	68
3.3.3	Résultats	71
3.4	Détection de points d'intérêt	74
3.4.1	Détecteur de Harris et géométrie différentielle	74
3.4.2	Géométrie différentielle dans un espace-échelle	76
3.4.3	Résultats	77
3.4.4	Complément 1 : H_{3D} couleur	87
3.4.5	Complément 2 : H_{3D} et signal monogène couleur	88
4	Conclusion	91
A	Espace-échelle de Poisson	93
A.1	Définition	93
A.2	Espace-échelle de Poisson	94
B	Noyaux et transformée de Fourier dans $\mathbb{R}_{2,0}$	95
B.1	Transformée de Fourier	95
B.2	Noyau de Hilbert	95
B.3	Noyau de Poisson	95
B.4	Noyau de Poisson conjugué	96
C	Noyaux et transformée de Fourier dans $\mathbb{R}_{3,0}$	97
C.1	Transformée de Fourier	97
C.2	Noyau de Riesz	97
C.3	Noyau de Poisson	98
C.4	Noyau de Poisson conjugué	98
D	Noyaux et transformée de Fourier dans $\mathbb{R}_{7,0}$	99
D.1	Transformée de Fourier	99
D.2	Transformée de Riesz	99
D.3	Noyau de Poisson	100
D.4	Noyau de Poisson conjugué	100
E	Analyse de la complexité	101
E.1	Signal Monogène Couleur	101
E.2	Représentation Polaire	101
E.3	Segmentation Multi-échelle	102
E.4	Flot Optique Couleur	102
F	Algorithme de Hostenes et Sobczyk	103
	Bibliographie	105

Introduction

Sommaire

1.1	Contexte de la thèse	5
1.2	Contenu de la thèse	6
1.3	Remarques	7

1.1 Contexte de la thèse

Cette thèse traite de l'analyse et du traitement d'images ou de séquences d'images couleur. Par delà les approches marginales usuellement employées, nous avons cherché à dégager une approche spécifique exploitant au maximum les informations colorimétriques. Notre démarche s'inscrit dans la continuité des travaux entrepris par S. Sangwine *et al.* [Sangwine 2007] utilisant notamment une représentation de l'espace couleur RGB par l'ensemble des quaternions imaginaires purs. Il s'avère qu'un cadre mathématique particulièrement adéquat pour appréhender la géométrie des couleurs est celui des algèbres de Clifford dont l'algèbre des quaternions est un exemple [Batard 2009, Sangwine 2007]. Les éléments d'une telle algèbre sont appelés multivecteurs ; ils représentent des "portions" de sous-espaces vectoriels de dimension arbitraire avec une norme et une orientation. L'un des intérêts, très exploité en vision et en robotique, est la possibilité de coder algébriquement des transformations géométriques à l'aide du calcul dit sans coordonnées.

La représentation des images couleurs que nous privilégions dans ce qui suit est basée sur une généralisation du signal analytique. Là encore, comme cela a été mis en évidence notamment par G. Sommer et M. Felsberg [Felsberg 2001], le cadre des algèbres de Clifford se révèle être particulièrement pertinent. La construction proposée du signal monogène couleur repose sur l'analogie qui peut être faite entre les conditions de Cauchy Riemann, qui définissent le caractère holomorphe d'une fonction, et l'équation de Dirac, faisant intervenir l'opérateur du même nom. La résolution des systèmes qui en découlent utilisent les noyaux de Riesz et de Poisson 2D. Il est bien connu que l'un des intérêts du signal analytique est la possibilité de définir la phase et l'amplitude instantanées d'un signal réel. Nous discutons, pour le signal monogène couleur, d'une représentation polaire, basée sur un produit géométrique, qui s'inscrit dans la même voie.

Les applications envisagées reposent majoritairement sur cette représentation polaire et sur les informations de couleur, mais nous le verrons aussi de structures

locales, s'y rattachant. Nous abordons successivement des problématiques liées au flot optique couleur, à la segmentation couleur multi-échelle, au suivi d'objets couleur et à la détection de points d'intérêt.

Mentionnons également que les travaux décrits dans cette thèse s'inscrivent dans l'approche "scale-space" du traitement d'images. Le signal monogène couleur est par construction multi-échelle pour le noyau de Poisson. Cet aspect est très présent dans les applications évoquées ci-dessus.

1.2 Contenu de la thèse

Nous avons distingué deux parties, l'une théorique, l'autre plus applicative. Nous commençons la première partie par des notions de base sur les algèbres de Clifford et les groupes de spineurs. Il existe somme toute peu de références liées au traitement d'images et de la vidéo où l'on trouve explicitées les définitions et les propriétés qui nous seront utiles dans la suite. Nous avons opté pour une présentation mathématique rigoureuse, parfois assez abstraite, illustrée par des exemples classiques. Nous décrivons ensuite les constructions des signaux analytique et monogène (cette dernière due à M. Felsberg) impliquant les algèbres de Clifford $\mathbb{R}_{2,0}$ et $\mathbb{R}_{3,0}$ ainsi que les opérateurs de Dirac associés. La fin de la première partie est consacrée à la construction du signal monogène couleur et à la représentation polaire qui en découle. Sans entrer plus dans les détails, signalons simplement que cette représentation polaire est obtenue à partir du produit géométrique dans $\mathbb{R}_{5,0}$ du signal monogène couleur avec un vecteur de référence contenant des informations de structures locales et de couleur fixées. Un tel produit est traditionnellement appelé un 2-verseur. On extrait de ce 2-verseur une fonction d'angle et une partie scalaire sur l'utilisation desquelles reposent les applications proposées dans la suite. Nous discutons également sur des exemples de l'interprétation en termes de phase de cette fonction d'angle. Nous verrons que beaucoup reste à faire dans cette optique, ce qui ouvre de nombreuses perspectives au travail présenté dans ce mémoire. Les propriétés des transformées et des noyaux utilisés sont décrites en annexe.

La première application abordée dans la seconde partie traite de l'extraction d'un flot optique couleur. Il ne s'agit pas ici de trouver un champ de vecteurs dense qui serait meilleur que les champs de vecteurs existants. Notre approche est différente des approches usuelles ; elle consiste à extraire, à partir de la fonction d'angle de la représentation polaire, le champ de vecteurs des déplacements d'objets d'une couleur fixée. La contrainte de conservation de l'intensité est remplacée par la contrainte de conservation de la fonction d'angle. Comme dans toutes les autres applications, on peut tirer profit du caractère multi-échelle. Deux exemples sont traités, l'un présentant de fortes variations de luminosité, l'autre étant très dégradé par la compression et les mouvements de caméra.

Nous nous intéressons ensuite à un problème de segmentation couleur multi-échelle. Pour cela nous considérons tout d'abord la surface de \mathbb{R}^4 paramétrée par la partie scalaire et la fonction de phase de la représentation polaire. La détection

des variations dans une couleur prescrite se fait via l'étude de la première forme fondamentale de la surface, plus précisément des valeurs propres de cette dernière. Une variante est proposée où le paramètre d'échelle est traité comme une variable supplémentaire, la surface précédente étant remplacée par une variété de dimension 3. La même analyse des valeurs propres de la première forme fondamentale correspondante conduit à une détection multi-échelle avec une détermination automatique du paramètre d'échelle significatif ; c'est ce dernier point qui justifie cette variante.

La troisième application est consacrée au suivi d'objets. La fonction d'angle est une nouvelle fois utilisée mais cette fois-ci comme mesure de similarité. Nous commençons par décrire une approche naïve de type "block-matching". Dans un second temps, nous incorporons l'information de similarité dans la fonction de vraisemblance d'un algorithme de filtrage particulière. Ces deux méthodes nous ont permis de valider la robustesse de la fonction d'angle aux changements d'illuminant ainsi que sa capacité à encoder l'information de couleur et de structure de façon pertinente.

Enfin, et en marge de ces travaux, nous abordons le problème de la détection des points d'intérêt suivant une approche multi-échelle. Nous commençons par resituer la définition du détecteur de Harris dans le cadre de la géométrie différentielle en faisant le lien entre ce dernier et une version "relaxée" du discriminant du polynôme caractéristique de la première forme fondamentale. La version multi-échelle fait intervenir une nouvelle fois la variété de dimension 3 obtenue en traitant le paramètre d'échelle comme une variable. Les détecteurs proposés sont définis à partir de l'étude des bifurcations des racines (à l'aide des formules de Cardan) d'une version "relaxée" du discriminant de Cardan du polynôme caractéristique de la première forme fondamentale de la variété.

1.3 Remarques

Comme nous l'avons dit précédemment, un des objectifs de cette thèse est d'aborder certaines des problématiques du traitement des images et de la vidéo en mettant fortement l'accent sur l'utilisation des informations colorimétriques. L'approche privilégiée, basée sur la définition et l'utilisation du signal monogène couleur, est nouvelle et n'a donc fait l'objet d'aucun travail antérieur. Nous le verrons au cours du texte, de nombreuses questions restent en suspens, notamment en ce qui concerne la définition de phase pour des images couleur. Les applications décrites montrent que cette approche donnent de bons résultats dans divers contextes. Clairement, les résultats contenus dans ce mémoire de thèse sont une première contribution à l'étude d'un "objet" complexe qui semble laisser entrevoir de nombreuses potentialités.

Algèbres de Clifford et représentations des signaux

Nous commençons par rappeler les définitions et les propriétés de base concernant les algèbres de Clifford et les groupes de spineurs. Les paragraphes 2.2 et 2.3. sont consacrés aux constructions des signaux analytique et monogène et à l'analyse de ces derniers. Le reste du chapitre concerne les deux principaux objets d'étude de notre travail à savoir le signal monogène couleur et la représentation polaire associée. Le lecteur trouvera en annexe les propriétés des transformées et des noyaux utilisés.

Sommaire

2.1	Rappels sur les algèbres de Clifford	9
2.1.1	Algèbres de Clifford : définitions et propriétés.	9
2.1.2	Exemples et compléments.	12
2.1.3	Les groupes de spineurs	13
2.1.4	L'opérateur de Dirac	16
2.2	Le signal analytique	17
2.2.1	Construction dans l'algèbre de Clifford $\mathbb{R}_{2,0}$	18
2.2.2	Représentation polaire d'un signal monodimensionnel	21
2.2.3	Analyse de signaux monodimensionnels	22
2.3	Le signal monogène	22
2.3.1	Construction dans l'algèbre de Clifford $\mathbb{R}_{3,0}$	24
2.3.2	Représentation polaire d'une image en niveau de gris	27
2.3.3	Analyse de signaux bidimensionnels	28
2.4	Le signal monogène couleur	30
2.4.1	Système colorimétrique	31
2.4.2	Construction dans l'algèbre de Clifford $\mathbb{R}_{5,0}$	31
2.4.3	Représentation polaire d'une image couleur	35
2.4.4	Fonction d'angle φ_V et représentation polaire χ	41

2.1 Rappels sur les algèbres de Clifford

2.1.1 Algèbres de Clifford : définitions et propriétés.

Les données de base sont un espace vectoriel E (sur \mathbb{R} pour ce qui nous intéresse), de dimension finie, et une forme quadratique Q sur E . Le but est de définir (et de

construire) une algèbre sur \mathbb{R} qui contienne E et telle que si u est un vecteur de E , le produit $uu = u^2$ (qui a un sens dans l'algèbre considérée) soit égal à la valeur $Q(u)$ de la forme quadratique évaluée sur u . Par suite, si u et v sont deux vecteurs de E plongés dans une telle algèbre, alors nécessairement : $uv + vu = 2B(u, v)$, en notant B la forme bilinéaire symétrique associée à Q .

Le **produit géométrique** des deux vecteurs u et v est alors défini par la formule :

$$uv = B(u, v) + u \wedge v$$

où

$$u \wedge v = \frac{1}{2}(uv - vu)$$

est la partie antisymétrique de ce dernier. On peut remarquer rapidement que l'algèbre recherchée, si elle se doit de contenir les scalaires de \mathbb{R} et les vecteurs de E , doit également contenir d'autres types d'objets. Par exemple, si E est l'espace vectoriel \mathbb{R}^2 , Q la forme quadratique euclidienne et (e_1, e_2) une base orthonormée, alors l'élément $e_1 \wedge e_2$ est de carré -1 ; il ne peut donc pas s'agir d'un scalaire ou d'un vecteur.

Il faut cependant prendre quelques précautions si l'on veut être rigoureux. Il existe en effet beaucoup d'algèbres A munies d'une application linéaire f de E dans A satisfaisant la condition $[f(u)]^2 = Q(u)$ pour tout vecteur u de E . De tels couples (A, f) sont les objets d'une catégorie et l'algèbre de Clifford de E relativement à la forme quadratique Q , notée dorénavant $(Cl(E, Q), i_{(E, Q)})$, est définie comme étant un objet initial de cette catégorie. C'est dire précisément que si (A, f) est un couple comme ci-dessus, il existe un unique morphisme d'algèbres g de $Cl(E, Q)$ dans A tel que $f = g \circ i_{(E, Q)}$. Comme tout objet initial d'une catégorie, le couple $(Cl(E, Q), i_{(E, Q)})$ est défini à isomorphisme (de la catégorie) près. Nous parlerons cependant, avec un léger abus, de l'algèbre de Clifford de E .

Remarque : L'application $i_{(E, Q)}$ est là pour envoyer E dans $Cl(E, Q)$. Nous verrons plus loin que cette application est une injection et qu'elle permet d'identifier par conséquent un vecteur u de E et son image $i_{(E, Q)}(u)$ dans $Cl(E, Q)$ (ce que nous avons fait au début de ce paragraphe).

Pour construire pratiquement l'algèbre de Clifford de E , on procède de la façon suivante. Soit :

$$T(E) = \mathbb{R} \oplus E \oplus (E \otimes E) \oplus \dots = \bigoplus_{i \geq 0} E^{\otimes i}$$

l'algèbre tensorielle¹ de E , et soit $\mathcal{I}(E, Q)$ l'idéal de $T(E)$ engendré par les éléments de la forme $u \otimes u - Q(u)$, u appartenant à E . Notons encore $i_{(E, Q)}$ l'application composée :

$$E \longrightarrow T(E) \longrightarrow T(E)/\mathcal{I}(E, Q)$$

Il n'est pas trop difficile de vérifier (voir [Postnikov 1985]) que :

1. Si F et G sont deux espaces vectoriels muni chacun d'une base $\{f_i\}_{i \in I}$ et $\{g_j\}_{j \in J}$, le produit tensoriel $F \otimes G$ de F et de G est l'espace vectoriel admettant pour base les produits formels $\{f_i g_j\}_{(i, j)}$. On note traditionnellement $f_i \otimes g_j$ le produit $f_i g_j$. Si $u = \sum \alpha_i f_i$ et $v = \sum \beta_j g_j$ alors $u \otimes v = \sum \alpha_i \beta_j f_i \otimes g_j$.

Théorème 1. *Le couple $(T(E)/\mathcal{I}(E, Q), i_{(E, Q)})$ est l'algèbre de Clifford de E .*

Par ailleurs, et c'est fondamental dans tout ce qui suit :

Théorème 2. *l'application $i_{(E, Q)}$ est injective et l'ensemble $\{e_{i_1} e_{i_2} \dots e_{i_k} ; i_1 < i_2 < \dots < i_k ; k \in \{1, 2, \dots, n\}\}$ avec l'unité 1 de $Cl(E, Q)$ est une base² de $Cl(E, Q)$.*

On en déduit en particulier que l'algèbre de Clifford de E est de dimension 2^n (si n est la dimension de E). La démonstration de ces deux propriétés repose sur le fait que $Cl(E, Q)$ est une algèbre \mathbb{Z}_2 -graduée. En voici les grandes lignes. On commence tout d'abord par utiliser le fait que $(Cl(E, Q), i_{(E, Q)})$ est un objet initial (unique à isomorphisme près) pour montrer qu'il existe un unique automorphisme α de $Cl(E, Q)$ qui satisfait : $\alpha \circ \alpha = Id$ et $\alpha \circ i_{(E, Q)} = \alpha_0$ où α_0 est l'application linéaire de E dans $Cl(E, Q)$ qui à un vecteur u associe l'élément $-i_{(E, Q)}(u)$. On utilise ensuite cet automorphisme pour décomposer l'algèbre $Cl(E, Q)$ en la somme directe $Cl(E, Q) = Cl(E, Q)^{(0)} \oplus Cl(E, Q)^{(1)}$ avec :

$$Cl(E, Q)^{(0)} = \{x \in Cl(E, Q), \alpha(x) = x\} \quad (\text{éléments pairs})$$

$$Cl(E, Q)^{(1)} = \{x \in Cl(E, Q), \alpha(x) = -x\} \quad (\text{éléments impairs})$$

De plus : $Cl(E, Q)^{(i)} Cl(E, Q)^{(j)} \subset Cl(E, Q)^{(i+j)}$, i et j étant des entiers modulo 2, ce qui signifie précisément que

Proposition 1. *L'algèbre $Cl(E, Q)$ est une algèbre \mathbb{Z}_2 -graduée.*

Avant d'aller plus loin, rappelons ce qu'est le **produit tensoriel gauche** de deux algèbres \mathbb{Z}_2 -graduées, $A = A^{(0)} \oplus A^{(1)}$ et $B = B^{(0)} \oplus B^{(1)}$. Il s'agit du produit tensoriel $A \otimes B = (A \otimes B)^{(0)} \oplus (A \otimes B)^{(1)}$, avec $(A \otimes B)^{(0)} = (A^{(0)} \otimes B^{(0)}) \oplus (A^{(1)} \otimes B^{(1)})$ et $(A \otimes B)^{(1)} = (A^{(0)} \otimes B^{(1)}) \oplus (A^{(1)} \otimes B^{(0)})$, ce produit tensoriel étant muni du produit gauche :

$$(a \otimes b)(c \otimes d) = (-1)^{ij}(ac \otimes bd)$$

avec $b \in B^{(i)}$ et $c \in A^{(j)}$. On note le produit tensoriel gauche de A et B , $A \widehat{\otimes} B$. L'intérêt de cette construction pour notre étude réside dans le résultat suivant. Soient E_1 et E_2 deux espaces vectoriels munis respectivement d'une forme quadratique Q_1 et Q_2 , soit encore $E_1 \oplus E_2$ la somme directe de ces deux espaces munie de la forme quadratique³ $Q_1 \oplus Q_2$, alors :

Proposition 2. *L'algèbre de Clifford $Cl(E_1 \oplus E_2, Q_1 \oplus Q_2)$ est isomorphe au produit tensoriel gauche $Cl(E_1, Q_1) \widehat{\otimes} Cl(E_2, Q_2)$.*

Connaissant ce résultat, on peut "dévisser" l'algèbre de Clifford d'un espace vectoriel E muni d'une forme quadratique Q en invoquant le théorème d'inertie de Sylvester (classification des formes quadratiques sur \mathbb{R}). L'algèbre $Cl(E, Q)$ est le produit tensoriel gauche de n algèbres élémentaires isomorphes chacune à l'un des modèles suivants : $\mathbb{R}[x]/(x^2 + \varepsilon)$, avec $\varepsilon = -1, 0$ ou 1 . On en déduit les deux propriétés citées plus haut (voir [Postnikov 1985] p. 223).

2. Les e_i sont les éléments d'une base Q -orthogonale de E .

3. Cette forme quadratique est définie par $Q_1 \oplus Q_2(u + v) = Q_1(u) + Q_2(v)$.

2.1.2 Exemples et compléments.

Dans ce qui suit, on note $\mathbb{R}_{p,q}$ (avec $n = p + q$) l'algèbre de Clifford de l'espace vectoriel \mathbb{R}^n muni de la forme quadratique :

$$Q(x_1, x_2, \dots, x_n) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_n^2$$

Premier exemple :

L'exemple le plus simple est celui de l'algèbre $\mathbb{R}_{0,1}$, *i.e.* l'algèbre de Clifford de \mathbb{R} muni de la forme quadratique $Q(x) = -x^2$. L'algèbre tensorielle de \mathbb{R} étant isomorphe à $\mathbb{R}[x]$, $\mathbb{R}_{0,1}$ est isomorphe au quotient $\mathbb{R}[x]/(x^2 + 1)$ c'est à dire à \mathbb{C} . Si e est un vecteur de \mathbb{R} de norme euclidienne 1, alors $e^2 = -1$. Le couple $(1, e)$ est une base de $\mathbb{R}_{0,1}$ et on réalise un isomorphisme entre cette dernière algèbre et \mathbb{C} en envoyant 1 sur 1 et e sur i . L'algèbre $\mathbb{R}_{1,0}$ est quant à elle isomorphe au quotient $\mathbb{R}[x]/(x^2 - 1)$.

Deuxième exemple :

Soit (e_1, e_2) une base orthonormée⁴ de \mathbb{R}^2 . Dans $\mathbb{R}_{0,2}$, on a les relations suivantes : $e_1^2 = e_2^2 = -1$. Il est facile de vérifier que l'on réalise un isomorphisme entre $\mathbb{R}_{0,2}$ et \mathbb{H} en envoyant 1 sur 1, e_1 sur i , e_2 sur j et e_1e_2 sur k . Les éléments pairs, de la forme $z = x1 + ye_1e_2$, forment une sous-algèbre de $\mathbb{R}_{0,2}$ isomorphe à \mathbb{C} . Puisque $(e_1e_2)e_1 = e_2$ et $(e_1e_2)e_2 = -e_1$, le produit à gauche par (e_1e_2) agit comme une rotation d'angle $\pi/2$ (le produit à droite agit comme une rotation d'angle $-\pi/2$). On montre aisément, en considérant les matrices :

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

que les deux algèbres $\mathbb{R}_{1,1}$ et $\mathbb{R}_{2,0}$ sont isomorphes à l'algèbre $\mathbb{R}(2)$ des matrices 2×2 à coefficients réels.

Troisième exemple :

Soit (e_1, e_2, e_3) une base orthonormée de \mathbb{R}^3 . L'algèbre $\mathbb{R}_{3,0}$ est de dimension 8 et admet pour base la famille $(1, e_1, e_2, e_3, e_1e_2, e_2e_3, e_1e_3, e_1e_2e_3)$. Les éléments pairs, de la forme $q = q_01 + q_1e_1e_2 + q_2e_2e_3 + q_3e_1e_3$, forment une sous-algèbre de $\mathbb{R}_{3,0}$ isomorphe à \mathbb{H} ; on constate en effet que $(e_1e_2)^2 = (e_2e_3)^2 = (e_1e_3)^2 = -1$. De la même manière, tous les éléments de la forme $q = q_01 + pB$ pour un bivecteur $B = q_1e_1e_2 + q_2e_2e_3 + q_3e_1e_3$ de norme 1 fixé, forment une

4. La base est orthonormée pour le produit scalaire euclidien.

sous-algèbre de $\mathbb{R}_{3,0}$ isomorphe à \mathbb{C} ; on constate en effet que $B^2 = -1$.

Considérons à nouveau le cas général d'un espace vectoriel E muni d'une forme quadratique Q . Nous avons vu plus haut qu'il existait un unique automorphisme α de $Cl(E, Q)$ satisfaisant certaines conditions. De la même façon, il existe un unique anti-automorphisme t de $Cl(E, Q)$ tel que : $t \circ t = Id$ et $t \circ i_{(E, Q)}(u) = i_{(E, Q)}(u)$ pour tout vecteur u de E . Pratiquement, les applications α et t sont définies sur les éléments de la base de $Cl(E, Q)$ par :

$$\begin{aligned} \alpha(1) &= 1 & \alpha(e_i) &= -e_i & \alpha(e_{i_1} e_{i_2} \dots e_{i_k}) &= (-1)^k e_{i_1} e_{i_2} \dots e_{i_k} \\ t(1) &= 1 & t(e_i) &= e_i & t(e_{i_1} e_{i_2} \dots e_{i_k}) &= e_{i_k} e_{i_{k-1}} \dots e_{i_1} \end{aligned}$$

A partir de α et de t , on construit une **conjugaison sur** $Cl(E, Q)$: le conjugué d'un élément x est $\bar{x} = t \circ \alpha(x)$, puis une **norme** : $N(x) = x\bar{x}$.

On a en particulier : $N(u) = -u^2 = -Q(u)$ pour tout vecteur u de E , et $N(e_{i_1} e_{i_2} \dots e_{i_k}) = (-1)^k Q(e_{i_1}) Q(e_{i_2}) \dots Q(e_{i_k})$.

Nous allons terminer ce paragraphe par quelques exemples de formules de calcul sans coordonnées. Remarquons tout d'abord que si n est un vecteur de E tel que $Q(n) \neq 0$, alors n est inversible, d'inverse $n^{-1} = n/n^2 = n/Q(n)$. Si maintenant u est un vecteur de E , on dispose d'une décomposition en partie parallèle et orthogonale relativement à n : $u = u_{\parallel} + u_{\perp}$, avec $u_{\parallel} \wedge n = 0$ et $u_{\perp} \cdot n = 0$ (on note \cdot la forme bilinéaire associée à la forme quadratique Q). On montre facilement que $u_{\perp} = (u \wedge n)n^{-1}$ et $u_{\parallel} = (u \cdot n)n^{-1}$.

On en déduit que la réflexion par rapport à l'hyperplan orthogonal à un vecteur inversible n de E est donnée par la formule : $s_n(u) = -nun^{-1}$. De même le symétrique d'un vecteur u par rapport à n est donné par $\sigma_n(u) = nun^{-1}$.

2.1.3 Les groupes de spineurs

2.1.3.1 Définitions et propriétés

Commençons par un exemple simple : soient n_1 et n_2 deux vecteurs unitaires de \mathbb{R}^3 , comment écrire dans le langage des algèbres de Clifford la rotation qui envoie n_1 sur n_2 en laissant le plan engendré par n_1 et n_2 globalement invariant ? On constate tout d'abord que : $n_2(1 + n_2 n_1) = n_2 + n_1 = (1 + n_2 n_1)n_1$ (car $n_2^2 = n_1^2 = 1$). Par ailleurs la norme de l'élément $(1 + n_2 n_1)$ est $2(1 + n_2 \cdot n_1)$. Finalement, l'élément pair de norme 1 (et par conséquent inversible) :

$$\tau = \frac{(1 + n_2 n_1)}{\sqrt{2(1 + n_2 \cdot n_1)}}$$

satisfait $\tau n_1 \tau^{-1} = n_2$. La rotation cherchée s'écrit $u \mapsto \tau u \tau^{-1}$.

Notons θ l'angle entre n_1 et n_2 ; en écrivant $2(1 + n_2.n_1) = 4 \cos^2(\theta/2)$, on vérifie aisément que :

$$\tau = \cos(\theta/2) + \sin(\theta/2) \frac{n_2 \wedge n_1}{|n_2 \wedge n_1|} = \exp \left[(\theta/2) \frac{n_2 \wedge n_1}{|n_2 \wedge n_1|} \right]$$

i.e. l'élément τ est l'exponentielle⁵ d'un bivecteur, nommément $(\theta/2)(n_2 \wedge n_1 / |n_2 \wedge n_1|)$.

On considère l'algèbre $\mathbb{R}_{n,0}$, de sorte que $N(u) = -u^2 = -Q(u) \leq 0$, pour tout vecteur u de \mathbb{R}^n . Le groupe $Spin(n)$ est défini de la façon suivante :

$$Spin(n) = \{x \in \mathbb{R}_{n,0}^{(0)}, xvx^{-1} \in \mathbb{R}^n \quad \forall v \in \mathbb{R}^n, N(x) = 1\}$$

Quelques remarques⁶. Comme ci-dessus $\mathbb{R}_{n,0}^{(0)}$ désigne l'ensemble des éléments pairs de l'algèbre $\mathbb{R}_{n,0}$. Pour un tel élément, la condition $N(x) = 1$ équivaut⁷ à $xt(x) = 1$. Cette dernière condition implique en particulier que l'élément x est inversible dans $\mathbb{R}_{n,0}$. L'objet qui nous intéresse dans ce qui suit est la représentation :⁸

$$\rho : Spin(n) \longrightarrow GL(n; \mathbb{R})$$

qui à x associe l'application $\rho(x) = \rho_x$ de \mathbb{R}^n dans \mathbb{R}^n définie par $\rho_x(v) = xvx^{-1}$. Comme $N(xvx^{-1}) = N(v)$, la représentation ρ est à valeurs dans $O(n, \mathbb{R})$. Nous allons démontrer le résultat suivant :

Théorème 3. *La représentation ρ est un morphisme de groupes surjectif de $Spin(n)$ dans $SO(n, \mathbb{R})$ de noyau isomorphe à \mathbb{Z}_2 .*

Commençons dans un premier temps par vérifier que le noyau de ρ est isomorphe à \mathbb{Z}_2 . Soit donc x un élément pair de $\mathbb{R}_{n,0}$ tel que $xv = vx$ pour tout vecteur v de \mathbb{R}^n . On peut écrire : $x = a_0 + e_1 b_1$ avec $a_0 \in \mathbb{R}_{n,0}^{(0)}$, $b_1 \in \mathbb{R}_{n,0}^{(1)}$, et ni a_0 ni b_1 ne contient le facteur e_1 . En choisissant $v = e_1$, on obtient : $a_0 e_1 + e_1 b_1 e_1 = e_1 a_0 + e_1^2 b_1$. La base (e_1, e_2, \dots, e_n) étant orthonormée, on a $e_i e_j = -e_j e_i$ pour $i \neq j$. On en déduit en particulier que $a_0 e_1 = e_1 a_0$ et $e_1 b_1 e_1 = -e_1^2 b_1$. Finalement,

$$a_0 e_1 + e_1 b_1 e_1 = e_1 a_0 + e_1^2 b_1 = e_1 a_0 - e_1^2 b_1$$

d'où $e_1^2 b_1 = 0$ et $b_1 = 0$. Ainsi l'élément x ne contient pas de terme impliquant e_1 . On montre de la même manière qu'il ne contient pas de terme impliquant e_2, e_3, \dots, e_n . Ceci signifie que x est un scalaire de $\mathbb{R}_{n,0}$, de norme 1, c'est à dire est soit 1, soit -1 .

5. Nous reviendrons plus tard sur l'application exponentielle.

6. Dans ce qui suit (e_1, e_2, \dots, e_n) est une base orthonormée (pour le produit scalaire euclidien) de l'espace \mathbb{R}^n .

7. Certains auteurs appellent *rotors* ce que nous appelons ici spineurs

8. Ceci signifie que ρ est un morphisme de groupes, ce qui est facilement vérifiable.

La deuxième étape consiste à s'assurer que ρ est à valeurs dans $SO(n, \mathbb{R})$. Sinon, il existe f , impropre, appartenant à $O(n, \mathbb{R})$ et x appartenant à $Spin(n)$ tels que $\rho_x = f$. D'après le théorème de Cartan-Dieudonné, il existe un nombre impair de vecteurs unitaires, $w_1, w_2, \dots, w_{2k-1}$, tels que :

$$f(v) = -(w_1 w_2 \dots w_{2k-1}) v (w_1 w_2 \dots w_{2k-1})^{-1}$$

L'élément $y = x^{-1}(w_1 \dots w_{2k-1})$ satisfait donc $-yv = vy$ pour tout vecteur v de \mathbb{R}^n . Comme précédemment, décomposons l'élément impair y en $y = a_1 + e_1 b_0$ avec $a_1 \in \mathbb{R}_{n,0}^{(1)}$, $b_0 \in \mathbb{R}_{n,0}^{(0)}$, et ni a_1 ni b_0 ne contient le facteur e_1 . En choisissant encore pour vecteur v le vecteur e_1 , on vérifie que b_0 est nul, donc que y ne contient pas de terme faisant intervenir e_1 . De même y ne contient pas de terme faisant intervenir e_2, \dots, e_n ; par suite y est un scalaire. Mais alors $x = \lambda w_1 w_2 \dots w_{2k-1}$ (λ étant réel), est un élément impair de $\mathbb{R}_{n,0}$, d'où une contradiction.

Il ne reste plus qu'à vérifier que ρ est effectivement surjective, ce qui découle trivialement du théorème de Cartan-Dieudonné⁹ [Gallier 2000].

2.1.3.2 Exemples.

Le groupe $Spin(2)$ est constitué des éléments x pairs et de norme 1 de l'algèbre $\mathbb{R}_{2,0}$ vérifiant : xvx^{-1} est un vecteur de \mathbb{R}^2 pour tout vecteur v de \mathbb{R}^2 . Les éléments pairs de $\mathbb{R}_{2,0}$ sont de la forme $x = x_0 1 + x_2 e_1 e_2$ avec $N(x) = x_0^2 + x_2^2$. Pour un tel élément, la condition xvx^{-1} est un vecteur est automatiquement satisfaite : si v est un vecteur de \mathbb{R}^2 alors xvx^{-1} est un élément impair de $\mathbb{R}_{2,0}$, donc un vecteur. Il s'ensuit que :

Proposition 3. *Le groupe $Spin(2)$ est isomorphe au groupe $U(1)$ des nombres complexes de module 1.*

On vérifie facilement que si $x = \cos \theta 1 + \sin \theta e_1 e_2$ alors xvx^{-1} est l'image de v par la rotation de \mathbb{R}^2 d'angle 2θ .

Tout élément pair, x , de l'algèbre $\mathbb{R}_{3,0}$ est de la forme $x = x_0 1 + x_1 e_1 e_2 + x_2 e_2 e_3 + x_3 e_1 e_3$ avec $N(x) = x_0^2 + x_1^2 + x_2^2 + x_3^2$. De même que précédemment, on va vérifier que si v est un vecteur de \mathbb{R}^3 alors xvx^{-1} est nécessairement un vecteur de \mathbb{R}^3 . Pour ce faire, on écrit $xvx^{-1} = u + \lambda \overline{e_1 e_2 e_3}$ avec u appartenant à \mathbb{R}^3 et λ appartenant à \mathbb{R} . On constate par ailleurs que $\overline{xvx^{-1}} = -xvx^{-1}$ car $\overline{v} = -v$ et $\overline{x} = x^{-1}$. De plus $\overline{e_1 e_2 e_3} = e_1 e_2 e_3$, d'où $xvx^{-1} = -u + \lambda e_1 e_2 e_3$. On en déduit que λ vaut 0. Il s'ensuit que

9. Deux remarques s'imposent :

- Si l'on sait que le groupe $Spin(n)$ est connexe alors le fait que ρ soit à valeurs dans $SO(n, \mathbb{R})$ est immédiat.

- Pour la surjectivité, on peut invoquer un résultat général : un morphisme de groupes de Lie connexes $\phi : G \rightarrow H$ est un revêtement de groupes de Lie (donc en particulier est surjectif) si G et H sont de même dimension et si le noyau de ϕ est discret. Nous avons vérifié que le noyau de ρ est discret ; il faudrait s'assurer une nouvelle fois que $Spin(n)$ est connexe et que sa dimension est $n(n-1)/2$.

Proposition 4. *Le groupe $Spin(3)$ est isomorphe au groupe \mathbb{H}^1 des quaternions de module 1.*

Tout élément pair, x , de l'algèbre $\mathbb{R}_{4,0}$ est de la forme $x = x_01 + x_1e_1e_2 + x_2e_1e_3 + x_3e_1e_4 + x_4e_2e_3 + x_5e_2e_4 + x_6e_3e_4 + x_7e_1e_2e_3e_4$. On note

$$i = e_1e_2 \quad j = e_2e_3 \quad k = e_1e_3 \quad i' = e_4e_3 \quad j' = e_4e_1 \quad k' = e_4e_2 \quad \mathbb{I} = e_1e_2e_3e_4$$

d'où les relations :

$$ij = k \quad jk = i \quad ki = j \quad i^2 = -1 \quad j^2 = -1 \quad k^2 = -1 \\ i\mathbb{I} = \mathbb{I}i = i' \quad j\mathbb{I} = \mathbb{I}j = j' \quad k\mathbb{I} = \mathbb{I}k = k'$$

On vérifie une nouvelle fois¹⁰ que si v est un vecteur de \mathbb{R}^4 et x est un élément pair de $\mathbb{R}_{4,0}$, alors xvx^{-1} est un vecteur de \mathbb{R}^4 . On en déduit que tout élément x de $Spin(4)$ s'écrit sous la forme

$$x = u + \mathbb{I}v$$

où u et v sont des quaternions tels que $u\bar{u} + v\bar{v} = 1$ et $u\bar{v} + v\bar{u} = 0$. Finalement, l'application qui à $u + \mathbb{I}v$ associe $(u + v, u - v)$ est un isomorphisme, d'où

Proposition 5. *Le groupe $Spin(4)$ est isomorphe au groupe $Spin(3) \times Spin(3)$.*

2.1.4 L'opérateur de Dirac

L'opérateur de Dirac est défini sur $\mathbb{R}_{n,0}$ par :

$$D_n = \sum_{k=1}^n e_k \frac{\partial}{\partial x_k}$$

où pour tout i et j dans $\{1, \dots, n\}$, $e_i e_j + e_j e_i = 2\delta_{ij}$.

L'équation $D_n f = 0$ peut être vu comme une généralisation des équations de Cauchy-Riemann pour la raison suivante. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}_{2,0}$ définie par

$$f(x, y) = f_1(x, y)e_1 + f_2(x, y)e_2$$

En appliquant l'opérateur de Dirac

$$D_2 = e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y}$$

à f , nous obtenons :

$$D_2 f(x, y) = D_2 \cdot f(x, y) + D_2 \wedge f(x, y) \\ = \frac{\partial f_1}{\partial x}(x, y) + \frac{\partial f_2}{\partial y}(x, y) + e_{12} \left[\frac{\partial f_2}{\partial x}(x, y) - \frac{\partial f_1}{\partial y}(x, y) \right]$$

Par conséquent, il apparaît que l'équation de Dirac $D_2 f = 0$ dans $\mathbb{R}_{2,0}$ est équivalente aux équations de Cauchy-Riemann pour les fonctions complexes

$$f(x, y) = f_1(x, y) + i f_2(x, y)$$

impliquant que f est holomorphe.

10. Ne pas en tirer de conclusions hâtives : ceci est faux en toute généralité.

En toute généralité, une fonction satisfaisant l'équation de Dirac est appelée monogène quand $n > 2$ et holomorphe quand $n = 2$ (voir [Felsberg 2002][Gilbert 1991]).

2.2 Le signal analytique

“L’auteur définit également la fréquence instantanée d’un signal, en utilisant la notion de signal analytique (obtenu par prolongement analytique du signal réel quand le temps est considéré comme une variable complexe). Ces notions de fréquence instantanée et de spectre instantané sont appelées à fournir une base théorique ferme aux recherches sur la modulation de fréquence, sur l’analyse harmonique continue, sur la compression de fréquences, ...” [Ville 1948]. Nous commençons par quelques exemples, puis nous montrons comment la construction proposée par J. Ville, et initiée par D. Gabor [Gabor 1946], s’insère dans le cadre des algèbres de Clifford.

Premier exemple :

Considérons le signal :

$$s(t) = \cos \omega t$$

Nous pouvons le considérer, indifféremment, comme partie réelle de :

$$\Psi(t) = e^{i\omega t}$$

ou de :

$$\Psi^*(t) = e^{-i\omega t}$$

Tant que t ne prend que des valeurs réelles, il n’y a pas de raison pour choisir l’une de ces deux formes plutôt que l’autre. Ce n’est que si t prend des valeurs complexes :

$$t = \tau + i\theta$$

que nous voyons apparaître des différences. Si, en effet, θ tend vers $+\infty$, Ψ tend vers zéro et Ψ^* vers l’infini.

Si nous convenons de ne conserver, dans la mesure du possible, que des fonctions régulières dans le demi-plan supérieur, nous choisirons donc :

$$\Psi(t) = e^{i\omega t}$$

Deuxième exemple :

Soit le signal :

$$s(t) = \frac{1}{1+t^2}$$

Nous pouvons le considérer indifféremment comme partie réelle des deux fonctions :

$$\frac{1}{1+it} \text{ ou } \frac{1}{1-it}$$

Une seule d'entre elles, la seconde, est régulière dans le demi-plan supérieur, nous choisirons donc :

$$\Psi(t) = \frac{1}{1 - it}$$

Troisième exemple :

Soit le signal :

$$s(t) = \begin{cases} 1 & \text{if } -1 < t < +1 \\ 0 & \text{if } t < -1 \text{ ou } t > 1 \end{cases}$$

Nous pouvons considérer $s(t)$ comme partie réelle de :

$$\Psi(t) = \frac{1}{\pi i} \log \frac{t-1}{t+1}$$

avec une détermination convenablement choisie. Cette fonction n'étant pas uniforme, nous la rendrons uniforme dans le demi-plan supérieur en excluant de ce demi-plan les points singuliers -1 et $+1$.

D'une manière générale, les principes ci-dessus conduisent à associer au signal $s(t)$ le signal en quadrature :

$$\begin{aligned} \sigma(t) &= \frac{PV}{\pi} \int_{-\infty}^{+\infty} \frac{s(\tau) d\tau}{t - \tau} \\ &= \frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \left(\int_{-\infty}^{t-\varepsilon} \frac{s(\tau) d\tau}{t - \tau} + \int_{t+\varepsilon}^{+\infty} \frac{s(\tau) d\tau}{t - \tau} \right) \end{aligned}$$

et la fonction :

$$\Psi(t) = s(t) + i\sigma(t)$$

La formule précédente définit $\Psi(t)$ comme fonction de la variable complexe t , holomorphe dans le demi-plan supérieur. Un signal tel que $\Psi(t)$ sera appelé signal analytique.

Réciproquement, nous pouvons alors définir $s(t)$ par :

$$s(t) = \Re e \left[|\psi(t)| e^{i \arg \Psi} \right]$$

Ce qui nous définit $s(t)$ comme le résultat de la modulation du signal $e^{i \arg \Psi}$ par le signal $|\Psi|$.

2.2.1 Construction dans l'algèbre de Clifford $\mathbb{R}_{2,0}$

Nous considérons le signal 1D $s : \mathbb{R} \rightarrow \mathbb{R}$ et nous lui associons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}_{2,0}$ à valeurs dans l'algèbre de Clifford $\mathbb{R}_{2,0}$ telle que $f(x) = s(x)e_2$. Maintenant, il s'agit de définir une fonction holomorphe f_a dans le demi-plan $\mathbb{R} \times \mathbb{R}^+$.

En utilisant la caractérisation décrite dans le paragraphe 2.1.4, nous cherchons donc une fonction f_a satisfaisant :

$$D_2 f_a = 0$$

et telle que la partie "réelle", qui correspond ici à la partie dans la direction e_2 , soit le signal lui-même, d'où la condition suivante :

$$f_a(x, 0) = k(x, 0)e_1 + s(x)e_2$$

où $k(x, y)$ est une fonction que nous devons construire.

De manière équivalente, nous cherchons une fonction harmonique u satisfaisant l'équation de Laplace :

$$D_2^2 u = 0$$

avec la condition au bord :

$$\frac{\partial u}{\partial y}(x, 0) = s(x)$$

dont le gradient $D_2 u$ sera la fonction holomorphe recherchée. Tout ceci nous amène donc à résoudre le système suivant :

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 & \text{si } y > 0 \\ e_2 \frac{\partial u}{\partial y} = f & \text{si } y = 0 \end{cases} \quad (2.1)$$

avec $\Delta = D_2^2$, où D_2 est l'opérateur de Dirac.

La première équation est l'équation de Laplace 2D restreinte au demi-plan $y > 0$ et la deuxième équation est la condition au bord. L'équation de Laplace est une équation aux dérivées partielles de type elliptique. La solution fondamentale de l'équation de Laplace 2D (au facteur $\frac{1}{\pi}$ près) est :

$$L_2(x, y) = \frac{\log(\sqrt{x^2 + y^2})}{\pi}$$

et son gradient associé est :

$$g(x, y) = D_2 L_2(x, y) = \frac{x}{\pi(x^2 + y^2)} e_1 + \frac{y}{\pi(x^2 + y^2)} e_2$$

Nous faisons le produit géométrique de g avec e_2 afin de séparer les deux termes dans des entités différentes, *i.e.* une partie scalaire et une partie bivectorielle, ce qui permettra à la fin d'obtenir une solution vectorielle :

$$g e_2 = \frac{x}{\pi(x^2 + y^2)} e_1 e_2 + \frac{y}{\pi(x^2 + y^2)}$$

Le deuxième terme (la partie scalaire de ce produit géométrique) est plus connu sous le nom de noyau de Poisson 1D [Stein 1993, Gilbert 1991] et le premier terme (la partie bivectorielle de ce produit géométrique) est appelé le noyau de Poisson 1D conjugué [Stein 1993, Gilbert 1991].

Pour étendre la fonction f au demi-plan $y > 0$, nous allons utiliser deux propriétés :

1. La fonction ge_2 est holomorphe dans le demi-plan supérieur puisque $D_2ge_2 = D_2D_2L_2e_2 = 0$.
2. Pour toutes fonctions intégrables h et k , on a $D_2(h*k) = (D_2h)*k = h*(D_2k)$.

Il faut donc trouver une fonction f_0 associée à f qui par convolution deviendra holomorphe dans le demi-plan supérieur, ce constat nous amène à poser :

$$f_0(x, y) = f(x)\delta_0(y) = s(x)\delta_0(y)e_2$$

où

$$\delta_0(y) = \begin{cases} 0 & \text{si } y > 0 \\ 1 & \text{si } y = 0 \end{cases}$$

L'étape suivante est de convoler cette fonction f_0 avec la fonction ge_2 :

$$ge_2 * f_0(x, y) = \iint \left(\frac{x'}{\pi(x'^2 + y'^2)}e_1e_2 + \frac{y'}{\pi(x'^2 + y'^2)} \right) f(x - x')\delta_0(y - y')dx'dy'$$

En intégrant par rapport à y' , cette convolution 2D se simplifie en une convolution 1D et nous obtenons l'expression suivante :

$$\begin{aligned} ge_2 * f_0(x, y) &= \int \left(\frac{x'}{\pi(x'^2 + y^2)}e_1e_2 + \frac{y}{\pi(x'^2 + y^2)} \right) f(x - x')dx' \\ &= \int \frac{x'}{\pi(x'^2 + y^2)}e_1e_2f(x - x')dx' + \int \frac{y}{\pi(x'^2 + y^2)}f(x - x')dx' \\ &= \left[\frac{xe_1e_2}{\pi(x^2 + y^2)} * f \right] (x) + \left[\frac{y}{\pi(x^2 + y^2)} * f \right] (x) \end{aligned}$$

Par suite, nous appliquons la transformée de Fourier à cette expression et nous obtenons (voir Annexe B pour le détail des calculs) :

$$\mathcal{F}\{ge_2 * f_0(x, y)\}(u) = H_{P_{1D}}(u)H_H(u)F(u) + H_{P_{1D}}(u)F(u)$$

avec

$$\begin{cases} H_H(u) = \frac{u}{|u|} \\ H_{P_{1D}}(u) = e^{-2\pi|u|y} \\ F(u) = \text{transformée de Fourier de } f \end{cases} \quad (2.2)$$

En appliquant la transformée de Fourier inverse, nous obtenons donc l'expression suivante que nous noterons f_a :

$$f_a(x, y) = h_{P_{1D}} * f(x, y) + h_{P_{1D}} * h_H * f(x, y) \quad (2.3)$$

où

$$\begin{cases} h_{P_{1D}} = \frac{y}{\pi(x^2 + y^2)} \text{ (noyau de Poisson 1D)} \\ h_H = \frac{e_1 e_2}{\pi x} \text{ (noyau de Hilbert).} \end{cases}$$

La variable y est un paramètre d'échelle et en le faisant tendre vers 0, nous obtenons le signal analytique classique :

$$f_a(x, 0) = s(x)e_2 + h_H * s(x)e_1 = s(x) + i(h_H * s(x))$$

2.2.2 Représentation polaire d'un signal monodimensionnel

Dans ce paragraphe, nous abordons la notion d'amplitude et de phase du signal analytique afin d'obtenir sa décomposition polaire dans l'algèbre de Clifford $\mathbb{R}_{2,0}$. Soit s un signal réel monodimensionnel et f la fonction vectorielle associée. Sa représentation holomorphe f_a est donc :

$$\begin{aligned} f_a(x, y) &= h_{P_{1D}} * f(x, y) + h_{P_{1D}} * h_H * f(x, y) \\ &= A_1 e_1 + A_2 e_2 \end{aligned}$$

avec

$$\begin{cases} A_1 = h_{P_{1D}} * \frac{1}{\pi x} * s(x) \\ A_2 = h_{P_{1D}} * s(x) \end{cases}$$

Le paramètre y étant un paramètre d'échelle, la fonction f_a correspond au signal analytique de la fonction s lissée par le noyau de Poisson.

La fonction f_a étant vectorielle, nous allons la mettre sous forme de 2-verseur¹¹ en regardant le produit géométrique de f_a et e_2 . En procédant ainsi, nous étudions le 2-verseur S qui envoie l'élément de base e_1 sur f_a .

$$S = f_a e_2 = A_2 + A_1 e_1 e_2$$

Cette transformation n'est pas anodine, en agissant ainsi, nous pouvons maintenant écrire :

$$S = A_1 + i A_2$$

Et nous retrouvons exactement l'expression du signal analytique dont l'amplitude \mathcal{A} et la phase φ s'écrivent :

$$\begin{aligned} \mathcal{A} &= |f_a e_2| = |f_a| |e_2| = |f_a| = \sqrt{A_1^2 + A_2^2} \\ \varphi &= \arctan \frac{A_2}{A_1} \end{aligned}$$

11. produit géométrique de deux vecteurs.

L'amplitude \mathcal{A} est donc la norme du vecteur f_a ou de manière équivalente le module de la fonction complexe. La phase φ correspond à l'angle de rotation du 2-verseur, ce qui est précisément l'argument de la fonction complexe.

Nous retrouvons donc un résultat classique sur le signal analytique, à savoir, sa représentation polaire :

$$f_a e_2 = \mathcal{A} e^{e_1 e_2 \varphi}$$

2.2.3 Analyse de signaux monodimensionnels

D'après l'analyse de Fourier, il est bien connu que tout signal $g \in L_2(\mathbb{R}, \mathbb{R})$ peut être approximé avec une erreur infiniment petite par sa série de Fourier

$$g(x) = \sum_{\nu} a_{\nu} \cos(2\pi\nu x + \phi_{\nu})$$

où ν correspond à une fréquence. Chaque composante de fréquence avec sa phase et son amplitude individuelles peut être analysée séparément. Ces fréquences d'intérêt sélectionnées comportent l'information de structure du signal et doivent être extraites du signal original. Pour cela, nous allons étudier ce signal à l'origine pour une fréquence arbitraire ν_0 . Le modèle local du signal 1D à l'origine 0 donne :

$$\begin{aligned} f_a(x, y)|_{x=0} &= h_{P_{1D}} * g(x)|_{x=0} e_2 + h_{P_{1D}} * h_H * g(x)|_{x=0} e_1 \\ &= a_{\nu_0}^y \cos(\phi_{\nu_0}^y) e_2 + a_{\nu_0}^y \sin(\phi_{\nu_0}^y) e_1 \end{aligned}$$

où $a_{\nu_0}^y$ et $\phi_{\nu_0}^y$ sont l'amplitude locale et la phase locale en 0 à l'échelle y . En d'autres termes, il s'agit de l'amplitude locale et de la phase locale du signal lissé par le noyau de Poisson.

En utilisant la représentation polaire introduite précédemment, l'amplitude locale et la phase locale de f_a en 0 sont :

$$\mathcal{A} = \sqrt{(a_{\nu_0}^y \cos \phi_{\nu_0}^y)^2 + (a_{\nu_0}^y \sin \phi_{\nu_0}^y)^2} = a_{\nu_0}^y$$

et

$$\varphi = \arctan \frac{\sin \phi_{\nu_0}^y}{\cos \phi_{\nu_0}^y} = \phi_{\nu_0}^y$$

Grâce au signal analytique, nous pouvons donc trouver de manière exacte l'amplitude locale et la phase locale de chaque fréquence.

2.3 Le signal monogène

Comme nous l'avons vu dans le paragraphe précédent, le signal analytique est particulièrement bien adapté à l'étude des signaux monodimensionnels. Evidemment la tentation est grande de vouloir généraliser la construction de J. Ville à des signaux bidimensionnels représentant, par exemple, des images en niveaux de gris. Une première approche consiste à étendre la transformée de Hilbert à la dimension

deux dont nous donnons ci-dessous trois exemples. Nous allons voir cependant que l'approche retenue précédemment permet d'obtenir une généralisation "monogène" très naturelle.

Premier exemple : transformée de Hilbert partielle [Hahn 1996]

La transformée de Hilbert partielle f_{Hi}^j d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ par rapport à l'orientation $j \in \{1, 2\}$ est donnée par :

$$f_{Hi}^j(x_1, x_2) = PV \int_{\mathbb{R}^2} \frac{f(x'_1, x'_2)}{\pi(x_j - x'_j)} dx'_j$$

Le signal complexe partiel f_{part}^j d'un signal f par rapport à $j \in \{1, 2\}$ est défini par :

$$f_{part}^j(x_1, x_2) = f(x_1, x_2) + i f_{Hi}^j(x_1, x_2)$$

Dans cette approche, l'auteur utilise la transformée de Hilbert dans chacune des directions x et y .

Deuxième exemple : transformée de Hilbert totale [Stark 1971]

La transformée de Hilbert totale f_{Hi} d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est donnée par :

$$f_{Hi}(x, y) = PV \int_{\mathbb{R}^2} \frac{f(x', y')}{\pi^2(x - x')(y - y')} dx' dy'$$

Le signal complexe total de f s'écrit alors :

$$f_{tot}(x, y) = f(x, y) + i f_{Hi}(x, y)$$

Dans ce travail, l'auteur définit le noyau $\frac{1}{\pi^2 xy}$ comme une extension 2D du noyau de Hilbert $\frac{1}{\pi x}$. Nous pouvons noter que la transformée de Hilbert totale peut être considérée comme l'application successive des transformées de Hilbert partielles par rapport aux deux directions x_1 et x_2 .

Troisième exemple : signal analytique quaternionique [Bülow 2001]

Le signal analytique quaternionique utilise les transformées de Hilbert partielles et totale d'un signal $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et est exprimé au travers des quaternions comme suit :

$$f_A^q(x, y) = f(x, y) + i f_{Hi}^1(x, y) + j f_{Hi}^2(x, y) + k f_{Hi}(x, y)$$

Dans toutes ces approches, il existe un problème majeur souligné dans [Felsberg 2002], en effet, elles ne permettent pas d'estimer correctement l'amplitude et la phase de signaux 2D bien que le signal analytique quaternionique soit très proche de la solution adéquate. Dans toutes ces méthodes, il n'y a pas d'invariance par rotation ce qui est la clé d'une solution optimale.

Dans [Felsberg 2002, Felsberg 2001], l'auteur propose une solution au problème en utilisant la solution fondamentale de l'équation de Laplace 3D. En exprimant le problème à l'aide de fonctions monogènes dans l'algèbre de Clifford $\mathbb{R}_{3,0}$, il en résulte une solution basée sur la transformée de Riesz qui possède toutes les propriétés requises en terme d'amplitude, de phase et d'invariance par rotation. Cette approche est présentée dans le paragraphe suivant.

2.3.1 Construction dans l'algèbre de Clifford $\mathbb{R}_{3,0}$

Considérons $s : \mathbb{R}^2 \rightarrow \mathbb{R}$ un signal réel bidimensionnel et posons $f : \mathbb{R}^2 \rightarrow \mathbb{R}_{3,0}$ la représentation vectorielle de ce signal dans l'algèbre de Clifford $\mathbb{R}_{3,0}$ définie par $f(x, y) = s(x, y)e_3$. Le problème se ramène à définir une fonction monogène f_m dans le demi-espace $\mathbb{R}^2 \times \mathbb{R}^+$.

En utilisant la caractérisation du paragraphe 2.1.4, nous cherchons donc une fonction f_m satisfaisant :

$$D_3 f_m = 0$$

et telle que la partie "réelle", qui correspond ici à la partie dans la direction e_3 , soit le signal lui-même, d'où la condition suivante :

$$f_m(x, 0) = k_1(x, 0)e_1 + k_2(x, 0)e_2 + s(x)e_3$$

où $k_1(x, y)$ et $k_2(x, y)$ sont des fonctions que nous devons construire.

De manière équivalente, nous cherchons une fonction harmonique u satisfaisant l'équation de Laplace :

$$D_3^2 u = 0$$

avec la condition au bord :

$$\frac{\partial u}{\partial z}(x, y, 0) = s(x)$$

dont le gradient $D_3 u$ sera la fonction monogène recherchée. Tout ceci nous amène donc à résoudre le système suivant :

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 & \text{si } z > 0 \\ e_3 \frac{\partial u}{\partial z} = f & \text{si } z = 0 \end{cases} \quad (2.4)$$

où $\Delta = D_3^2$ et $D_3 = e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y} + e_3 \frac{\partial}{\partial z}$.

La première équation est l'équation de Laplace 3D restreinte au demi-espace $z > 0$ et la seconde équation est la condition au bord.

La solution fondamentale de l'équation de Laplace 3D (au facteur $\frac{1}{2\pi}$) est :

$$L_3(x, y) = -\frac{1}{2\pi\sqrt{x^2 + y^2 + z^2}}$$

et son gradient associé est :

$$\begin{aligned} g(x, y, z) &= D_3 L_3(x, y, z) \\ &= \frac{x}{2\pi(x^2 + y^2 + z^2)^{3/2}} e_1 + \frac{y}{2\pi(x^2 + y^2 + z^2)^{3/2}} e_2 + \frac{z}{2\pi(x^2 + y^2 + z^2)^{3/2}} e_3 \end{aligned}$$

Nous faisons le produit géométrique de g avec e_3 afin de séparer les trois termes dans des entités différentes (comme pour la construction du signal analytique) ce qui permettra à la fin d'obtenir une solution vectorielle :

$$ge_3 = \frac{x}{2\pi(x^2 + y^2 + z^2)^{3/2}} e_1 e_3 + \frac{y}{2\pi(x^2 + y^2 + z^2)^{3/2}} e_2 e_3 + \frac{z}{2\pi(x^2 + y^2 + z^2)^{3/2}}$$

Le dernier terme (la partie scalaire) est plus connu sous le nom de noyau de Poisson 2D [Stein 1993, Gilbert 1991], le premier et le deuxième terme (partie bivectorielle) sont les noyaux de Poisson 2D conjugués [Stein 1993, Gilbert 1991] (voir Fig. 2.1).

Pour étendre la fonction f au demi-espace $z > 0$, nous allons utiliser les deux mêmes

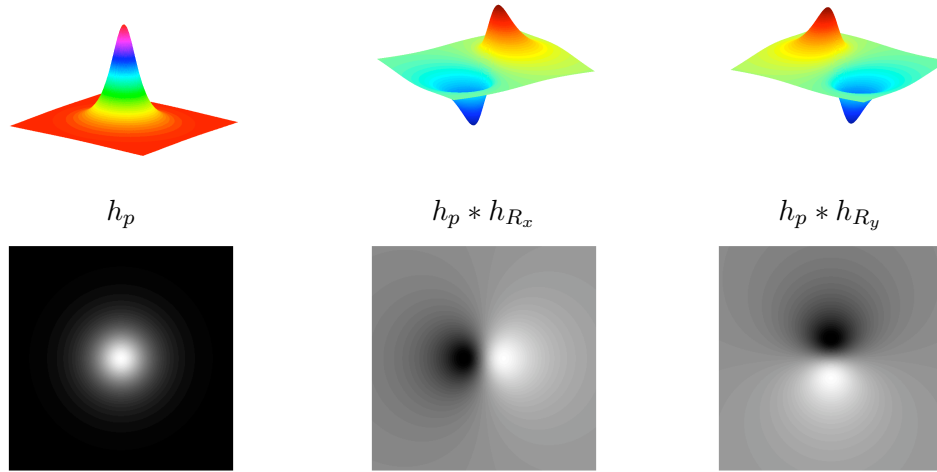


FIGURE 2.1 – Représentation 3D (première ligne) et 2D (deuxième ligne) du noyau de Poisson, du noyau de Poisson conjugué en x et du noyau de Poisson conjugué en y respectivement.

propriétés que dans 2.2.1, à savoir :

1. La fonction ge_3 est monogène dans le demi-espace supérieur puisque $D_3 ge_3 = D_3 D_3 L_3 e_3 = 0$.

2. Pour toutes fonctions intégrables h et k , on a $D_3(h*k) = (D_3h)*k = h*(D_3k)$. Il faut donc trouver une fonction f_0 associée à f qui par convolution deviendra monogène dans le demi-espace supérieur, posons alors :

$$f_0(x, y) = f(x, y)\delta_0(z) = s(x, y)\delta_0(z)e_3$$

et faisons la convolution 3D suivante :

$$ge_3 * f_0(x, y, z) = \iiint \frac{(x'e_1e_3 + y'e_2e_3 + z')}{2\pi(x'^2 + y'^2 + z'^2)^{3/2}} f(x - x', y - y')\delta_0(z - z')dx'dy'dz'$$

En intégrant par rapport à z' , cette convolution 3D se réduit à une convolution 2D, et nous obtenons :

$$\begin{aligned} ge_3 * f_0(x, y, z) &= \iint \frac{(x'e_1e_3 + y'e_2e_3 + z)}{2\pi(x'^2 + y'^2 + z^2)^{3/2}} f(x - x', y - y')dx'dy' \\ &= \iint \frac{(x'e_1e_3 + y'e_2e_3)}{2\pi(x'^2 + y'^2 + z^2)^{3/2}} f(x - x', y - y')dx'dy' \\ &+ \iint \frac{z}{2\pi(x'^2 + y'^2 + z^2)^{3/2}} f(x - x', y - y')dx'dy' \\ &= \left[\frac{xe_1e_3 + ye_2e_3}{2\pi(x^2 + y^2 + z^2)^{3/2}} * f \right] (x, y) + \left[\frac{z}{2\pi(x^2 + y^2 + z^2)^{3/2}} * f \right] (x, y) \end{aligned}$$

Par suite, nous appliquons la transformée de Fourier à cette expression et nous obtenons (voir Annexe C pour le détail des calculs) :

$$\mathcal{F}\{ge_3 * f_0(x, y, z)\}(u, v) = H_{P_{2D}}(u, v)H_R(u, v)F(u, v) + H_{P_{2D}}(u, v)F(u, v)$$

avec

$$\begin{cases} H_R(u, v) = \frac{ue_1 + ve_2}{\sqrt{u^2 + v^2}} I_2^{-1} \\ H_{P_{2D}}(u, v) = e^{-2\pi\sqrt{u^2 + v^2}z} \\ F(u, v) = \text{transformée de Fourier de } f \end{cases} \quad (2.5)$$

Par transformée de Fourier inverse, on obtient donc l'expression suivante que l'on notera f_m (voir [Felsberg 2002]) :

$$f_m(x, y, z) = h_{P_{2D}} * f(x, y) + h_{P_{2D}} * h_R * f(x, y) \quad (2.6)$$

où

$$\begin{cases} h_{P_{2D}} = \frac{z}{2\pi(x^2 + y^2 + z^2)^{3/2}} \\ h_R = \frac{xe_1 + ye_2}{2\pi(x^2 + y^2)^{3/2}} e_3 \end{cases}$$

La fonction $h_{P_{2D}}$ est le noyau de Poisson 2D et la fonction h_R est le noyau de Riesz, une extension en dimension deux du noyau de Hilbert. La variable z est un paramètre d'échelle et en le faisant tendre vers 0, le signal monogène est obtenu.

2.3.2 Représentation polaire d'une image en niveau de gris

La construction que nous venons de décrire permet, comme c'est le cas pour le signal analytique, d'associer à un signal bidimensionnel une représentation polaire. Cette dernière est caractérisée par une amplitude, une phase et une orientation.

Soit $I : \mathbb{R}^2 \rightarrow \mathbb{R}$ une image en niveaux de gris. Nous lui associons sa représentation monogène que nous noterons f_m telle que

$$\begin{aligned} f_m(x, y, s) &= h_{P_{2D}} * I(x, y) + h_{P_{2D}} * h_R * I(x, y) \\ &= A_1 e_1 + A_2 e_2 + A_3 e_3 \end{aligned}$$

avec

$$\begin{cases} A_1 = h_{P_{2D}} * h_{R_x} * I(x, y) \\ A_2 = h_{P_{2D}} * h_{R_y} * I(x, y) \\ A_3 = h_{P_{2D}} * I(x, y) \end{cases}$$

Les composantes A_1 et A_2 contiennent respectivement des informations liées aux structures verticales et horizontales dues au noyau de Riesz. La composante A_3 est l'image elle-même et le tout est lissé par un noyau de Poisson 2D, ce qui donne au signal monogène un caractère multi-échelle dans l'espace-échelle de Poisson (voir Annexe A).

Le signal monogène obtenu est un vecteur dans $\mathbb{R}_{3,0}$ (voir Fig. 2.2) et pour l'étudier, nous allons regarder le produit géométrique entre e_3 et f_m . En procédant ainsi, nous étudions le 2-verseur S qui envoie l'élément de base e_3 sur f_m .

$$S = f_m e_3 = A_3 + A_1 e_1 e_3 + A_2 e_2 e_3$$

L'amplitude d'un tel signal est :

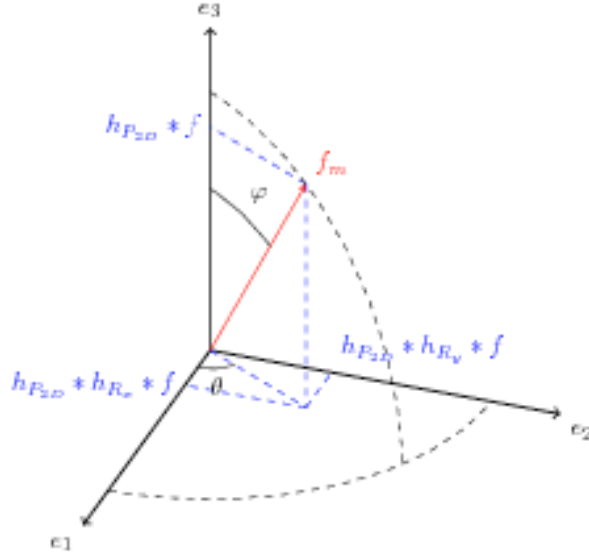
$$\mathcal{A} = |f_m e_3| = |f_m| |e_3| = |f_m| = \sqrt{A_1^2 + A_2^2 + A_3^2}$$

Si nous écrivons S différemment :

$$S = A_3 + \frac{A_1 e_1 e_3 + A_2 e_2 e_3}{|A_1 e_1 e_3 + A_2 e_2 e_3|} |A_1 e_1 e_3 + A_2 e_2 e_3|$$

alors en utilisant l'exemple 3 du paragraphe 2.1.2, nous pouvons écrire :

$$S = A_3 + i \sqrt{A_1^2 + A_2^2}$$

FIGURE 2.2 – Représentation du signal monogène f_m dans $\mathbb{R}_{3,0}$.

et nous définissons la phase du signal monogène par :

$$\varphi = \arctan \frac{\sqrt{A_1^2 + A_2^2}}{A_3}$$

Enfin M. Felsberg définit une nouvelle caractéristique qu'il nomme orientation du signal et qui est donnée par l'expression :

$$\theta = \arctan \frac{A_2}{A_1}$$

Pour conclure, le signal monogène peut comme le signal analytique s'écrire sous forme polaire :

$$f_m e_3 = \mathcal{A} e^{B\varphi}$$

où $B = \frac{A_1 e_1 e_3 + A_2 e_2 e_3}{|A_1 e_1 e_3 + A_2 e_2 e_3|}$ est le bivecteur engendrant le plan de rotation entre e_3 et f_m .

2.3.3 Analyse de signaux bidimensionnels

2.3.3.1 Classification des signaux 2D

Les signaux 2D $f \in L_2 = L_2(\mathbb{R}, \mathbb{R})$ sont classés en régions locales $N \subseteq \Omega$ de différentes dimensions intrinsèques [Felsberg 2002][Bülow 2001] (qui correspondent à leur codimension). La dimension intrinsèque exprime le nombre de degré de liberté nécessaire à décrire la structure locale. Les signaux constants sont de dimension zéro

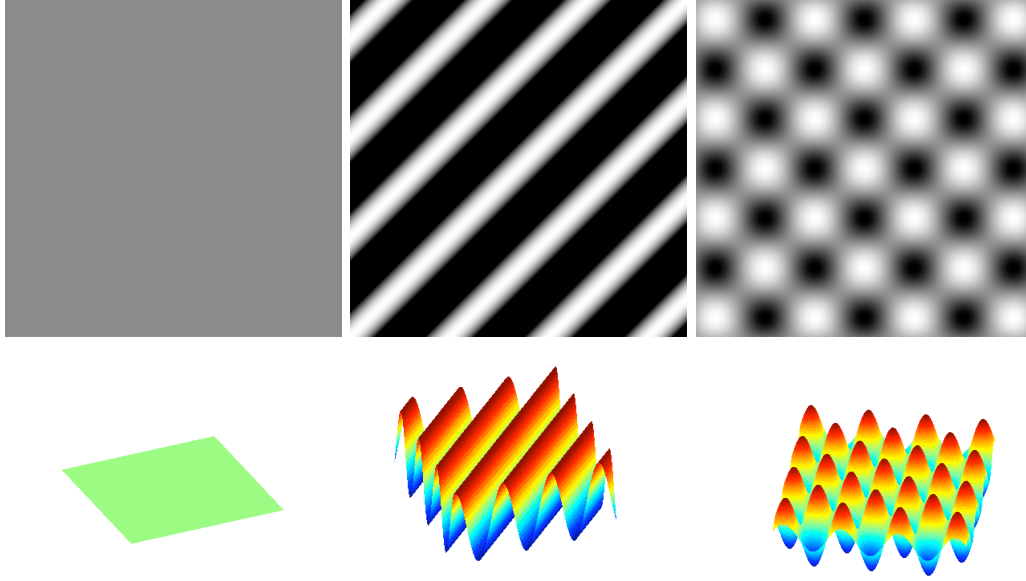


FIGURE 2.3 – Première ligne de gauche à droite : un signal constant (i0D), un signal orienté arbitrairement (i1D) et un signal i2D qui consiste en deux signaux i1D superposés. Deuxième ligne de gauche à droite : ces mêmes signaux respectivement représentés en 3D.

(i0D), les lignes et les contours sont de dimension intrinsèque un (i1D) et toutes les autres formes possibles sont de dimension intrinsèque deux (i2D) (voir Figure 2.3).

$$\begin{aligned} \text{i0D} &= \{f \in L_2, f(x_i, y_i) = f(x_j, y_j), \forall (x_i, y_i), (x_j, y_j) \in N\} \\ \text{i1D} &= \{f \in L_2, f(x, y) = g(x \cos \theta + y \sin \theta), \forall (x, y) \in N\} \\ \text{i2D} &= L_2 \setminus (\text{i0D} \cup \text{i1D}) \end{aligned}$$

avec $g \in L_2(\mathbb{R}, \mathbb{R})$. En général, les signaux i2D peuvent seulement être modélisés par une infinité de signaux i1D superposés.

2.3.3.2 Analyse de signaux i1D

Le modèle de signal i1D (qui est un signal 1D orienté arbitrairement dans le domaine 2D) s'écrit :

$$g(x, y) = a \cos(\omega(x \cos \theta + y \sin \theta) + \phi)$$

où a est l'amplitude, ω la fréquence, θ l'orientation et ϕ la phase.

Dans l'espace échelle de Poisson, ce modèle est donné pour chaque paramètre d'échelle $s > 0$ par :

$$f(x, y, s) = h_{P_{2D}} * g(x, y) = a_s \cos(\omega_s(x \cos \theta_s + y \sin \theta_s) + \phi_s)$$

Le signal monogène associé est donc :

$$f_m(x, y, s) = f(x, y, s) + h_R * f(x, y, s)$$

Le modèle local de ce signal à l'origine $(0, 0) \in \mathbb{R}^2$ donne :

$$\begin{aligned} f_m(x, y, s)|_{(x,y,s)=(0,0,s)} &= f(x, y, s)|_{(x,y,s)=(0,0,s)} + h_R * f(x, y, s)|_{(x,y,s)=(0,0,s)} \\ &= a_s \cos \phi_s e_3 + a_s \sin \phi_s \cos \theta_s e_1 + a_s \sin \phi_s \sin \theta_s e_2 \end{aligned}$$

En utilisant la représentation polaire introduite précédemment, l'amplitude locale, la phase locale et l'orientation locale de f_m sont :

$$\left\{ \begin{array}{l} \mathcal{A} = \sqrt{(a_s \cos \phi_s)^2 + (a_s \sin \phi_s \cos \theta_s)^2 + (a_s \sin \phi_s \sin \theta_s)^2} = a_s \\ \varphi = \arctan \frac{\sqrt{(a_s \sin \phi_s \cos \theta_s)^2 + (a_s \sin \phi_s \sin \theta_s)^2}}{a_s \cos \phi_s} = \phi_s \\ \Theta = \arctan \frac{a_s \sin \phi_s \sin \theta_s}{a_s \sin \phi_s \cos \theta_s} = \theta_s \end{array} \right.$$

Et de manière générale, on obtient la phase et la fréquence de la manière suivante :

$$\left\{ \begin{array}{l} \varphi = \omega_s(x \cos \theta_s + y \sin \theta_s) + \phi_s \\ \Omega = \sqrt{\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial y}\right)^2} = \omega_s \end{array} \right.$$

Le modèle est donc bien adapté pour l'analyse des signaux i1D puisqu'il permet de retrouver l'amplitude et la phase d'un signal i1D tout en étant invariant par rotation.

Malheureusement, le signal monogène ne distingue pas les signaux i1D des signaux i2D et c'est pour palier ce problème que le signal monogène conforme¹² a été introduit [Wietzke 2010].

2.4 Le signal monogène couleur

Il s'agit ici du principal apport théorique de cette thèse. La construction que nous allons décrire dans la suite de ce paragraphe justifie *a posteriori* l'approche retenue, basée sur l'utilisation du formalisme des algèbres de Clifford. Comme nous l'avons déjà mentionné, ce dernier offre un cadre d'étude très générique, en ce sens qu'il permet de s'affranchir de particularités liées à la dimension. Il n'est donc pas étonnant que, dans la continuité des deux constructions précédentes, nous puissions obtenir une représentation "monogène" d'un signal s de \mathbb{R}^2 à valeurs dans \mathbb{R}^3 . Une construction beaucoup plus générale pourrait d'ailleurs être envisagée. Il va sans dire que ce qui nous préoccupe est le traitement de signaux correspondant à des images couleur.

12. terme dû à l'utilisation des algèbres de Clifford conformes

Soit $s : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ un signal bidimensionnel à valeurs dans \mathbb{R}^3 tel que :

$$s(x, y) = (f_3(x, y), f_4(x, y), f_5(x, y))$$

Et soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}_{5,0}$ le signal vectoriel associé à s défini par

$$f(x_1, x_2) = f_3(x_1, x_2)e_3 + f_4(x_1, x_2)e_4 + f_5(x_1, x_2)e_5$$

Dans notre formulation du problème, l'espace couleur considéré est engendré par e_3, e_4, e_5 et tout système colorimétrique tels que RGB ou CIE L*a*b* peut être considéré.

2.4.1 Système colorimétrique

Il est bien connu qu'il existe de multiples variantes pour définir les couleurs [Fairchild 2005]. Dans la suite, nous serons amenés à considérer deux espaces couleur particuliers couramment utilisés en traitement d'images.

Le premier est l'espace RVB (Rouge, Vert, Bleu) défini à partir des trois couleurs primaires correspondant aux trois types de cônes de l'oeil humain. Cet espace, traditionnellement appelé le "cube RVB" car chaque couleur appartient à l'intervalle $[0, 255]$, est plongé dans la partie vectorielle engendrée par les vecteurs e_3, e_4 et e_5 de l'algèbre $\mathbb{R}_{5,0}$ (voir la figure 2.4).

Le second est l'espace CIE L*a*b*. Celui-ci est traditionnellement qualifié de perceptuel car sa définition repose sur la façon dont on perçoit les couleurs. Une couleur est représentée par des coordonnées (x, y, z) appartenant à $[-\alpha, \alpha]^2 \times [0, 100]$, le couple (x, y) mesurant la chrominance et z étant approximativement proportionnel à la racine cubique de la luminance. Bien qu'il n'existe pas de transformations linéaires entre RVB et L*a*b*, nous plongeons L*a*b* dans $\mathbb{R}_{5,0}$ de la même façon que RVB (voir la figure 2.5). Il serait intéressant, nous ne l'abordons pas ici, de faire varier la forme quadratique de l'espace vectoriel \mathbb{R}^5 pour tenir compte de la non linéarité.

2.4.2 Construction dans l'algèbre de Clifford $\mathbb{R}_{5,0}$

En utilisant la construction précédente du signal monogène, nous pouvons construire une représentation pour des images couleur dans l'algèbre de Clifford $\mathbb{R}_{5,0}$, de la manière suivante :

1. Soit $s(x, y) = (f_3(x, y), f_4(x, y), f_5(x, y))$ une image couleur dans le système $f_3 f_4 f_5$.
2. Soit $f(x, y) = f_3(x, y)e_3 + f_4(x, y)e_4 + f_5(x, y)e_5$ la fonction vectorielle dans $\mathbb{R}_{5,0}$ associée à s .
3. Soit γ la fonction monogène dans $\mathbb{R}^2 \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$ telle que sa partie "réelle" sur le bord du domaine \mathbb{R}^2 soit l'image couleur f :

$$\begin{aligned} \gamma(x, y, 0, 0, 0) &= k_1(x, y, 0, 0, 0)e_1 + k_2(x, y, 0, 0, 0)e_2 \\ &\quad + f_3(x, y)e_3 + f_4(x, y)e_4 + f_5(x, y)e_5 \end{aligned}$$

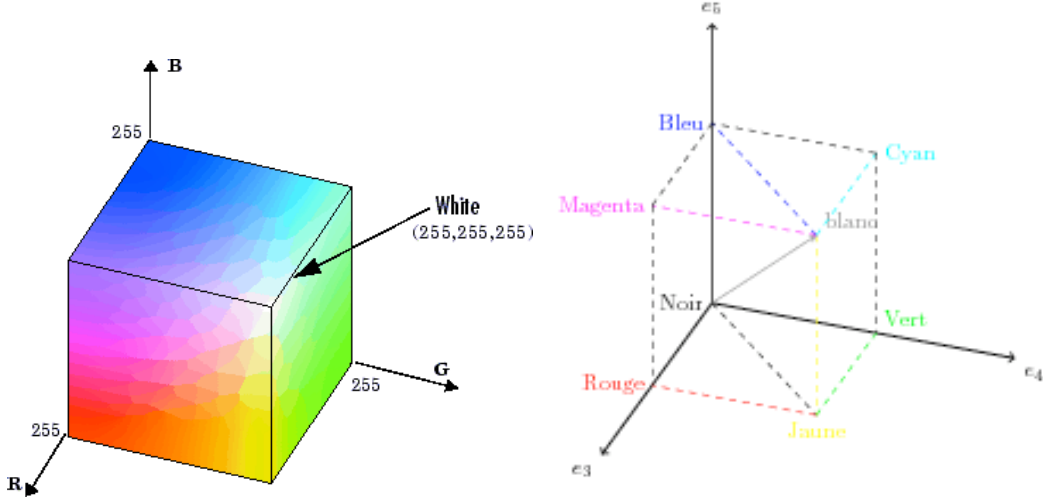


FIGURE 2.4 – Représentation du système colorimétrique RGB dans \mathbb{R}^3 (à gauche) et dans $\mathbb{R}_{5,0}$ à droite.

où les fonctions k_1 et k_2 sont inconnues.

Sachant que γ doit satisfaire l'équation de Dirac $D_5\gamma = 0$, nous sommes amenés à étudier le problème suivant dont le gradient de la solution D_5u sera notre fonction γ recherchée :

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + \frac{\partial^2 u}{\partial x_4^2} + \frac{\partial^2 u}{\partial x_5^2} = 0 & \text{si } x_3 > 0, x_4 > 0 \text{ et } x_5 > 0 \\ e_3 \frac{\partial u}{\partial x_3} + e_4 \frac{\partial u}{\partial x_4} + e_5 \frac{\partial u}{\partial x_5} = f(x_1, x_2) & \text{si } x_3 = x_4 = x_5 = 0 \end{cases} \quad (2.7)$$

où $\Delta = D_5^2$.

Comme nous voulons un signal multi-échelle avec une échelle indépendante selon les canaux traités (x_3 pour f_3 , x_4 pour f_4 et x_5 pour f_5) il est préférable de ne pas utiliser la solution fondamentale de l'équation de Laplace 5D mais de diviser le système d'équations précédents en 3 sous-systèmes de la manière suivante :

$$(i = 3, 4, 5) \begin{cases} \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_i^2} = 0 & \text{si } x_i > 0 \\ e_i \frac{\partial u}{\partial x_i} = f_i(x_1, x_2)e_i & \text{si } x_i = 0 \end{cases} \quad (2.8)$$

Soit :

$$p(x_1, x_2, x_i) = -\frac{1}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}}$$

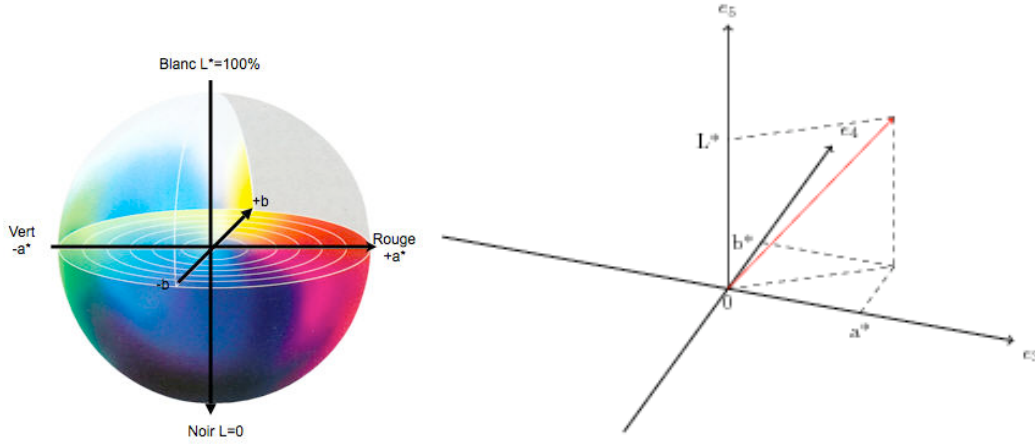


FIGURE 2.5 – Représentation du système colorimétrique CIE $L^*a^*b^*$ dans \mathbb{R}^3 à gauche et dans $\mathbb{R}_{5,0}$ à droite.

la solution de l'équation de Laplace 3D dont le gradient vectoriel associé est

$$g(x_1, x_2, x_i) = D_5 p(x_1, x_2, x_i) = \frac{x_1 e_1 + x_2 e_2 + x_i e_i}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}}$$

Afin de satisfaire les conditions au bord que nous avons imposé, nous faisons le produit géométrique entre la fonction g et le vecteur de base e_i dans le but de séparer le terme régularisant (noyau de Poisson) et le terme de structure (noyau de Poisson conjugué), nous obtenons alors :

$$\begin{aligned} g e_i &= g \cdot e_i + g \wedge e_i \\ &= \frac{x_i}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}} + \frac{x_1 e_1 e_i + x_2 e_2 e_i}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}} \\ &= h_p^i(x_1, x_2, x_i) + h_{cp}^i(x_1, x_2, x_i) \end{aligned}$$

En utilisant la fonction suivante :

$$k_i(x_1, x_2, x_i) = f_i(x_1, x_2) \delta_0(x_i) e_i$$

La convolution $g e_i * k_i$ est donnée par la somme de :

$$\iiint h_{cp}^i(x'_1, x'_2, x'_i) f_i(x_1 - x'_1, x_2 - x'_2) \delta_0(x_i - x'_i) e_i dx'_1 dx'_2 dx'_i$$

et

$$\iiint h_p^i(x'_1, x'_2, x'_i) f_i(x_1 - x'_1, x_2 - x'_2) \delta_0(x_i - x'_i) e_i dx'_1 dx'_2 dx'_i$$

En intégrant par rapport à x'_i , nous obtenons :

$$ge_i * k_i = \iint h_{cp}^i(x'_1, x'_2, x_i) f_i(x_1 - x'_1, x_2 - x'_2) e_i dx'_1 dx'_2 \\ + \iint h_p^i(x'_1, x'_2, x_i) f_i(x_1 - x'_1, x_2 - x'_2) e_i dx'_1 dx'_2$$

Nous renvoyons le lecteur à l'Annexe D pour la définition et les propriétés de la transformée de Fourier \mathcal{F} que nous utilisons ci-après. En appliquant \mathcal{F} à la convolution précédente, nous obtenons :

$$\mathcal{F}\{ge_i * k_i\}(u_1, u_2, x_i) = H_p^i(u_1, u_2) F_i(u_1, u_2) + H_p^i(u_1, u_2) H_R^i(u_1, u_2) F_i(u_1, u_2)$$

où H_p^i et H_R^i sont les transformées de Fourier des noyaux de Poisson 2D et de Riesz respectivement et $F_i = \mathcal{F}f_i$. Enfin, en utilisant la transformée de Fourier inverse, nous obtenons :

$$\mathcal{F}^{-1}\{\mathcal{F}\{ge_i * k_i\}\} = h_p^i * f_i + h_p^i * h_R^i * f_i$$

qui est la solution du système 2.8 pour $i = 1, 2, 3$. En effet, par définition de p , nous avons :

$$D_5(ge_i * k_i) = D_5(g * f_i \delta(x_i)) = D_5^2 p * f_i \delta(x_i) = 0$$

Ces solutions sont monogènes puisqu'elles satisfont l'équation de Dirac correspondant aux sous-espaces $E_i = \text{span}\{e_1, e_2, e_i\}$ ($i = 3, 4, 5$) et par conséquent l'équation de Dirac dans $\mathbb{R}_{5,0}$. La somme des solutions f_c est toujours monogènes dans $\mathbb{R}_{5,0}$ et elle satisfait les conditions aux bords du problème initial 2.7.

Nous appellerons f_c le signal monogène couleur d'un signal s . Il a la forme suivante :

$$f_c = h_p^3 * f_3 e_3 + h_p^4 * f_4 e_4 + h_p^5 * f_5 e_5 \\ + h_p^3 * h_R^3 * f_3 e_3 + h_p^4 * h_R^4 * f_4 e_4 + h_p^5 * h_R^5 * f_5 e_5 \quad (2.9)$$

où

$$\begin{cases} h_p^i = \frac{x_i}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}}, (i = 3, 4, 5) \\ h_R^i = \frac{x_1 e_1 + x_2 e_2}{2\pi(x_1^2 + x_2^2)^{3/2}} e_i \end{cases}$$

La fonction h_p^i est le noyau de Poisson 2D et la fonction h_R^i est le noyau de Riesz.

2.4.3 Représentation polaire d'une image couleur

Soit $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ une image couleur définie par $s(x_1, x_2) = (f_3(x_1, x_2), f_4(x_1, x_2), f_5(x_1, x_2))$. Le signal monogène couleur associé à s est de la forme :

$$f_c = A_1 e_1 + A_2 e_2 + A_3 e_3 + A_4 e_4 + A_5 e_5$$

où

$$\begin{cases} A_1 = h_p^3 * h_{R_{x_1}} * f_3 e_3 + h_p^4 * h_{R_{x_1}} * f_4 e_4 + h_p^5 * h_{R_{x_1}} * f_5 e_5, \\ A_2 = h_p^3 * h_{R_{x_2}} * f_3 e_3 + h_p^4 * h_{R_{x_2}} * f_4 e_4 + h_p^5 * h_{R_{x_2}} * f_5 e_5, \\ A_i = h_p^i * f_i, i = 3, 4, 5 \end{cases}$$

et

$$h_{R_{x_i}} = \frac{x_i}{2\pi(x_1^2 + x_2^2)^{3/2}}$$

Il contient deux sortes d'information :

- les fonctions A_1 et A_2 correspondent aux structures verticales et horizontales lissées au sens de Riesz (voir Fig. 2.6).
- les fonctions A_3 , A_4 et A_5 correspondent aux plans couleurs lissés dans l'espace colorimétrique engendré par $\{e_3, e_4, e_5\}$.

Si $V = ue_1 + ve_2 + ae_3 + be_4 + ce_5 \in \mathbb{R}_{5,0}$ est un vecteur unitaire contenant des informations de structure (u, v) et de couleur (a, b, c) alors le produit géométrique $f_c V$ est par définition un 2-verseur de l'algèbre $\mathbb{R}_{5,0}$ donné par :

$$f_c V = \langle f_c V \rangle_0 + \langle f_c V \rangle_2$$

où $p_V = \langle f_c V \rangle_0$ est la partie scalaire et $\langle f_c V \rangle_2$ est la partie bivectorielle. Le produit géométrique $f_c V$ peut être considéré comme un nombre complexe :

$$f_c V = \langle f_c V \rangle_0 + i |\langle f_c V \rangle_2|$$

En effet, si A et B sont deux vecteurs de $\mathbb{R}_{5,0}$, avec le produit géométrique $AB = A.B + A \wedge B$, nous pouvons toujours trouver un vecteur B_1 tel que $A \wedge B = A \wedge B_1 = AB_1$. Alors $(A \wedge B)^2 = (AB_1)^2 = -A^2 B_1^2$ est un nombre réel. En particulier :

$$\begin{aligned} AB &= A.B + \sqrt{A^2 B_1^2} \frac{AB_1}{\sqrt{A^2 B_1^2}} \\ &= \langle AB \rangle_0 + |\langle AB \rangle_2| \frac{\langle AB \rangle_2}{|\langle AB \rangle_2|} \end{aligned}$$

$$\text{avec } \left(\frac{\langle AB \rangle_2}{|\langle AB \rangle_2|} \right)^2 = -1.$$

Nous pouvons maintenant définir une amplitude et un angle de rotation associés au 2-verseur $f_c V$:

$$\varphi_V = \arg(f_c V) = \arctan \left(\frac{|\langle f_c V \rangle_2|}{\langle f_c V \rangle_0} \right) \quad (2.10)$$

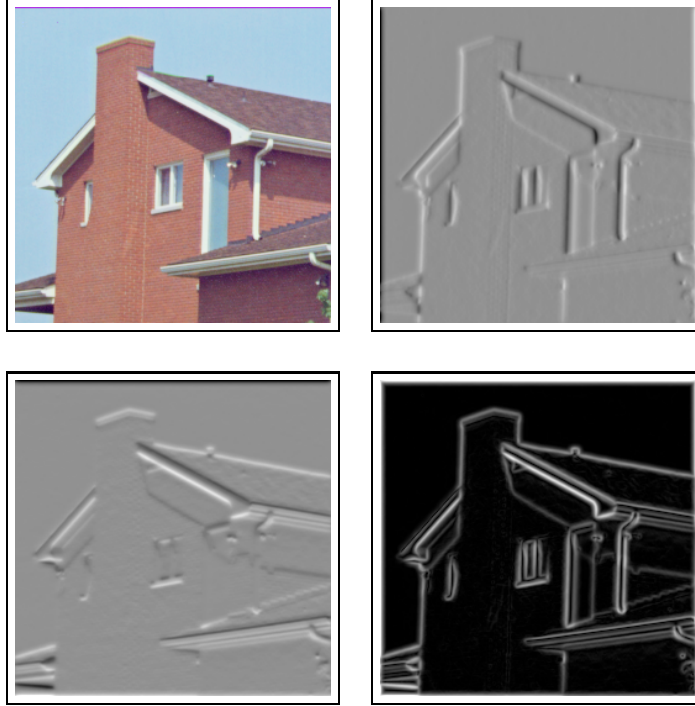


FIGURE 2.6 – De gauche à droite, de haut en bas : image originale, composantes A_1 , A_2 et $\sqrt{A_1^2 + A_2^2}$ respectivement.

et

$$\mathcal{A}_V = |f_c||V| = |f_c| \quad (2.11)$$

Ainsi, $f_c V$ est un 2-verseur polaire que nous pouvons écrire :

$$f_c V = \mathcal{A}e^{B\varphi_V} \quad (2.12)$$

où $B = \frac{\langle f_c V \rangle_2}{|\langle f_c V \rangle_2|}$ est le bivecteur unitaire engendrant le plan de rotation entre les vecteurs V et f_c .

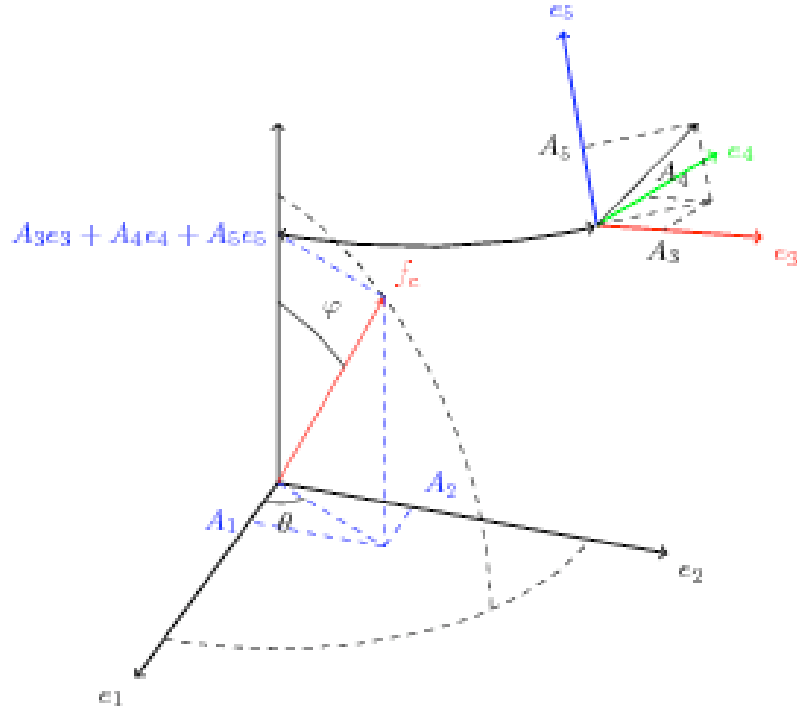
L'angle φ_V mesure la distance angulaire entre f_c et un vecteur choisi V dans $\mathbb{R}_{5,0}$, *i.e.* il donne une mesure de similarité entre un pixel muni de son information de couleur et de structure et un vecteur contenant une couleur ainsi qu'une structure choisies.

Définition 1. La représentation polaire d'une image couleur s est l'application $\chi : \mathbb{R}^5 \longrightarrow \mathbb{R}_{5,0}$ définie par $\chi(V) = f_c V$.

Nous présentons deux exemples pour illustrer cette représentation polaire.

Exemple 1 :

Dans cet exemple, nous considérons l'image couleur $I(x, y) =$

FIGURE 2.7 – Représentation du signal couleur monogène f_c dans $\mathbb{R}_{5,0}$.

$(r(x, y), v(x, y), b(x, y))$ de la figure 2.8 dans l'espace couleur RVB et nous lui associons sa représentation polaire $\chi(V)$ en prenant $V = e_3$, *i.e.* V code la couleur rouge. Calculons explicitement $\chi(V)$:

$$\begin{aligned}
 \chi(V) &= f_c V \\
 &= f_c e_3 \\
 &= (A_1 e_1 + A_2 e_2 + A_3 e_3 + A_4 e_4 + A_5 e_5) e_3 \\
 &= A_3 + A_1 e_1 e_3 + A_2 e_2 e_3 + A_4 e_4 e_3 + A_5 e_5 e_3 \\
 &= A_3 + \frac{A_1 e_1 e_3 + A_2 e_2 e_3 + A_4 e_4 e_3 + A_5 e_5 e_3}{|A_1 e_1 e_3 + A_2 e_2 e_3 + A_4 e_4 e_3 + A_5 e_5 e_3|} |A_1 e_1 e_3 + A_2 e_2 e_3 + A_4 e_4 e_3 + A_5 e_5 e_3| \\
 &= A_3 + B \sqrt{A_1^2 + A_2^2 + A_4^2 + A_5^2}
 \end{aligned}$$

Nous obtenons donc une amplitude \mathcal{A} et un angle φ_V tels que $\chi(V) = \mathcal{A} e^{B\varphi_V}$ avec

$$B = \frac{A_1 e_1 e_3 + A_2 e_2 e_3 + A_4 e_4 e_3 + A_5 e_5 e_3}{|A_1 e_1 e_3 + A_2 e_2 e_3 + A_4 e_4 e_3 + A_5 e_5 e_3|}.$$

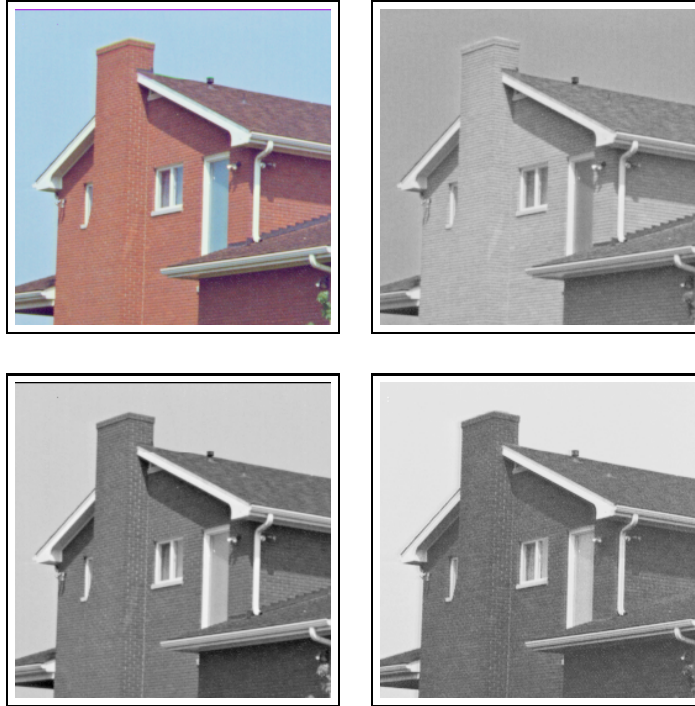


FIGURE 2.8 – Décomposition dans RVB. De gauche à droite : image originale, canal rouge, canal vert et canal bleu.

De plus, il est intéressant de remarquer que :

$$\begin{aligned}
 \chi(V) &= A_3 + B\sqrt{A_1^2 + A_2^2 + A_4^2 + A_5^2} \\
 &= \mathcal{A}e^{B\varphi_V} \\
 &= \mathcal{A}(\cos \varphi_V + B \sin \varphi_V)
 \end{aligned}$$

Ceci implique en particulier que la partie "réelle" $\mathcal{R}e(\chi(V)) = A_3 = \mathcal{A} \cos \varphi_V$. Nous pouvons donc reconstruire le canal rouge lissé à une certaine échelle grâce à \mathcal{A} et φ_V . La figure 2.9, ligne du bas, illustre cette reconstruction et montre l'erreur absolue ($|A_3 - \mathcal{A} \cos \varphi_V|$) entre le canal rouge original et celui reconstruit. Il est à noter que l'erreur absolue maximale est d'environ 9.58×10^{-10} .

Dans la figure 2.9, la ligne du haut montre φ_V et \mathcal{A} . Comme nous pouvons le constater, les angles sont les plus faibles (zones sombres) sur les parties de la maison proches de la couleur rouge.

Exemple 2 :

Pour les applications traitées dans la partie 3, nous utilisons seulement des vecteurs V dont les deux premières composantes u et v sont égales à 0, comme dans l'exemple précédent. Cela s'explique par le fait que nous sommes principalement

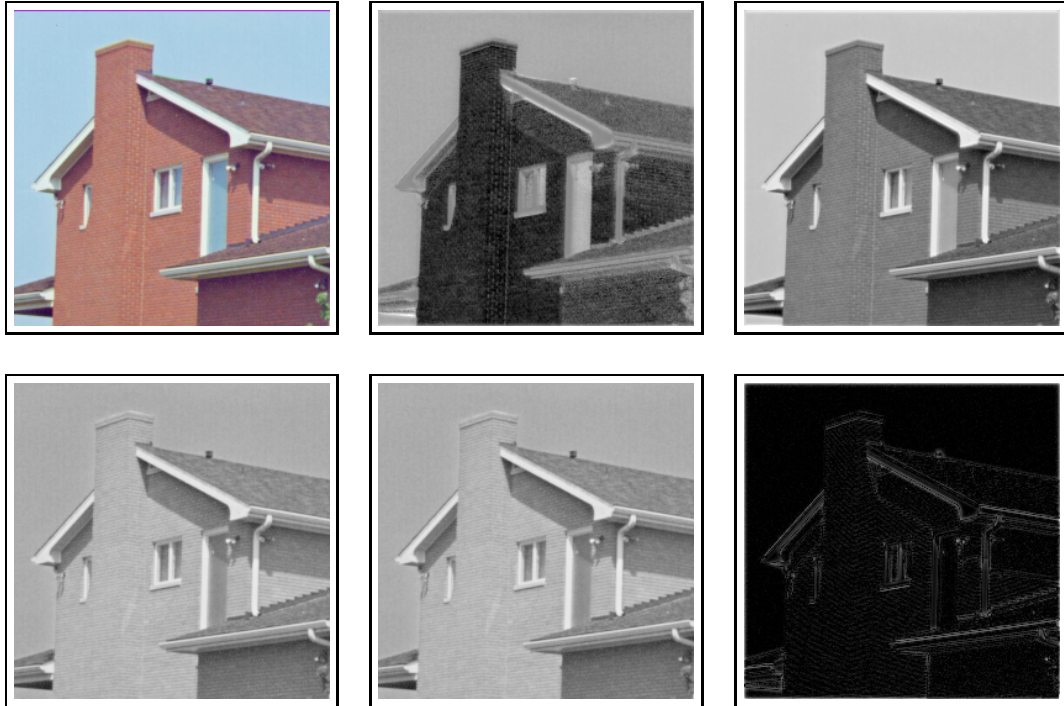


FIGURE 2.9 – Représentation polaire $\chi(V)$ d’une image couleur dans l’espace RVB avec $V = e_3$ (*i.e.* rouge). Première ligne de gauche à droite : image originale, φ_V et amplitude $|f_c|$. Deuxième ligne de gauche à droite : canal rouge, $|f_c| \cos \varphi_V$ et l’erreur absolue.

intéressés par les zones homogènes en couleur, bien que f_c contienne une information locale de structure.

L’exemple qui suit montre des résultats de φ_V avec V contenant des composantes de structure u et v non triviales. Ces dernières sont choisies de la manière suivante :

$$u = v = \max_{(x,y)}(\max(|A_1(x,y)|), \max(|A_2(x,y)|))$$

La fonction d’angle φ_V est donc petite quand les variations horizontales et verticales dans une couleur donnée (rouge, jaune et bleue dans cet exemple) sont fortes. Cela nous permet donc de détecter des coins dans une couleur donnée. La figure 2.10, ligne du haut, montre les angles φ_V et nous constatons que les angles les plus faibles (zones les plus sombres) sont bien sur les coins de la couleur choisie. En prenant les quatre plus petites valeurs dans chacune des images d’angles, nous obtenons la ligne du bas de la figure 2.10.

A ce stade, il n’est pas clair que la fonction d’angle φ_V soit une phase, au sens usuel du terme, couleur. Ce n’est pas un problème pour les applications que nous traitons dans la suite. Nous donnons néanmoins dans le paragraphe suivant des exemples de calculs concernant l’analyse de signaux et comment la décomposition

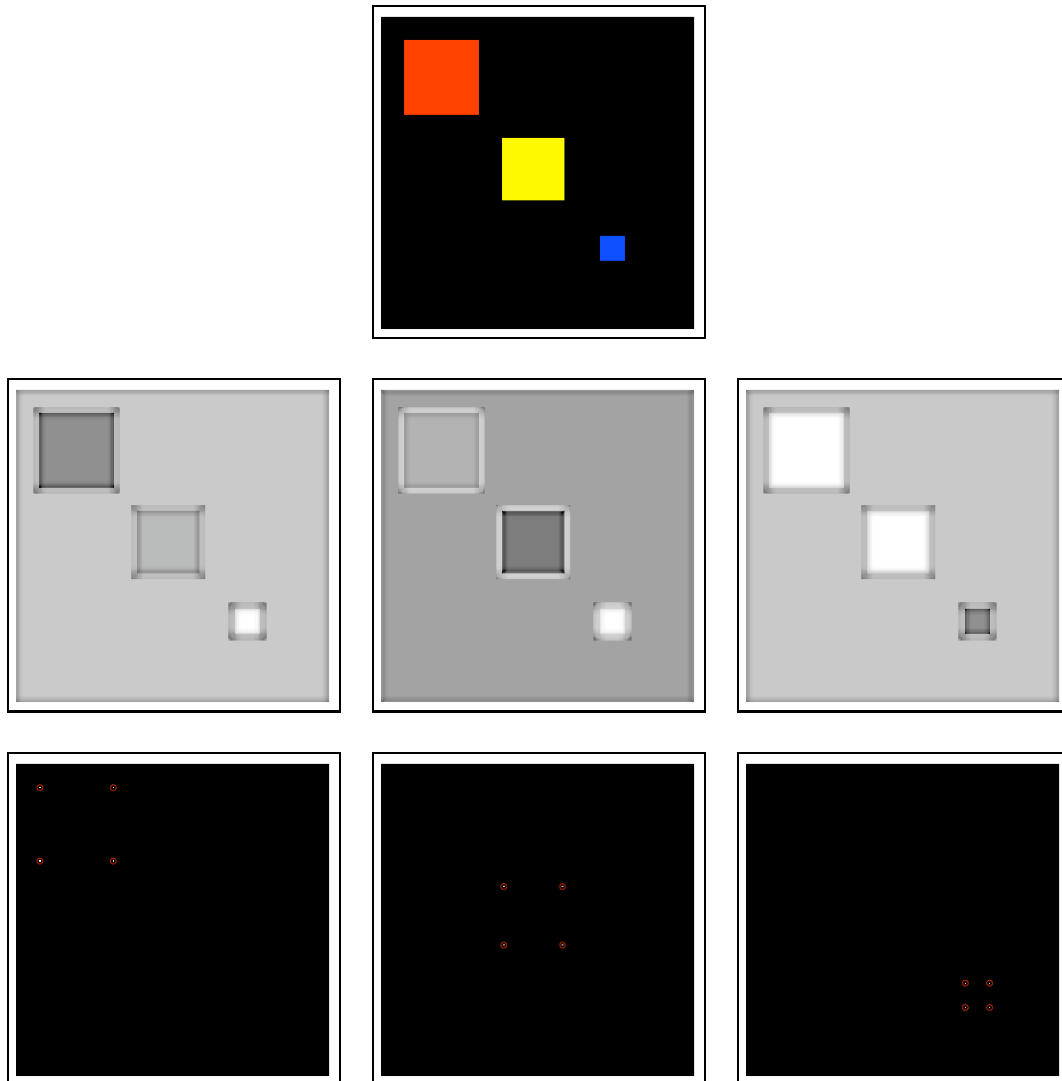


FIGURE 2.10 – Exemples pour $u = v \neq 0$. De haut en bas, de gauche à droite : image originale, angles φ_v dans la couleur rouge, jaune et bleu respectivement, valeurs minimales de φ_v entourées par des cercles rouges.

polaire peut être mise en relation avec la phase locale du signal monogène définie dans [Felsberg 2001].

2.4.4 Fonction d'angle φ_V et représentation polaire χ

Comme mentionné précédemment, la fonction d'angle φ_V est principalement utilisée dans les applications qui suivent comme une mesure de similarité. Nous pensons cependant que la représentation polaire que nous avons introduite devrait permettre à terme de dégager une notion rigoureuse mathématiquement de phase couleur. Nous nous bornons ici à quelques exemples de synthèse et évoquons rapidement une approche plus générale du problème.

Soit le signal $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ de composantes f_3, f_6 et f_9 la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}_{9,0}$ définie par :

$$f(x_1, x_2) = f_3(x_1, x_2)e_3 + f_6(x_1, x_2)e_6 + f_9(x_1, x_2)e_9.$$

Le signal monogène couleur correspondant avec le paramètre d'échelle τ est :

$$\begin{aligned} \tilde{f}_c(x_1, x_2, \tau) = & h_p * f_3e_3 + h_p * f_6e_6 + h_p * f_9e_9 \\ & + h_p * h_R^3 * f_3e_3 + h_p * h_R^6 * f_6e_6 + h_p * h_R^9 * f_9e_9 \end{aligned} \quad (2.13)$$

où

$$\left\{ \begin{array}{l} h_p = \frac{\tau}{2\pi(x_1^2 + x_2^2 + \tau^2)^{3/2}} \text{ est le noyau de Poisson 2D} \\ h_R^i = \frac{x_1e_{i-2} + x_2e_{i-1}}{2\pi(x_1^2 + x_2^2)^{3/2}} e_i \text{ est le noyau de Riesz (i=3,6,9)} \end{array} \right.$$

Puisque nous avons les égalités suivantes :

$$\left\{ \begin{array}{l} h_p * h_R^3 * f_3e_3 = h_p * h_{R_{x_1}} * f_3e_1 + h_p * h_{R_{x_2}} * f_3e_2 \\ h_p * h_R^6 * f_6e_6 = h_p * h_{R_{x_1}} * f_6e_4 + h_p * h_{R_{x_2}} * f_6e_5 \\ h_p * h_R^9 * f_9e_9 = h_p * h_{R_{x_1}} * f_9e_7 + h_p * h_{R_{x_2}} * f_9e_8 \end{array} \right.$$

il s'avère que les structures horizontales et verticales des différents canaux sont dans des composantes différentes et indépendantes. Comme précédemment, si $V = ae_3 + be_6 + ce_9$ nous considérons le 2-verseur de $\mathbb{R}_{9,0}$:

$$\tilde{\chi}(V) = \tilde{f}_c V = \tilde{A}_V e^{i\tilde{\varphi}_V}$$

Maintenant, décrivons trois exemples simples de signaux i1D, voir la figure 2.11. Les signaux i1D sont donnés par les mêmes fonctions que pour le signal monogène (voir paragraphe 2.3.3.2) :

$$f_i(x_1, x_2) = a_i \cos(x_1 \cos \theta_i + x_2 \cos \theta_i + \varphi_i)$$

avec $i = 3, 6, 9$ (voir [Wietzke 2009]).

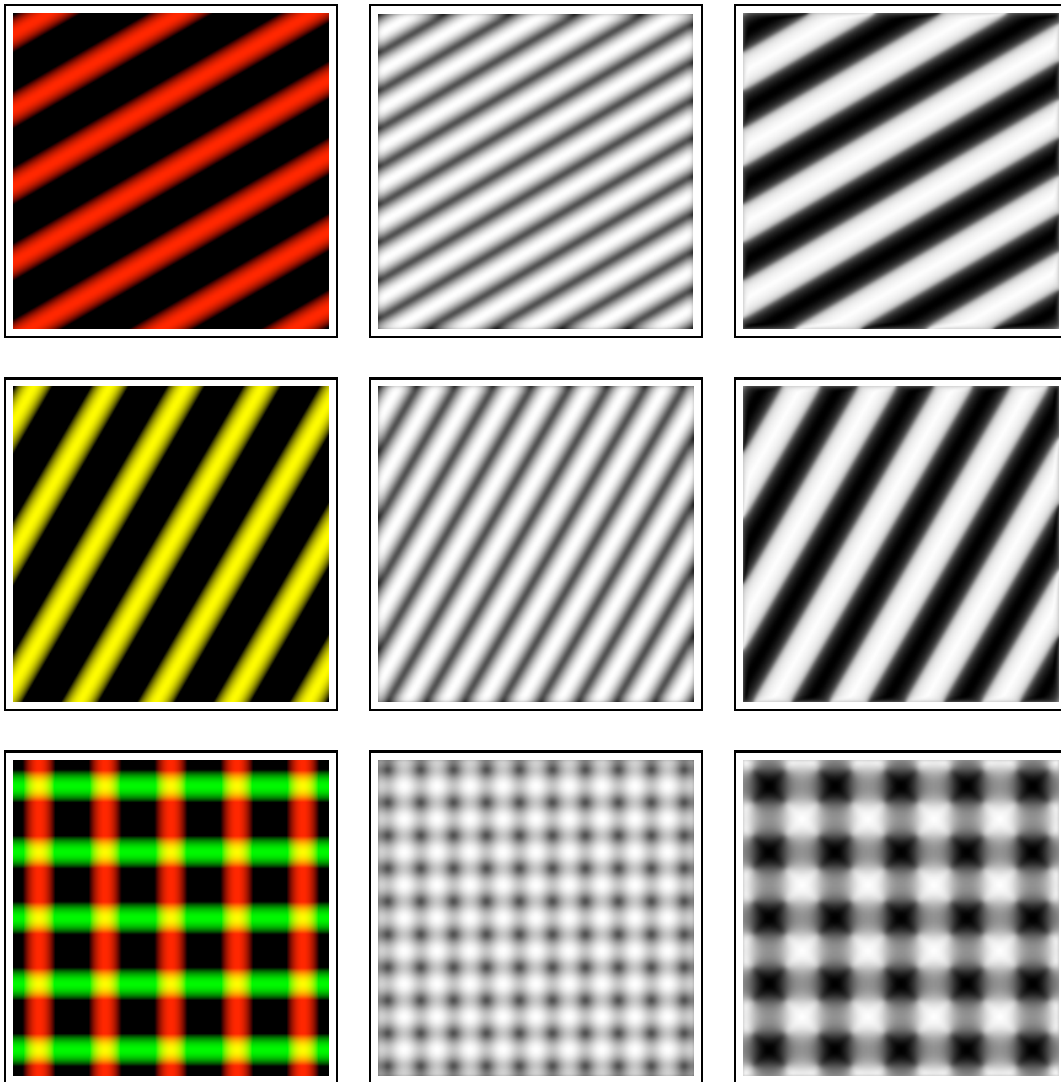


FIGURE 2.11 – Première colonne : images originales. Deuxième colonne : amplitude $\tilde{\mathcal{A}}_V$. Troisième colonne : angle $\tilde{\varphi}_V$.

Premier exemple :

Soit $\varphi_6 = \varphi_9 = \pi/2$ (voir Fig. 2.11 première colonne). La fonction test au point $(0, 0)$ donne $f_3 = a_3 \cos \varphi_3$. Le signal monogène couleur correspondant est :

$$\tilde{f}_c = h_p * h_{R_{x_1}} * f_3 e_1 + h_p * h_{R_{x_2}} * f_3 e_2 + h_p * f_3 e_3$$

C'est précisément le signal monogène couleur de la première composante de f (qui est le canal rouge de l'image couleur considérée). Pour $V = e_3$, nous obtenons :

$$\begin{aligned} \tilde{\chi}(V) &= \tilde{f}_c e_3 \\ &= h_p * f_3 + h_p * h_{R_{x_1}} * f_3 e_1 e_3 + h_p * h_{R_{x_2}} * f_3 e_2 e_3 \end{aligned}$$

et l'amplitude, la phase et l'orientation sont donnés à l'échelle $\tau = 0$ par

$$\left\{ \begin{array}{l} \tilde{A}_V = \sqrt{f_3^2 + (h_{R_{x_1}} * f_3)^2 + (h_{R_{x_2}} * f_3)^2} \\ \tilde{\varphi} = \arctan \frac{\sqrt{(h_{R_{x_1}} * f_3)^2 + (h_{R_{x_2}} * f_3)^2}}{f_3} \\ \tilde{\theta} = \arctan \frac{h_{R_{x_2}} * f_3}{h_{R_{x_1}} * f_3} \end{array} \right.$$

Plus précisément (voir [Wietzke 2009] pour des calculs similaires) :

$$\left\{ \begin{array}{l} \tilde{A}_V = a_3 = \sqrt{(a_3 \cos \varphi_3)^2 + (a_3 \sin \varphi_3 \cos \theta_3)^2 + (a_3 \sin \varphi_3 \sin \theta_3)^2} \\ \tilde{\varphi} = \varphi_3 = \arctan \frac{\sqrt{(a_3 \sin \varphi_3 \cos \theta_3)^2 + (a_3 \sin \varphi_3 \sin \theta_3)^2}}{a_3 \cos \varphi_3} \\ \tilde{\theta} = \theta_3 = \arctan \frac{a_3 \sin \varphi_3 \sin \theta_3}{a_3 \sin \varphi_3 \cos \theta_3} \end{array} \right.$$

Cela montre que le signal monogène couleur peut être utilisé pour retrouver l'amplitude locale, l'orientation locale et la phase locale du signal original. Notons que les mêmes résultats peuvent être obtenus en travaillant directement avec le signal monogène f_m appliqué à la fonction f_3 .

Deuxième Exemple :

Soient les fonctions tests (voir la figure 2.11 deuxième ligne)

$$f_3 = f_6 = a \cos \varphi \cos(\pi/4) \quad \text{et} \quad f_9 = 0$$

Si nous choisissons $V = \frac{1}{\sqrt{2}}(e_3 + e_6)$ alors

$$\left\{ \begin{array}{l} \tilde{\mathcal{A}}_V = a = \sqrt{a^2 \cos \varphi + 2\left(\frac{a}{\sqrt{2}} \sin \varphi \cos \theta\right)^2 + 2\left(\frac{a}{\sqrt{2}} \sin \varphi \sin \theta\right)^2} \\ \tilde{\varphi}_V = \varphi = \arctan \frac{\sqrt{(a \sin \varphi \cos \theta)^2 + (a \sin \varphi \sin \theta)^2}}{a \cos \varphi} \\ \tilde{\theta}_V = \theta = \arctan \frac{a \sin \varphi \sin \theta}{a \sin \varphi \cos \theta} \end{array} \right.$$

Cet exemple justifie l'introduction du signal \tilde{f}_c puisque ces résultats ne peuvent être obtenus en utilisant le signal f_c de $\mathbb{R}_{5,0}$. Cela montre que le signal monogène couleur peut résoudre le problème inverse pour les structures linéaires dans chacun des plans couleur. L'avantage ici est qu'aucune combinaison des canaux rouge, vert et bleu n'est requis étant donné que l'information couleur est portée par le vecteur V .

Troisième exemple :

Ce dernier exemple a pour but de montrer la pertinence de la représentation polaire quand les différents canaux comportent des signaux i1D paramétrés différemment (voir Fig. 3 troisième ligne). Soient f_3 , f_6 et f_9 définies par :

$$\left\{ \begin{array}{l} f_3(x_1, x_2) = a \cos(x_1 \cos \theta_1 + x_2 \cos \theta_1 + \varphi_1) \\ f_6(x_1, x_2) = a \cos(x_1 \cos \theta_2 + x_2 \cos \theta_2 + \varphi_2) \\ f_9(x_1, x_2) = 0 \end{array} \right.$$

La figure 2.11, troisième ligne, donne les résultats fournis par la représentation polaire $\tilde{f}_c V$ où V est le vecteur contenant la couleur jaune. Le principal avantage par rapport au signal monogène, est que la fonction d'angle $\tilde{\varphi}_V$ contient non seulement des informations géométriques et structurelles mais aussi une information de couleur. Les zones noires correspondent aux petites valeurs de $\tilde{\varphi}_V$ et donc indiquent les régions correspondant à la couleur choisie.

Dans [Felsberg 2001] la représentation polaire du signal monogène f_m est définie par :

$$f_m(x_1, x_2) = |f_m(x_1, x_2)| \exp(\psi^\perp(x_1, x_2))$$

où

$$\exp(\psi^\perp) = \cos(\varphi) + \cos(\theta) \sin(\varphi)i + \sin(\theta) \sin(\varphi)j$$

agit comme produit de rotations . Cela est intimement lié au fait que le groupe des quaternions unitaires est un groupe de spineurs Spin(3) de l'algèbre de Clifford $\mathbb{R}_{3,0}$.

Une telle approche peut être généralisée dans notre contexte. Le 2-verseur $f_c V$ peut être écrit :

$$f_c V = \mathcal{A}_V \eta_V$$

où η_V est un 2-verseur unitaire. En utilisant le fait que chaque spineur est l'exponentielle d'un bivecteur nous amène à :

$$f_c V = \mathcal{A}_V \exp(\beta_V)$$

pour un bivecteur β_V . Ce bivecteur peut être décomposé avec l'algorithme d'Hestenes et Sobczyk (voir Annexe F) pour introduire des rotations caractéristiques. Cette décomposition semble être plus adaptée à décrire la représentation polaire χ en terme de traitement du signal. Cela permet en particulier d'introduire des axes de rotations similaires aux axes utilisés dans [Felsberg 2001] mais maintenant dépendants du vecteur de référence choisi V . Puisque dans cette thèse, nous nous intéressons principalement à des applications sur des images et vidéos couleur qui dépendent de la fonction φ_V considérée comme mesure de similarité, nous n'aborderons pas en détail ce point de vue.

Applications

Ce chapitre est consacré à la mise en application des constructions théoriques du chapitre précédent. Comme nous l'avons déjà mentionné les exemples que nous traitons reposent essentiellement sur l'utilisation de la fonction d'angle φ_V . Il est bien évident que nous n'utilisons qu'une partie des informations contenues dans le signal monogène couleur et que d'autres propriétés, comme la décomposition/reconstruction à partir de la représentation polaire, pourraient être mises à profit. Ceci étant, il s'avère néanmoins que la fonction φ_V contient déjà beaucoup d'informations de structures et de couleur et qu'il est intéressant de la tester sur des applications - flot optique, détection de contours, suivi d'objets, points d'intérêt - qui sont le plus couramment abordées en utilisant des images en niveau de gris.

Sommaire

3.1 Flot optique couleur	48
3.1.1 Introduction	48
3.1.2 Méthode CLG (Combined Local-Global)	49
3.1.3 Méthode CLG modifiée	50
3.1.4 Résultats	50
3.1.5 Nouvelle contrainte	51
3.1.6 Résultats	53
3.2 Segmentation multi-échelle d'images couleur	55
3.2.1 Détection de contours métrique	55
3.2.2 Méthode de segmentation	56
3.2.3 Résultats	57
3.2.4 Détection de contours métrique et multi-échelle	57
3.2.5 Méthode de segmentation	60
3.2.6 Segmentation multi-échelle	61
3.2.7 Résultats	63
3.3 Suivi d'objets couleur	65
3.3.1 Un algorithme naïf de block-matching	65
3.3.2 Filtrage particulière	68
3.3.3 Résultats	71
3.4 Détection de points d'intérêt	74
3.4.1 Détecteur de Harris et géométrie différentielle	74
3.4.2 Géométrie différentielle dans un espace-échelle	76
3.4.3 Résultats	77
3.4.4 Complément 1 : H_{3D} couleur	87
3.4.5 Complément 2 : H_{3D} et signal monogène couleur	88

3.1 Flot optique couleur

3.1.1 Introduction

En traitement de la vidéo, extraire le flot optique consiste à trouver le champ de vecteurs associés aux déplacements. Une des approches très couramment utilisée repose sur le fait que l'intensité de l'image doit être intégrale première du champ recherché. On obtient à partir du développement de Taylor :

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t + o(\cdot)$$

l'équation dite de contrainte :

$$\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} = 0$$

où $(u, v) = \left(\frac{\delta x}{\delta t}, \frac{\delta y}{\delta t} \right)$.

Evidemment, cette simple équation n'est pas suffisante pour déterminer les deux inconnues u et v (problème d'ouverture). Pour les gradients non nuls de l'image, il est possible de déterminer la composante parallèle à $\nabla I = (I_x, I_y)$, *i.e.* normale aux contours dans l'image. Il s'agit du flot normal donné par :

$$w_n = -\frac{I_t}{|\nabla I|} \frac{\nabla I}{|\nabla I|}$$

3.1.1.1 Lucas/Kanade

Afin de palier le problème d'ouverture, Lucas et Kanade [Lucas 1981] proposent une approche locale qui consiste à considérer le flot optique constant dans un certain voisinage de taille ρ . Dans ce cas, il est possible de déterminer les deux constantes u et v au voisinage du point (x, y, t) en minimisant la fonction :

$$E_{LK}(u, v) = G_\rho * ((I_x u + I_y v + I_t)^2)$$

Dans cette équation, G_ρ est une gaussienne d'écart-type ρ , nous notons I_x , I_y et I_t les dérivées spatiales et temporelles de l'image.

Un minimum (u, v) de E_{LK} satisfait $\frac{\partial E_{LK}}{\partial u} = 0$ et $\frac{\partial E_{LK}}{\partial v} = 0$. Cela donne le système linéaire suivant :

$$\begin{pmatrix} G_\rho * I_x^2 & G_\rho * I_x I_y \\ G_\rho * I_x I_y & G_\rho * I_y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -G_\rho * I_x I_t \\ -G_\rho * I_y I_t \end{pmatrix}$$

Ce système peut être résolu tant que la matrice est inversible, ce qui n'est pas le cas dans les régions homogènes où le gradient est nul. Dans d'autres régions, la plus petite valeur propre de la matrice peut être proche de zéro et le problème d'ouverture reste donc présent. Tout cela produit un flot optique non dense, ce qui constitue un inconvénient non négligeable de cette méthode locale.

3.1.1.2 Horn et Schunk

Pour avoir un flot optique dense, Horn et Schunk [Horn 1981] proposent une méthode globale utilisant un schéma de régularisation. Cette méthode différentielle est la base de nombreuses recherches dans le domaine du flot optique et est considérée comme le travail fondateur. Il s'agit de minimiser la fonction d'énergie suivante :

$$E_{HS}(u, v) = \int_{\Omega} (I_x u + I_y v + I_t)^2 + \alpha(|\nabla u|^2 + |\nabla v|^2) dx dy$$

où $\alpha > 0$ est un poids de lissage qui sert en tant que paramètre de régularisation. Plus les valeurs de α sont grandes, plus la pénalisation des gradients du flot est forte ce qui donne un champs de vecteurs plus lisse.

Pour minimiser cette fonction convexe, nous utilisons les équations d'Euler-Lagrange :

$$\begin{cases} \Delta u - \frac{1}{\alpha}(I_x^2 u + I_x I_y v + I_x I_t) = 0 \\ \Delta v - \frac{1}{\alpha}(I_x I_y u + I_y^2 v + I_y I_t) = 0 \end{cases}$$

avec Δ opérateur de Laplace :

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

La solution de ces équations n'est pas seulement unique, elle bénéficie également d'un effet de "remplissage". Aux endroits où $|\nabla I| \approx 0$, il n'y a pas d'estimation du flot local possible, mais le terme de régularisation $|\nabla u|^2 + |\nabla v|^2$ résout le problème en prenant l'information dans le voisinage. Le flot optique produit est directement dense ce qui est clairement un avantage par rapport aux méthodes locales.

Le flot optique est un domaine très étudié dans la communauté de vision par ordinateur. Beaucoup de méthodes ont été proposées depuis les travaux fondateurs de Horn/Schunck [Horn 1981], et de Lucas/Kanade [Lucas 1981] mais la plupart d'entre elles repose sur des images en niveau de gris. De nos jours, de plus en plus de recherches essayent d'incorporer l'information de couleur pour améliorer le flot optique. Notre formalisme permet justement d'apporter des éléments de réponse et peut potentiellement être utilisé dans n'importe quelle méthode différentielle de flot optique, c'est ce que nous présentons dans la suite.

3.1.2 Méthode CLG (Combined Local-Global)

A. Bruhn et J. Weickert [Bruhn 2005] sont les premiers à combiner l'approche globale de Horn/Schunck et l'approche locale de Lucas/Kanade dans une nouvelle méthode qui donne un champs de vecteurs mouvements dense à 100% et robuste aux bruits. Leur idée est de minimiser une fonction d'énergie pour estimer un champs de vecteurs mouvements :

$$E(w) = \int_{\Omega} (\psi_1(w^T J_{\rho}(\nabla_3 I)w) + \alpha \psi_2(|\nabla w|^2)) dx dy$$

avec

$$(P) \begin{cases} w = (u, v, 1)^T, \\ |\nabla w|^2 = |\nabla u|^2 + |\nabla v|^2, \\ \nabla_3 I = (I_x, I_y, I_t)^T, \\ J_\rho(\nabla_3 I) = K_\rho * (\nabla_3 I \nabla_3 I^T) \end{cases}$$

où Ω est le domaine de l'image, α est un paramètre de régularisation, K_ρ est un noyau gaussien avec un écart-type ρ , ψ_1 et ψ_2 sont deux pénalités non quadratiques de la forme suivante :

$$\psi_i = 2\beta_i^2 \sqrt{1 + \frac{s^2}{\beta_i^2}}, i \in \{1, 2\}$$

où β_1 and β_2 sont des paramètres d'échelle. Les auteurs donnent dans [Bruhn 2005] un algorithme efficace pour la résolution.

3.1.3 Méthode CLG modifiée

En considérant la fonction d'angle $\varphi_V(x, y, t)$ correspondant à la représentation polaire $\chi(V)$ d'une séquence d'images couleur, où V est un vecteur constant codant une couleur choisie, nous remplaçons la fonction d'intensité I par φ_V (voir [Zang 2007] pour une approche similaire) ce qui donne l'équation suivante :

$$\varphi_V(x, y, t + 1) = \varphi_V(x, y, t)$$

En calculant un développement de Taylor du premier ordre, cela nous mène à l'équation de flot optique :

$$\varphi_x u + \varphi_y v + \varphi_t = 0.$$

Nous proposons de minimiser la fonction d'énergie suivante :

$$E(w) = \int_{\Omega} (\psi_1(w^T J_\rho(\nabla_3 \varphi_V)w) + \alpha \psi_2(|\nabla w|^2)) dx dy$$

Les paramètres sont les mêmes que dans (P) et la méthode de résolution est la même que pour A. Bruhn *et al* [Bruhn 2005].

3.1.4 Résultats

La séquence utilisée dans cette application est la séquence de deux images successives "Dimetrodon" et la vérité-terrain associée provenant de la base de données [Baker 2007]. Cette séquence est difficile car elle est composée de textures superposées.

Les paramètres (a, b, c) du vecteur de référence $V_{ref} = ue_1 + ve_2 + ae_3 + be_4 + ce_5$ sont choisis en prenant la médiane des valeurs de chaque plan de couleur et $(u, v) = (0, 0)$ pour les composantes de structure. D'autres choix comme la moyenne ou bien même une couleur choisie *a priori* auraient pu être utilisés. Nous appliquons la méthode CLG 2D avec pénalité quadratique et nous gardons les paramètres par défaut pour toutes les séquences d'images. Pour faire simple, nous choisissons de ne pas optimiser les paramètres de la CLG qui peuvent être différents suivant que l'on

traite des images en niveau de gris ou couleur, et nous voulons étudier la robustesse de notre approche aux bruits et au changement d'illuminant. Dans un premier temps, un bruit gaussien ($\sigma = 0.1$) est ajouté à chaque image en niveau de gris ou couleur. Les résultats (table 3.1 et figure 3.1) montre que notre méthode donne 29% d'erreur angulaire (AAE=Average Angular Error)[Bruhn 2005] en moins. Dans un deuxième temps, nous vérifions que notre fonction d'angle φ_V est robuste au changement d'illuminant en simulant des conditions d'éclairage plus faible. Cela est fait en divisant chaque valeur de pixel, couleur ou niveau de gris, de la deuxième image de la séquence par 2. Les résultats montre que notre méthode n'augmente pas l'erreur angulaire AAE tandis que la méthode standard produit un flot très altéré.

TABLE 3.1 – 2D-CLG – Erreur angulaire moyenne (AAE=Average Angle Error)

Condition	Image	
	Niveau de gris	Couleur
Normale	10.44	10.52
Bruitée	23.10	16.40
Illumination	88.44	10.52

3.1.5 Nouvelle contrainte

Dans cette section, nous nous sommes intéressés à trouver le champs de vecteurs correspondant au mouvement d'objets d'une couleur choisie. L'approche étend le travail décrit précédemment [Demarcq 2009a] et repose sur la méthode CLG présentée dans [Bruhn 2005] avec l'ajout d'une nouvelle contrainte : l'intégrale première considérée dans la fonction d'énergie est la fonction d'angle φ_V .

Soit $\varphi_V(x, y, t)$ la fonction d'angle correspondant à la représentation polaire $\chi(V)$ d'une vidéo couleur où V est un vecteur constant codant une couleur donnée. Remplacer la fonction d'intensité I par φ_V amène à l'énergie suivante (voir aussi [Zang 2007])

$$E(w) = \int_{\Omega} (w^T J_{\rho}(\nabla_3 \varphi_V) w) + \alpha(|\nabla_3 w|^2) dx dy dt$$

Comme déjà dit, la fonction φ_V peut être vue comme une mesure de similarité. L'idée est d'implémenter le champs de vecteurs associé aux régions où les valeurs de φ_V sont petites. Pour cette raison, il est nécessaire d'ajouter un terme de pénalité qui annule la norme du champs de vecteurs où les valeurs de φ_V sont grandes. Nous noterons φ à la place de φ_V dans la suite. L'énergie résultante est donnée par :

$$E = \int_{\Omega} (K_{\rho} * (\varphi_x u + \varphi_y v + \varphi_t)^2) + \alpha(|\nabla u|^2 + |\nabla v|^2) + \beta g(\varphi)(u^2 + v^2) dx dy$$

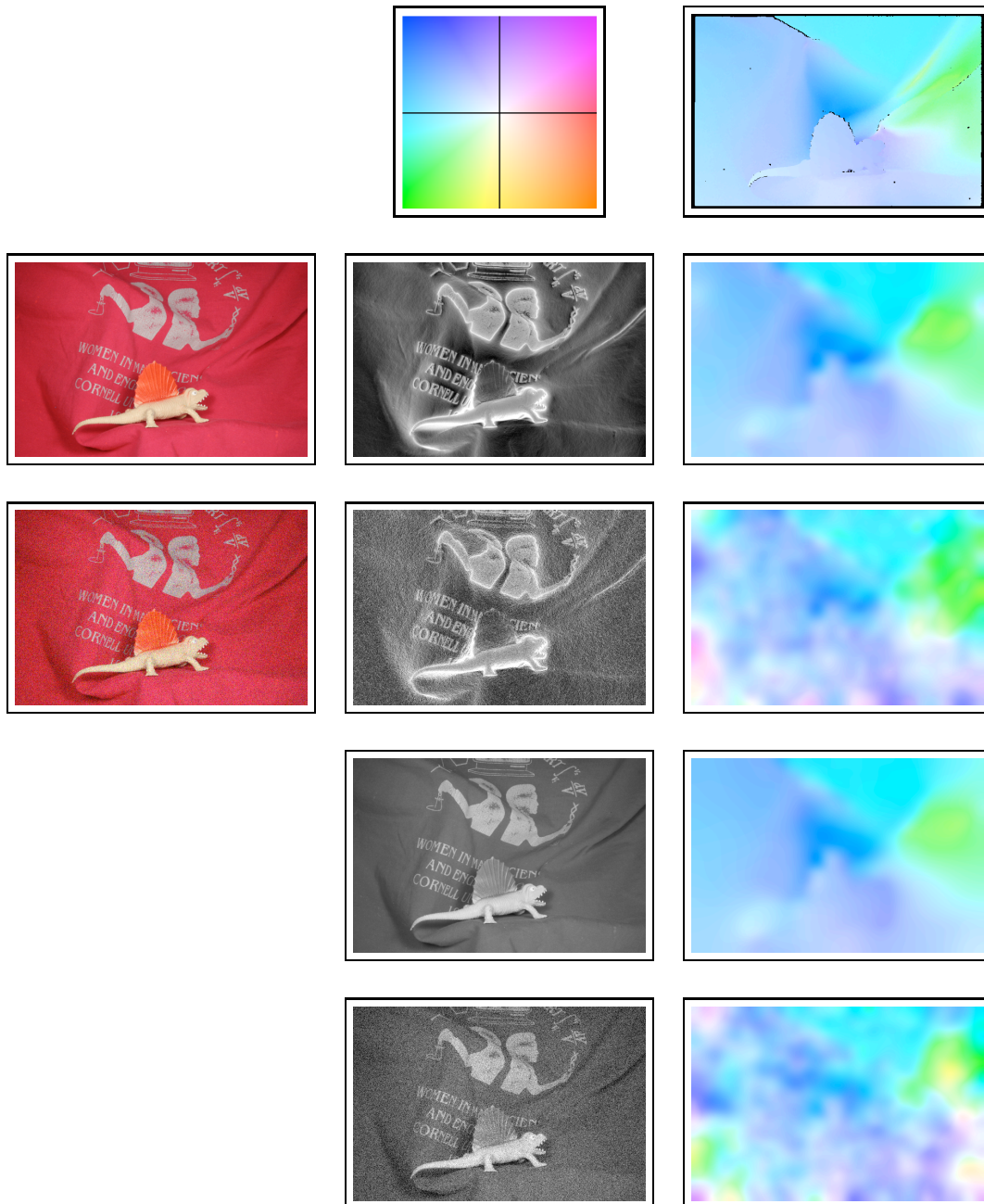


FIGURE 3.1 – De gauche à droite, première ligne : codage couleur du flot optique, vérité-terrain. Deuxième ligne : image originale couleur, φ_V , CLG couleur. Troisième ligne : image couleur bruitée, φ_V , CLG couleur. Quatrième ligne : image originale en niveau de gris, CLG. Cinquième ligne : image en niveau de gris bruitée, CLG.

où la fonction g est la suivante :

$$g(s) = (\tan(s))^{10}$$

Les équations d'Euler-Lagrange sont :

$$\begin{aligned} \Delta u - \frac{1}{\alpha} ((K_\rho * (\varphi_x^2) + \beta g(\varphi)) u + K_\rho * (\varphi_x \varphi_y) v \\ + K_\rho * (\varphi_x \varphi_t)) &= 0 \\ \Delta v - \frac{1}{\alpha} (K_\rho * (\varphi_x \varphi_y) u + (K_\rho * (\varphi_y^2) + \beta g(\varphi)) v \\ + K_\rho * (\varphi_y \varphi_t)) &= 0 \end{aligned}$$

Comme nous pouvons le voir, cela amène une légère modification de l'algorithme original dans [Bruhn 2005] puisque nous ajoutons seulement le terme $\beta g(\varphi)$ dans les équations de résolution.

3.1.6 Résultats

Il est important de noter que l'approche proposée diffère des approches usuelles. Nous ne cherchons pas un champs de vecteurs dense qui serait meilleur que les flots optiques déjà existants. L'idée dans cette application est de montrer comment utiliser l'information couleur pour obtenir les vecteurs mouvements d'objets d'une certaine couleur. Nous pouvons également souligner le fait que notre méthode est très robuste au changement d'illuminant dû au choix de φ comme intégrale première et au bruit dû à l'aspect multi-échelle du signal monogène couleur.

Deux exemples sont traités. Le premier concerne une vidéo de plantes dans un aquarium (voir la figure 3.2 et [web a]). Le mouvement des plantes est dû au mouvement de l'eau qui crée également de forts changements d'illuminant. Le but ici est d'extraire le flot optique des feuilles rouges.

Le deuxième exemple est une vidéo d'un trafic routier en Inde (voir la figure 3.3 et [web b]). Cette dernière contient d'importants bruits dûs à la compression JPEG et à une basse résolution. Nous sommes intéressés par l'extraction du champs de vecteurs associé aux mouvements des taxis jaunes. Dans les deux cas, la caméra n'est pas fixe et introduit une perturbation dans le mouvement. Pour une meilleure visualisation, l'orientation des vecteurs de flot optique sera codée par la teinte et sa norme par la saturation dans le système colorimétrique Hue, Saturation Value (HSV) (voir la figure 3.2).

La figure 3.2 montre les résultats pour le premier exemple. Il est évident que la méthode de Horn/Schunk (deuxième ligne, gauche) donne de mauvais résultats dûs aux forts changements d'illumination. Le mouvement le plus significatif est celui de l'ombre sur le sable. Le champs de vecteurs de la méthode CLG (deuxième ligne, milieu) est meilleur puisqu'il prend en compte plus de mouvements. Il est à noter qu'il est très difficile de retrouver les vecteurs correspondant aux feuilles. Le

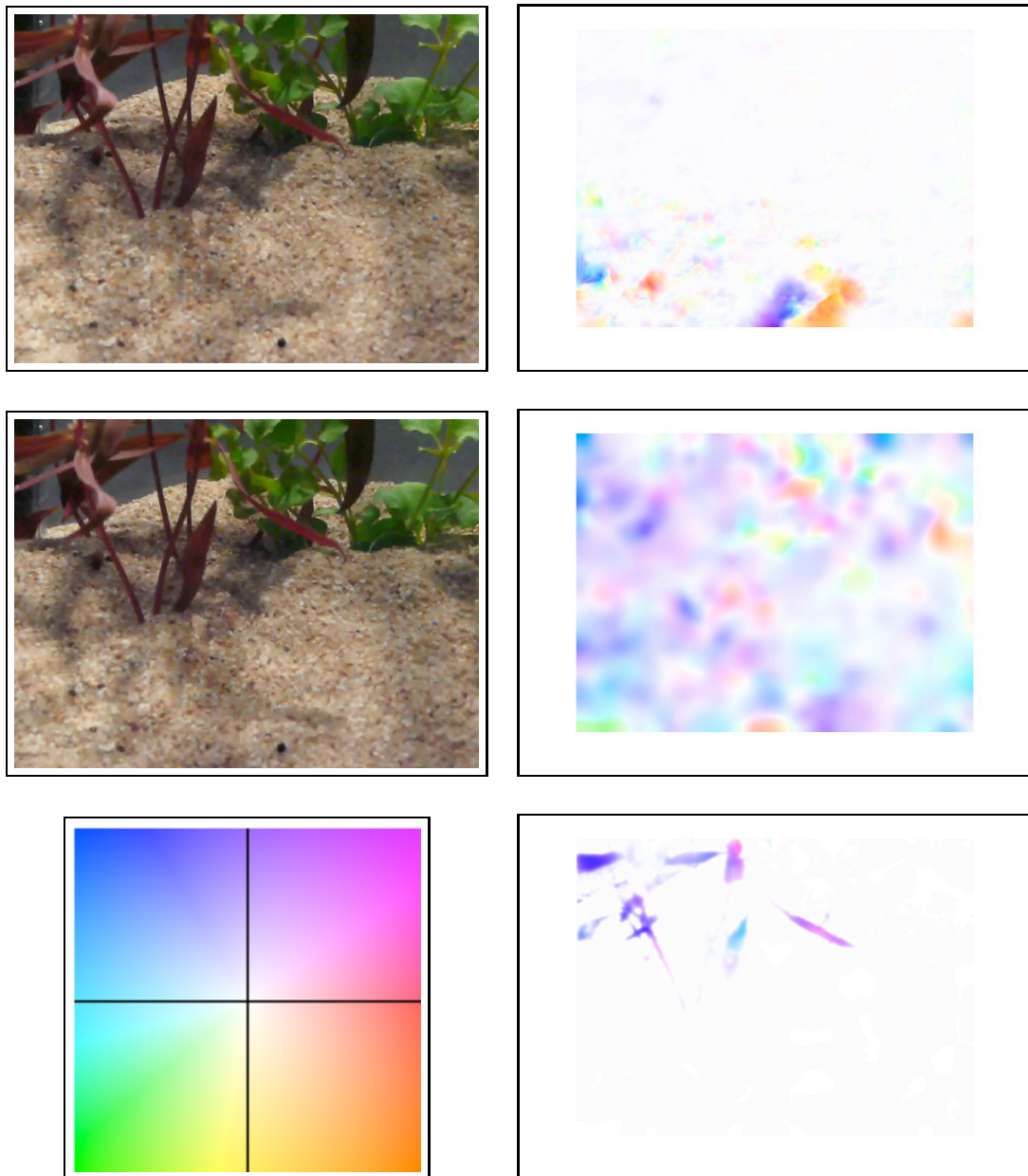


FIGURE 3.2 – De haut en bas. Première colonne : deux images de la séquence aquarium et codage couleur. Deuxième colonne : résultat pour les méthodes Horn/Schunk, CLG et notre méthode avec comme couleur choisie le rouge.

flot optique couleur obtenu avec la représentation polaire (deuxième ligne, droite) apparaît très pertinente et robuste aux changements d'illuminant. Le paramètre d'échelle utilisé est 1.

Pour le deuxième exemple (voir la figure 3.3), le paramètre d'échelle est aussi 1. Le champ de vecteurs correspondant aux taxis apparaît clairement et se révèle être robuste aux bruits.



FIGURE 3.3 – De gauche à droite. Image de la séquence des taxis, notre méthode avec comme couleur choisie le jaune.

3.2 Segmentation multi-échelle d'images couleur

Il existe de nombreuses méthodes de segmentation. Dans ce qui suit, nous nous intéressons à la détection des contours dans une image couleur correspondant à de fortes variations de l'image dans une couleur fixée. Ces variations sont mesurées à l'aide de la représentation polaire associée au signal monogène couleur. Notre approche est résolument géométrique; elle consiste à mesurer les variations métriques tout d'abord d'une surface, puis d'une variété de dimension 3, immergées respectivement dans \mathbb{R}^4 et \mathbb{R}^5 . Nous profitons ici du caractère multi-échelle du signal monogène couleur : dans un premier temps, la détection est à échelle fixée et dans un second "au travers de l'espace échelle".

3.2.1 Détection de contours métrique

Une méthode classique pour faire de la détection de contours est d'utiliser l'information de métrique donnée par la première forme fondamentale [DiZeno 1986] d'une image vue comme une surface. Pour cela, nous considérons une image multi-dimensionnelle définie sur un domaine D de \mathbb{R}^2 avec pour composantes (f_1, \dots, f_n) comme une surface S de dimension 2 paramétrisée par :

$$\psi : (x, y) \rightarrow (x, y, f_1(x, y), \dots, f_n(x, y))$$

plongée dans \mathbb{R}^{n+2} muni de la métrique :

$$g = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

La métrique de l'espace ambiant induit une métrique sur S , appelée première forme fondamentale de S , qui prend la forme suivante :

$$dS^2 = dx^2 + dy^2 + \lambda_1 df_1^2 + \dots + \lambda_n df_n^2$$

Les variations de l'image sont assimilées aux vecteurs tangents de S et une mesure de ces variations est donnée par dS^2 . Le reste de la méthode consiste en la sélection des plus grandes variations locales appelées contours. Plus précisément, soit $I(q)$ la matrice de représentation de la métrique dS^2 au point $q = \psi(p)$ ($p \in \mathbb{R}^2$) dans le système de coordonnées donné par les dérivées partielles $(d_p\psi(1,0), d_p\psi(0,1))$. Alors $I(q)$ a la forme suivante :

$$I(q) = \begin{pmatrix} 1 + \sum_{i=1}^n \left[\frac{\partial f_i}{\partial x}(q) \right]^2 & \sum_{i=1}^n \frac{\partial f_i}{\partial x} \frac{\partial f_i}{\partial y}(q) \\ \sum_{i=1}^n \frac{\partial f_i}{\partial x} \frac{\partial f_i}{\partial y}(q) & 1 + \sum_{i=1}^n \left[\frac{\partial f_i}{\partial y}(q) \right]^2 \end{pmatrix} \quad (3.1)$$

Soient $\lambda_+(q)$ et $\lambda_-(q)$, $\lambda_+ \geq \lambda_-$ deux valeurs propres de $I(q)$ et $\theta_+(q)$, $\theta_-(q)$ les vecteurs propres correspondants. La force de contour est alors donnée par :

$$w(q) = \sqrt{\lambda_+(q) - \lambda_-(q)} \quad (3.2)$$

et nous dirons que $p \in D$ est un point de contour si la fonction w possède un maximum local au point $\psi(p)$ dans la direction donnée par $\theta_+(\psi(p))$ [Cumani 1991].

3.2.2 Méthode de segmentation

Tout d'abord, nous avons besoin de choisir le système colorimétrique que nous souhaitons utiliser pour notre signal monogène couleur. Dans cette application, nous choisissons l'espace CIE $L^*a^*b^*$ et nous considérons chaque pixel d'une image couleur comme un vecteur dans $\mathbb{R}_{5,0}$ à l'aide du signal monogène couleur à l'échelle s (la même pour chaque composante, *i.e.* $s = x_3 = x_4 = x_5$ dans la formule 8). Alors le calcul géométrique nous autorise à considérer le produit géométrique entre chaque vecteur et un vecteur de référence $V = ue_1 + ve_2 + ae_3 + be_4 + ce_5$. Comme dans le paragraphe 2.4.3, une fonction d'angle $\varphi_V(x, y)$ est obtenue et associée à la partie scalaire $p(x, y) = \langle f_c V(x, y) \rangle_0$ donne une information d'angle et d'amplitude respectivement.

Schéma algorithmique

1. Soit $V = ue_1 + ve_2 + ae_3 + be_4 + ce_5$.
2. Calcul de $\varphi(x, y)$ et $p(x, y)$ comme dans 2.4.3.
3. Etant donnée S paramétrée par $\psi : (x, y) \rightarrow (x, y, \varphi(x, y), p(x, y))$ dans \mathbb{R}^4 muni de la métrique $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$:
 - Calcul de la matrice de représentation de la métrique dS^2 comme dans 3.1 ;
 - Calcul de $w(q)$ comme dans 3.2.

Enfin, les variations à la fois dans la fonction d'angle et dans la partie scalaire sont assimilées aux vecteurs tangents de S et nous pouvons mesurer les contours relativement à une couleur choisie V .

3.2.3 Résultats

Nous présentons des résultats obtenus sur des images provenant de la base de données de Berkeley [Martin 2001]. Nous souhaitons montrer la pertinence de notre approche sur ce type d'images difficiles. En premier lieu, nous utilisons l'image des jouets [Gevers 2006] (figure 3.4) et nous souhaitons une détection des contours des objets jaunes.

En prenant le vecteur V codant le jaune dans l'espace $L^*a^*b^*$ sans aucune structure *i.e.* $V = ae_3 + be_4 + ce_5$, nous appliquons la méthode décrite ci-dessus avec la métrique euclidienne ($\lambda_1 = 1, \lambda_2 = 1$) et obtenons les images de la figure 3.4. Comme nous pouvons le constater, quand nous prenons le paramètre d'échelle égal à 0, nous n'obtenons pas que les objets jaunes mais tous les objets dont la variation en jaune est forte. Quand le paramètre d'échelle augmente, nous avons dans un premier temps un lissage des contours et ensuite la disparition de blocs bleus. Finalement nous gardons que les objets jaunes recherchés mais aussi les objets verts clairs à cause de la forte variation dans le jaune entre l'objet et le fond.

Par la suite, nous choisissons deux images de la base de données de Berkeley : l'avion et le jardin (voir la figure 3.5). Pour la première image (première ligne), nous voulons les contours rouges. En utilisant l'aspect multi-échelle et en prenant la métrique ($\lambda_1 = 2, \lambda_2 = 0.5$) pour renforcer la fonction d'angle relativement à l'amplitude, nous obtenons les résultats visibles dans la figure 3.5. Pour la seconde image (deuxième ligne), nous voulons les contours verts. En utilisant l'aspect multi-échelle et en prenant la métrique euclidienne, les résultats de la figure 3.5 montre une segmentation très intéressante comparée à celles obtenues dans [Martin 2001].

3.2.4 Détection de contours métrique et multi-échelle

Cette partie présente une nouvelle approche de segmentation multi-échelle d'images couleur. Basée sur le travail précédent [Demarcq 2009b], le signal mono-

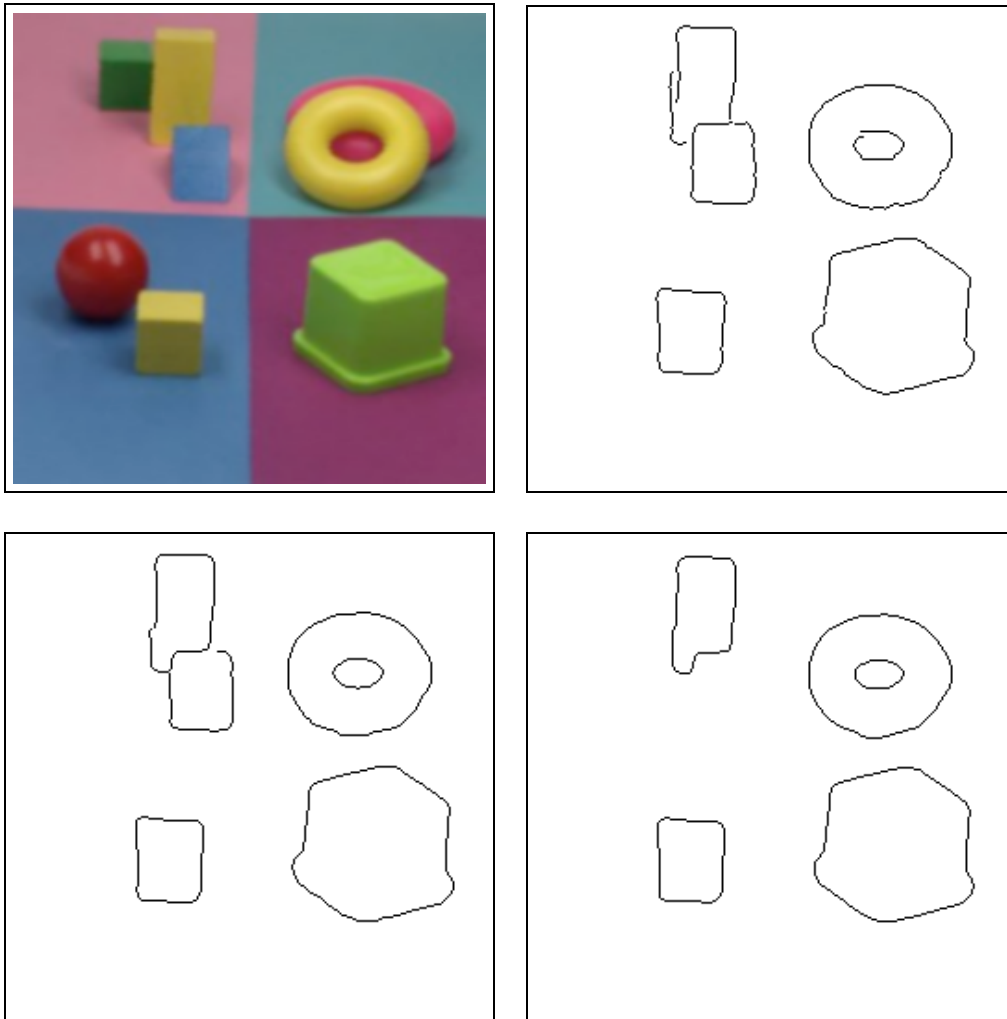


FIGURE 3.4 – De gauche à droite : image des jouets, segmentation aux échelles 0, 1.5 et 3.5 (le seuillage par hystéresis est le même pour chaque échelle).

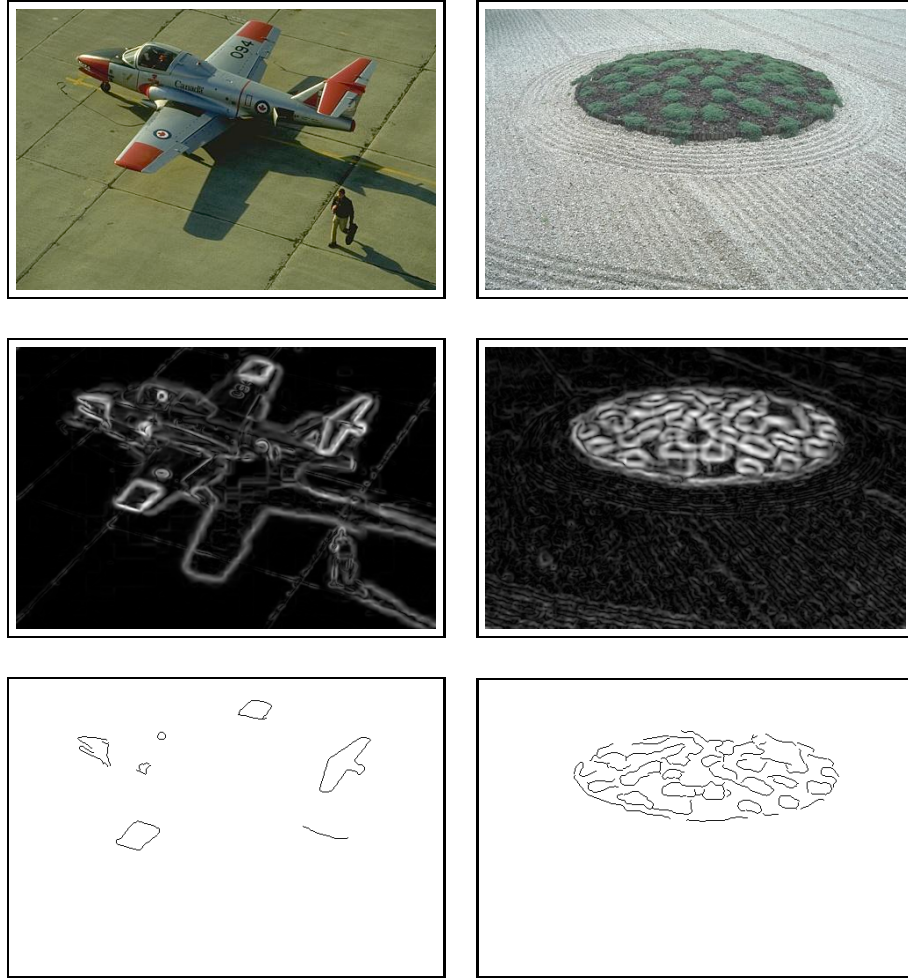


FIGURE 3.5 – De haut en bas. Première colonne, image de l'avion, la fonction w et le résultat à l'échelle 3.5. Deuxième colonne, l'image du jardin, la fonction w et le résultat à l'échelle 3.5.

gène couleur ainsi que de la géométrie différentielle sont utilisés pour modéliser les variations dans une image couleur. Le principal avantage de cette approche comparé à la précédente est qu'elle est capable de faire une segmentation automatique dans l'espace-échelle, *i.e.* sans avoir à choisir une échelle particulière.

La méthode précédente peut être étendue à une famille de surfaces paramétrées par un paramètre réel τ . Nous considérons la variété de dimension 3 plongée par

$$\psi : (x, y, \tau) \rightarrow (x, y, \tau, f_1(x, y, \tau), \dots, f_n(x, y, \tau))$$

dans \mathbb{R}^{n+3} muni de la métrique

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Dans ce cas, la première forme fondamentale est :

$$dS^2 = dx^2 + dy^2 + d\tau^2 + \lambda_1 df_1^2 + \dots + \lambda_n df_n^2$$

La matrice de représentation $I(q)$ de la métrique dS^2 en $q = \psi(p)$ ($p \in \mathbb{R}^3$) est de la forme suivante :

$$I(q) = \begin{pmatrix} 1 + \sum_{i=1}^n \left[\frac{\partial f_i}{\partial x}(q) \right]^2 & \sum_{i=1}^n \frac{\partial f_i}{\partial x} \frac{\partial f_i}{\partial y}(q) & \sum_{i=1}^n \frac{\partial f_i}{\partial x} \frac{\partial f_i}{\partial \tau}(q) \\ \sum_{i=1}^n \frac{\partial f_i}{\partial y} \frac{\partial f_i}{\partial x}(q) & 1 + \sum_{i=1}^n \left[\frac{\partial f_i}{\partial y}(q) \right]^2 & \sum_{i=1}^n \frac{\partial f_i}{\partial y} \frac{\partial f_i}{\partial \tau}(q) \\ \sum_{i=1}^n \frac{\partial f_i}{\partial \tau} \frac{\partial f_i}{\partial x}(q) & \sum_{i=1}^n \frac{\partial f_i}{\partial \tau} \frac{\partial f_i}{\partial y}(q) & 1 + \sum_{i=1}^n \left[\frac{\partial f_i}{\partial \tau}(q) \right]^2 \end{pmatrix} \quad (3.3)$$

Soient $\lambda_1(q)$, $\lambda_2(q)$ et $\lambda_3(q)$, $\lambda_1(q) \leq \lambda_2(q) \leq \lambda_3(q)$, les trois valeurs propres de $I(q)$ et $\theta_1(q)$, $\theta_2(q)$, $\theta_3(q)$ les vecteurs propres correspondants. La force de contour considérée est donnée par :

$$w(q) = \sum_{i=1}^3 \lambda_i \quad (3.4)$$

et nous dirons que p est un contour si la fonction w a un maximum local en $\psi(p)$ dans la direction donnée par $\theta_3(\psi(p))$. Notons que la fonction w est un peu différente de celle utilisée dans le cas 2D puisque nous utilisons trois valeurs propres. Ce type de fonction est aussi utilisé dans [Sapiro 1997][Tschumperlé 2006] pour faire de la diffusion anisotropique.

L'idée est d'appliquer ce type de segmentation quand τ est un paramètre d'échelle. Cela permet une détection de contours dans l'image, ces derniers étant aussi, d'une certaine manière, des contours dans l'espace-échelle, *i.e.* les contours les plus pertinents au travers des échelles.

3.2.5 Méthode de segmentation

Nous considérons une image couleur dans le système colorimétrique CIE L*a*b*. Soit V un vecteur de $\mathbb{R}_{5,0}$ et $\chi(V) = f_c V$ le verneur polaire correspondant. De ces données, nous introduisons la variété de dimension trois plongée par

$$\psi : (x, y, \tau) \rightarrow (x, y, \tau, \varphi_V(x, y, \tau), p_V(x, y, \tau))$$

dans \mathbb{R}^5 muni de la métrique

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

(p_V est la partie scalaire de $\chi(V)$). Les variations de la famille de surfaces définie par p_V et φ_V et paramétrée par le paramètre d'échelle τ , sont assimilée aux vecteurs tangents et une mesure de ces variations est donnée par $dS^2 = dx^2 + dy^2 + d\tau^2 + \lambda d\varphi^2 + \mu dp^2$. Ces dernières sont les variations dans l'espace échelle de Poisson. En appliquant la méthode définie dans le paragraphe 3.2.4, nous obtenons une segmentation métrique et multi-échelle.

Remarques

- La détection de contours dépend du vecteur V , *i.e.* d'une couleur choisie. Ceci implique que les contours d'une couleur choisie sont détectés.
- Pour simplifier, nous travaillons avec seulement un paramètre d'échelle en posant : $x_3 = x_4 = x_5 = \tau$ dans la construction du signal monogène couleur. Le cas général peut être traité théoriquement de la même manière en considérant la variété de dimension cinq correspondante. Néanmoins, les interprétations paraissent moins évidentes.
- Nous considérons la variété donnée par les composantes p_V et φ_V du verneur polaire $\chi(V)$ au lieu de travailler directement avec le signal monogène couleur. Il apparait en fait que ce verneur contient toute l'information nécessaire pour réaliser la segmentation vis à vis d'une couleur choisie.
- Les paramètres λ et μ contrôlent l'influence de φ_V et p_V dans la détection.

3.2.6 Segmentation multi-échelle

L'algorithme de segmentation que nous proposons ici est une extension du classique détecteur de Canny [Canny 1986], défini pour les images en niveau de gris, donc bidimensionnelles, à un espace d'échelle tridimensionnel. Cette méthode est constituée de quatre étapes successives :

- Lissage gaussien de l'image pour réduire le bruit
- Calcul du vecteur gradient en chaque pixel
- Suppression des non-maxima locaux
- Seuillage par hystéresis

La suppression des non-maxima locaux est un point clé de cette approche, et son extension à l'espace d'échelle peut être faite d'au moins deux façons. Ces points particuliers peuvent être caractérisés à partir des dérivées d'ordre deux et trois [Lindeberg 1998] ou plus directement en terme de vecteur gradient [Monga 1990] comme dans l'algorithme original de Canny. Le principe est ici de se déplacer dans la direction de la normale au contour, telle que donnée par le vecteur gradient, et de retenir comme points contours potentiels ceux dont la norme est localement la plus grande. L'espace d'échelle étant ici représenté par un

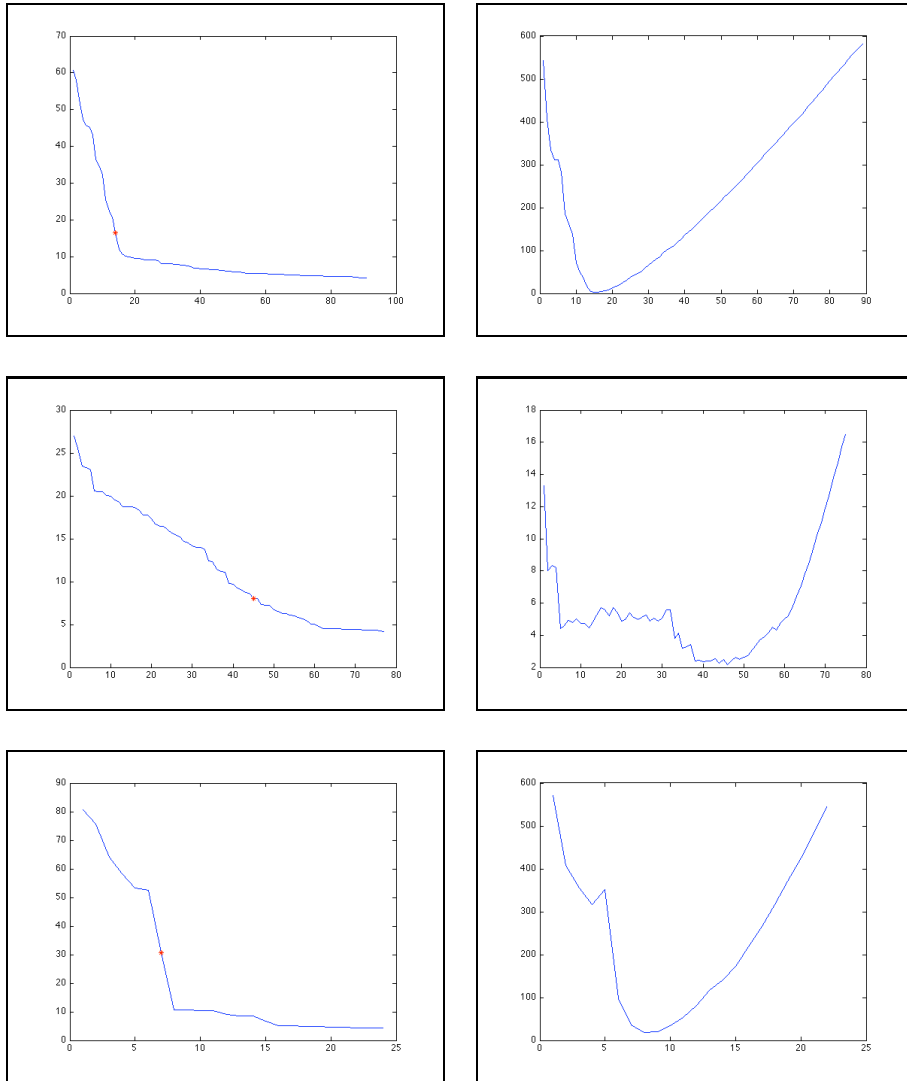


FIGURE 3.6 – De haut en bas. Première colonne : courbe des composantes connexes triées par ordre décroissant en fonction de w pour les images d'avion, de jardin et des jouets respectivement. Deuxième colonne : Fonctions d'erreur correspondantes.

empilement d'images bidimensionnelles, chaque pixel possède 26 voisins. En pratique, pour un pixel donné, la norme du vecteur gradient en ce point est comparée à celle des deux voisins dans la direction de ce même vecteur gradient. Il faut donc considérer un voisin "avant" le pixel courant et un voisin "après". La norme de ces voisins est obtenue par combinaison trilineaire des 8 pixels qui définissent le cube 3D dans lequel il se situe [Monga 1990]. La transposition dans notre cadre est directe, en substituant $\theta_3(\psi(p))$ à la direction du vecteur gradient et l'intensité du contour $w(p)$ à la norme du gradient.

À l'issue de ce processus de suppression des non-maxima locaux nous obtenons une liste de composantes connexes parmi lesquelles les contours pertinents sont sélectionnés. Classiquement, ceci est réalisé par seuillage par hystérésis dont les paramètres sont, en général, délicats à ajuster. Nous proposons de simplifier cette étape en exploitant l'aspect multi-échelle de notre signal. Pour cela, les N composantes connexes sont triées par ordre décroissant de la moyenne de l'intensité de contour ω à travers les échelles. Les plus pertinentes sont alors les n premières. Le graphe de cette mesure est naturellement décroissant (voir la figure 3.6 première ligne), et nous pouvons visuellement vérifier qu'il est constitué de deux parties quasi-linéaires reliées par un coude. Nous proposons que ce point remarquable soit utilisé pour définir le meilleur compromis entre le nombre de composantes connexes et la qualité de la segmentation. Sa détection est faite très simplement en utilisant un modèle de courbe défini par deux segments de droites. L'origine du premier segment est le premier point de la courbe, le coude est à la fois son extrémité et l'origine du second segment dont l'extrémité est le dernier point de la courbe. La qualité de l'ajustement du modèle à la courbe est mesuré par le calcul de l'erreur quadratique moyenne (EQM). Le coude optimum est déterminé en sélectionnant celui qui, parmi les N candidats, donne l'EQM la plus faible (voir Figure 3.6 seconde ligne).

L'analyse de la complexité algorithmique (temps et espace mémoire) de la segmentation est donnée dans l'Appendice E.

3.2.7 Résultats

Notre méthode de segmentation est testée sur les trois images utilisées précédemment. Les premiers résultats concernent l'approche décrite au paragraphe 3.2.2 (voir également [Demarcq 2009b]). La surface 2D est donnée par :

$$\psi : (x, y) \rightarrow (x, y, \varphi_V(x, y), p_V(x, y))$$

où $\chi(V)$ est implémenté à une échelle choisie. La figure 3.4 montre les contours détectés dans l'image de jouets pour différentes échelles avec V , vecteur de $\mathbb{R}_{5,0}$, codant la couleur jaune sans aucune structure ($u = v = 0$). Il faut noter que le seuil sur w est indépendant de l'échelle mais a besoin d'être choisie. Ici, les paramètres λ et μ de la métrique ambiante sont égaux à 1. Les contours à l'échelle 0 correspondent aux fortes variations de la couleur jaune dans l'image. Quand le paramètre d'échelle augmente, les contours sont lissés (échelle 1.5) et le cube bleu disparaît (échelle 3.5) puisque les variations correspondantes ne sont plus assez significatives. Les fortes

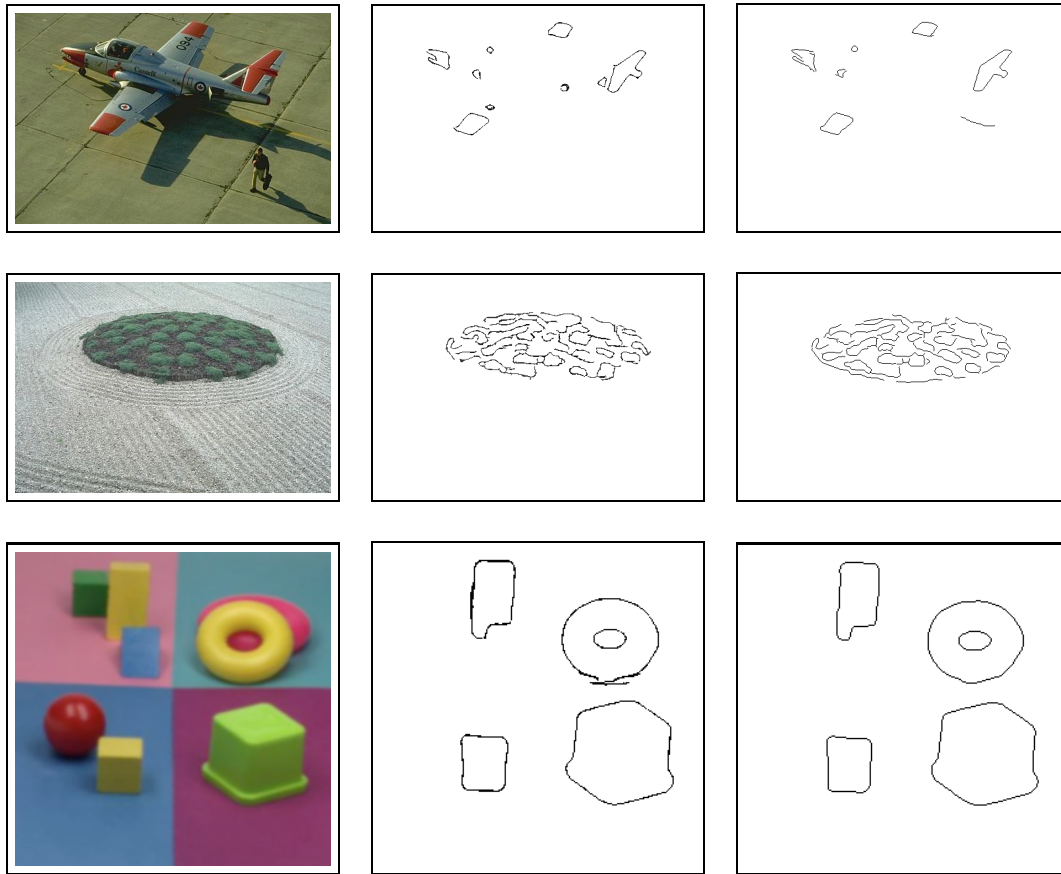


FIGURE 3.7 – Première colonne : images originales. Deuxième colonne : résultats utilisant la segmentation multi-échelle. Troisième colonne : résultats utilisant une échelle fixée (échelle=3.5).

variations en jaune entre le cube vert clair et le fond sont toujours pris en compte. La première (resp. deuxième) ligne de la figure 3.5 montre les niveaux de la fonction w pour V codant la couleur rouge (resp. verte) et la détection de contours à l'échelle 3.5 avec λ égale 2 (resp. 1) et μ égale 0.5 (resp. 1).

Maintenant, détaillons les résultats obtenus avec l'approche développée dans les paragraphes 3.2.4 à 3.2.6 utilisant la variété de dimension trois :

$$\psi : (x, y, \tau) \rightarrow (x, y, \tau, \varphi_V(x, y, \tau), p_V(x, y, \tau))$$

Les résultats de notre méthode de segmentation multi-échelle sont illustrés dans la figure 3.7. Nous avons procédé de la manière suivante, l'image d'avion est segmentée dans la couleur rouge, l'image de jardin dans le vert et l'image de jouets dans le jaune. La métrique est la même pour chaque image et égale à $\lambda = \mu = 0.1$, les paramètres d'échelle sont pris entre 1 et 5 avec un pas de 0.5. Ces contours sont obtenus en utilisant la méthode décrite au paragraphe 3.2.6 qui nous permet de sélectionner le nombre pertinent de composantes connexes (voir la figure 3.6). Les courbes associées

à l'image de jardin sont plus difficiles à interpréter que les autres ; cela est clairement dû à la grande complexité des contours verts dans l'image. Néanmoins chacune des trois courbes de la deuxième ligne possède un minimum global ce qui donne les contours pertinents dans la segmentation multi-échelle. Les contours obtenus avec cette méthode automatique et multi-échelle donne des résultats intéressants comparés à la méthode avec échelle fixe décrite dans [Demarcq 2009b]. En effet, les contours de la troisième ligne correspondent à la segmentation à l'échelle 3.5. Celles de la deuxième ligne sont ceux obtenus avec la nouvelle méthode. Nous pouvons voir que le résultat de l'image de jouets est cohérent avec le choix de l'échelle 3.5 (le cube bleu disparaît). Les résultats pour les images de jardin et d'avion apparaissent plus pertinents que pour notre méthode à échelle fixe.

3.3 Suivi d'objets couleur

Le suivi d'un objet en mouvement dans une séquence d'images, également appelé poursuite de cible, consiste à estimer l'état de cet objet (ou cible) sur chaque image de la séquence à partir d'observations. Les applications en vision par ordinateur en sont nombreuses et les approches très diverses. Sans prétendre être exhaustif, nous citerons la surveillance intelligente de parking ou de magasins [Dockstader 2001, KaewTrakulPong 2003, Siebel 2003], la surveillance de réseau routier [Magee 2004, Rittscher 2000, Pai 2004], l'analyse de mouvements humains pour animer des "avatars" synthétiques [Moeslund 2001] et pour finir, la robotique mobile qui a été dans les années passées l'application reine [Koller-Meier 2001, Schulz 2001]. Les premiers algorithmes étaient basés sur le filtre de Kalman [Harris 1993] et plus récemment, la technique du Mean-Shift [Comaniciu 2000] ou le WSL [Jepson 2003] sont devenus des approches intéressantes. Mais ces dix dernières années, le filtrage particulière [Cappé 2007] est devenu l'outil le plus populaires en suivi d'objets dans les vidéos. Ces filtres, comme les filtres de Kalman, estiment la densité de probabilité *a posteriori*, c'est-à-dire la loi du processus d'état conditionnellement aux mesures acquises depuis l'instant initial jusqu'à l'instant présent. Le filtre de Kalman s'applique dans le cas où les évolutions de l'état et des mesures sont régies par des modèles linéaires et bruits additifs Gaussiens. Les méthodes de filtrage particulières ont été développées afin d'estimer des distributions de probabilité *a posteriori* multi-modales telles que celles couramment rencontrées dans les applications de suivi.

Dans ce contexte, nous définissons un nouveau critère de similarité à partir du signal monogène couleur. Dans un premier temps, nous l'utilisons dans un algorithme naïf de suivi de type block-matching, et dans un deuxième temps, dans un filtre particulière simple dit "bootstrap". La méthode est testée sur des images couleur et en niveau de gris.

3.3.1 Un algorithme naïf de block-matching

Notre méthode de block-matching repose sur l'idée principale qui consiste à voir une image couleur non pas comme un ensemble de pixels mais comme un champs

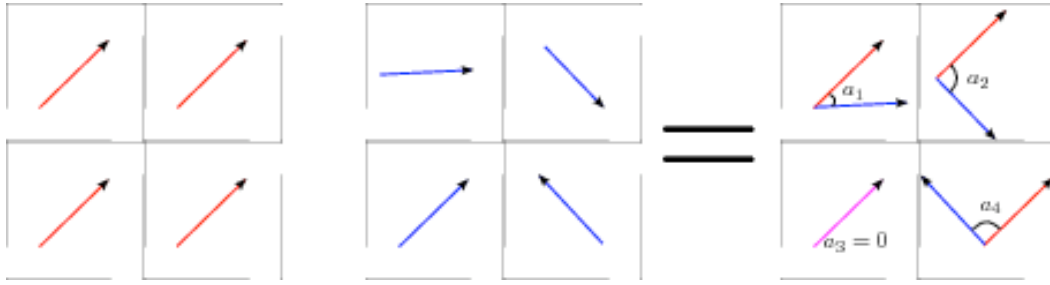


FIGURE 3.8 – Produit géométrique entre deux blocs de 4 pixels vus comme des vecteurs dans $\mathbb{R}_{5,0}$. Les angles sont représentés par a_1 , a_2 , a_3 et a_4 .

de vecteurs dans l’algèbre de Clifford $\mathbb{R}_{5,0}$. De par cette représentation, chacun des vecteurs contient à la fois une information de couleur et de structure. Dans ce contexte, nous définissons un objet par un bloc de taille $N \times M$. En faisant l’hypothèse classique que l’apparence d’un objet ne change pas significativement d’une image à l’autre, nous prenons le bloc définissant l’objet et nous recherchons le plus similaire dans l’image suivante en utilisant le critère suivant :

$$\bar{\varphi} = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M \varphi_{ij} \quad (3.5)$$

où φ_{ij} est la fonction d’angle (voir l’équation 2.10) entre chaque vecteur des deux blocs (voir la figure 3.8). Afin d’en montrer la pertinence, nous allons utiliser un algorithme naïf qui consiste en l’application de l’approche de block-matching décrite précédemment dans un voisinage de 30×30 pixels. Pour chaque pixel de l’image courante, nous extrayons le bloc correspondant et calculons le critère de similarité. La plus petite valeur est choisie comme position de la cible à suivre. En appliquant cette méthode itérativement et en mettant à jour le modèle de l’objet à chaque image, nous obtenons de bons résultats (voir la figure 3.9). Évidemment, cet algorithme est simple et possède un grand coût de calcul mais cela démontre que notre critère est adapté comme mesure de similarité.

Nous choisissons comme séquence test, une vidéo provenant du projet aqu@thèque [Bichot 2005] qui est un système de reconnaissance automatique de poissons. Elle contient des difficultés telles que des changements de luminosité, des occlusions et de fortes textures. Dans la figure 3.9, nous initialisons le processus de suivi en choisissant manuellement un bloc autour de l’œil du poisson bleu et nous appliquons l’algorithme précédent. Nous pouvons constater que la cible est maintenue jusqu’à la fin malgré un changement de direction brutal du poisson dans le milieu de la séquence ainsi qu’un fort changement de luminosité. L’efficacité de notre critère pour suivre les structures et la couleur dans un environnement complexe est illustrée dans cette séquence.

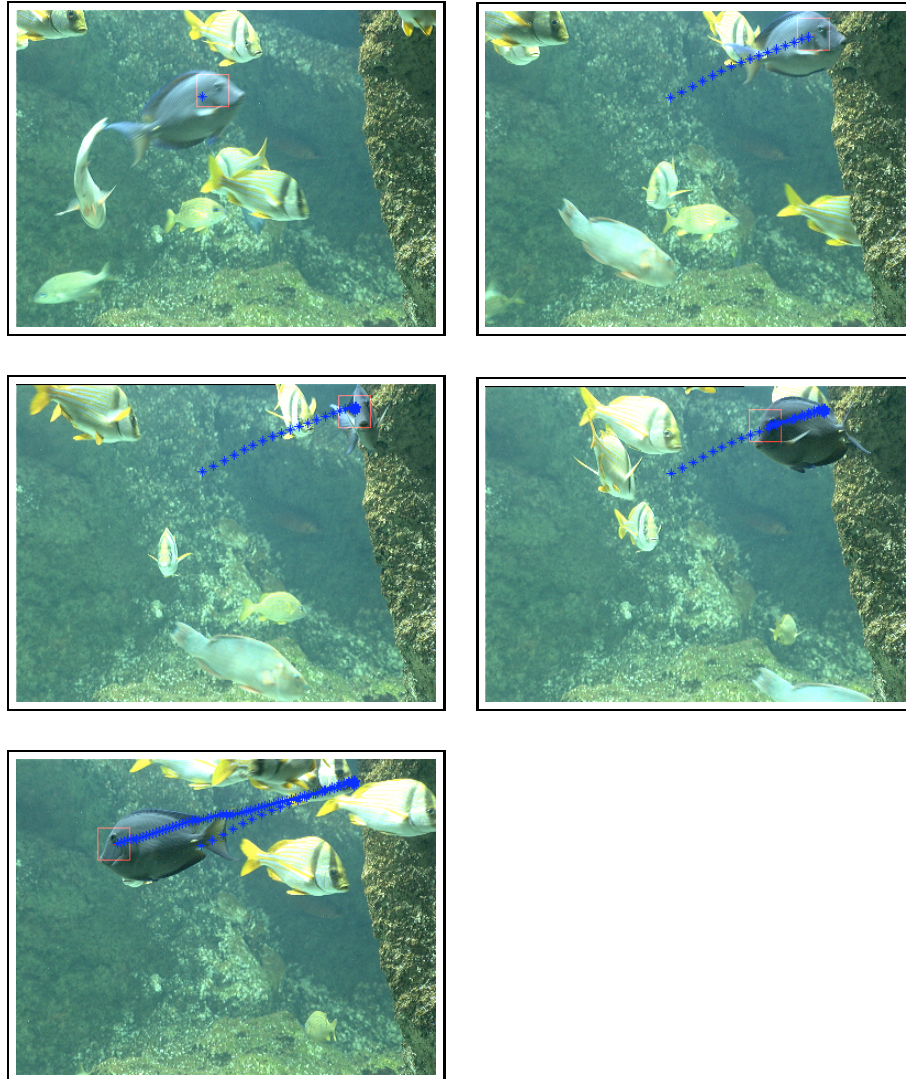


FIGURE 3.9 – Résultats du suivi pour les images #1, #16, #31, #51 et #101 de la séquence Aqu@thèque. Les étoiles bleues représentent la trajectoire du poisson suivi.

3.3.2 Filtrage particulière

Dans le cadre de notre étude, nous nous intéressons au suivi d'une seule¹ cible par les méthodes de filtrage particulière. L'état \mathbf{x}_t de l'objet² regroupe un ensemble de caractéristiques descriptives comme par exemple sa position, sa vitesse ou encore sa forme. Son estimation s'effectue à partir d'observations ou mesures issues de l'image. La poursuite de cible est donc un problème d'inférence qui peut se formaliser comme suit :

- nous nous intéressons à l'objet caractérisé par le vecteur d'état inconnu \mathbf{x}_t à l'instant t ;
- nous disposons d'un ensemble d'observations bruitées mesurées jusqu'à l'instant t et notées $\mathbf{z}_{1:t} = \{\mathbf{z}_1, \dots, \mathbf{z}_t\}$;
- l'objectif du suivi est de déterminer \mathbf{x}_t conditionnellement à $\mathbf{z}_{1:t}$.

Soit (Ω, P) un espace probabilisé. Considérons le système composé des deux processus d'état $(\mathbf{x}_t)_{t \in N}$ et de mesures $(\mathbf{z}_t)_{t \in N^*}$ à valeurs respectivement dans R^{n_x} et R^{n_z} . Le processus d'état $(\mathbf{x}_t)_{t \in N}$ est inconnu et supposé Markovien. Le processus de mesures $(\mathbf{z}_t)_{t \in N^*}$ est également appelé processus d'observations.

La dynamique du processus d'état est régie par l'équation d'évolution

$$\mathbf{x}_t = F_t(\mathbf{x}_{t-1}, \mathbf{v}_t) \quad (3.6)$$

où F_t désigne une fonction, linéaire ou non, décrivant l'évolution de l'état et V_t est un bruit blanc de dynamique, Gaussien ou non. Le processus de mesures $(\mathbf{z}_t)_{t \in N^*}$ est lié au processus d'état par l'équation

$$\mathbf{z}_t = H_t(\mathbf{x}_t, \mathbf{w}_t)$$

où H_t désigne une fonction d'observations, linéaire ou non, et \mathbf{w}_t est un bruit blanc de mesures, indépendant de V_t , Gaussien ou non. Comme les bruits de dynamique \mathbf{v}_t et de mesures \mathbf{w}_t sont blancs et indépendants, le processus d'état est Markovien et les observations sont également indépendantes. Dans la suite, le temps est discrétisé et nous désignons par instant courant l'instant t . \mathbf{x}_0 désigne l'état initial de la cible obtenu manuellement ou automatiquement. La poursuite de cible consiste à estimer

1. Il existe, naturellement, des méthodes dédiées au suivi multi-cibles. Une difficulté supplémentaire du suivi multi-cibles par rapport au suivi d'un seul objet réside dans l'association des mesures ou observations extraites de l'image aux différents objets d'intérêt. La variation du nombre de cibles, liée à l'apparition et/ou disparition d'objets dans la scène, représente un autre problème. Les méthodes les plus fréquemment rencontrées sont "le filtre conjoint à associations de données probabilistes" ou *JPDAF* [Bar-Shalom 1988], le "filtre à hypothèses multiples" ou *MHT* [Reid 1979], et les extensions des filtres particuliers au suivi multi-cibles [Isard 2001, Hue 2003, Schulz 2001].

2. Nous utiliserons une notation de type "ingénieur", classique dans le domaine du filtrage, où la variable aléatoire est confondue avec sa réalisation, ceci n'entraînant pas ici de confusion tout en simplifiant l'écriture.

l'état du processus à l'instant courant à partir des observations dont nous disposons depuis l'instant initial jusqu'à l'instant courant. D'un point de vue formel, il s'agit d'estimer la densité de probabilité *a posteriori* $p(\mathbf{x}_t|\mathbf{z}_{1:t})$, c'est-à-dire la distribution de l'état \mathbf{x}_t du processus à l'instant t conditionnellement aux observations $\mathbf{z}_{1:t}$ mesurées depuis l'instant initial jusqu'à l'instant t . La densité de probabilité *a posteriori* $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ peut être obtenue récursivement en deux étapes, prédiction et correction, que nous allons maintenant décrire. Supposons connue la densité de probabilité *a posteriori* $p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1})$ à l'instant précédent. Dans un premier temps, l'étape de prédiction consiste à utiliser l'équation de dynamique 3.6 pour obtenir la densité de probabilité *prédite* $p(\mathbf{x}_t|\mathbf{z}_{1:t-1})$ à l'instant courant :

$$p(\mathbf{x}_t|\mathbf{z}_{1:t-1}) = \int_{R^{n_x}} p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1})d\mathbf{x}_{t-1} \quad (3.7)$$

À l'instant courant, une nouvelle mesure \mathbf{z}_t devient disponible. Cette mesure permet, dans un second temps, de corriger la densité de probabilité *prédite* $p(\mathbf{x}_t|\mathbf{z}_{1:t-1})$ selon la règle de Bayes afin d'obtenir la densité *a posteriori* $p(\mathbf{x}_t|\mathbf{z}_{1:t})$:

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \frac{p(\mathbf{z}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{z}_{1:t-1})}{p(\mathbf{z}_t|\mathbf{z}_{1:t-1})} \quad (3.8)$$

Le filtre de Kalman fournit la solution exacte dans le cas où F_t et H_t sont linéaires et les bruits \mathbf{v}_t et \mathbf{w}_t Gaussiens. Cependant, les problèmes réels de poursuite de cible présentent des conditions difficiles comme, par exemple, la présence d'un milieu perturbé, qui ne permettent pas d'appliquer ces filtres. C'est notamment pour pallier ce problème que les méthodes de filtrage particulaire, encore appelées méthodes de Monte-Carlo séquentielles, ont été proposées en vision par ordinateur à la fin des années 90. Ces méthodes sont le sujet de nombreuses études [Isard 1998a, Arulampalam 2002, Hue 2003]. Elles approchent la distribution de probabilité *a posteriori* $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ par une somme pondérées de distributions de Dirac centrées en des points appelés *particules* :

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \sum_{n=1}^N \pi_t^n \delta(\mathbf{x}_t - \mathbf{x}_t^n)$$

où π_t^n représente le poids de la particule \mathbf{x}_t^n , $\delta(\cdot)$ désigne la distribution de Dirac et N le nombre de particules utilisées pour estimer la densité *a posteriori* $p(\mathbf{x}_t|\mathbf{z}_{1:t})$. A chaque itération, l'état moyen de la cible est estimé par :

$$\hat{\mathbf{x}}_t = \sum_{n=1}^N \pi_t^n \mathbf{x}_t^n$$

Les particules représentent des hypothèses sur l'état de la cible. Le poids associé à une particule est proportionnel à la probabilité de cette hypothèse. Les méthodes de filtrage particulaire consistent à propager la liste des particules pondérées au cours du temps de sorte que les particules explorent l'espace d'état de la cible et se

concentrent par un mécanisme de sélection dans les régions de forte vraisemblance de l'espace d'état [Arulampalam 2002]. Du point de vue probabiliste, cette propagation s'effectue grâce à une densité de probabilité appelée loi de proposition ou densité d'importance, et notée $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{z}_t)$. Le poids de chaque particule est mis à jour par la formule suivante :

$$\pi_t^n \propto \pi_{t-1}^n \frac{p(\mathbf{z}_t|\mathbf{x}_t^n)p(\mathbf{x}_t^n|\mathbf{x}_{t-1}^n)}{q(\mathbf{x}_t^n|\mathbf{x}_{t-1}^n, \mathbf{z}_t)} \quad (3.9)$$

où $p(\mathbf{z}_t|\mathbf{x}_t^n)$ est la vraisemblance de l'observation, et $p(\mathbf{x}_t^n|\mathbf{x}_{t-1}^n)$ la loi de dynamique qui correspond à la partie déterministe du mouvement des particules. L'ensemble des poids est ensuite normalisé sous la contrainte : $\sum_{n=1}^N \pi_t^n = 1$. À cette dernière étape près, cette mise à jour se résume à la multiplication du poids à l'instant précédent par la vraisemblance de l'observation et le modèle de dynamique, divisé par la densité d'importance. Dans le cas, classique, où densité d'importance et loi de dynamique se confondent (l'observation n'intervient pas dans la mise à jour des particules), l'équation 3.9 se réduit à :

$$\pi_t^n \propto \pi_{t-1}^n p(\mathbf{z}_t|\mathbf{x}_t^n)$$

Ce filtre est connu sous le nom de filtre Bootstrap [Gordon 1993], qui est le premier filtre particulaire publié en 1993, ou Condensation ("CONDitional DENSITY propaGATION") [Isard 1998b], qui est le filtre qui a popularisé cette approche. Ces filtres restent les modèles les plus usuels aujourd'hui en suivi d'objets. Le suivi par filtrage particulaire s'articule donc autour du schéma classique prédiction/correction, à savoir :

- une étape d'initialisation dont l'objectif est de localiser l'objet à suivre,
- une étape de prédiction qui consiste à prévoir l'état de l'objet en s'appuyant sur un modèle de son mouvement,
- une étape de correction qui ajuste l'état prédit en se fondant sur des mesures.

Dans ce volet applicatif de notre thèse, nous utilisons un filtre "Bootstrap" construit autour de la mesure de mise en correspondance de blocs définie à partir de la phase du signal monogène. Nous avons choisi ce filtre très simple parce qu'il n'intègre pas les observations dans l'étape de prédiction et que, à modèle de dynamique fixé, ses performances ne dépendent que la qualité de la mesure de vraisemblance, et donc de notre fonction d'angle. L'état de l'objet suivi est défini par les coordonnées horizontales et verticales de sa boîte englobante et par sa vitesse. Les particules sont donc définies par $\mathbf{x}_t^{(n)} = (x_t, y_t, \dot{x}_t, \dot{y}_t)$, où chaque composante suit une distribution Gaussienne. La hauteur et largeur de la boîte englobante sont fixes, sans rotation ni mise à l'échelle. Les paramètres du modèle de dynamique sont mis à jour à partir des positions estimées précédemment. De plus, pour limiter les problèmes de dérive du modèle nous gardons en mémoire les dix dernières cibles identifiées. Ainsi, pour chaque particule, la vraisemblance est calculée par moyennage du critère obtenu sur ces dix cibles de références $\bar{\phi}$:

$$l(\mathbf{y}_t|\mathbf{x}_t^{(n)}) = \mathcal{N}(\bar{\phi}|0, \sigma^2)$$

où $\mathcal{N}(\cdot|0, \sigma^2)$ est une distribution Gaussienne de moyenne nulle et de variance σ^2 .

3.3.3 Résultats

Ce paragraphe montre qu'un filtre bootstrap construit autour de notre critère donne des résultats intéressants dans des vidéos couleur. Mais sachant que la majeure partie de la littérature s'appuie sur des approches en niveau de gris, nous proposons également des comparaisons entre notre méthode et certains de ces algorithmes. Les résultats montrent que notre méthode appliquée de la même manière sur les images en niveau de gris ou couleur donne de bons résultats.

Séquence couleur :

Notre méthode est initialisée comme pour l'approche naïve (figure 3.9). Le résultat du suivi est pratiquement le même, il n'y a pas perte de la cible quand le poisson fait demi-tour dans la zone à forts changements d'illuminant (voir la figure 3.10 images #31, #51). Cela confirme la robustesse de notre critère aux changements d'illuminant. De plus, il peut être observé que tout au long de la séquence, de l'image #1 à l'image #101, le bloc reste parfaitement centré sur l'oeil du poisson. Cette bonne localisation est due aussi bien à l'encodage de la structure que de la couleur dans notre critère.

Séquence en niveau de gris :

Dans un récent travail, D. Ross propose une grande variété de vidéo en niveau de gris [Ross 2007] qui inclut des environnements variés. Il compare sa méthode de suivi avec l'approche Mean-Shift et WSL afin de montrer la pertinence de son approche. Nous testons notre méthode sur une de ces séquences (voir la figure 3.11) et montrons que notre méthode donne des résultats similaires à ceux de l'IVT. Ici, nous proposons des résultats de suivi par filtrage bootstrap sur la séquence "Plush Toy" [Ross 2007]. Dans cette séquence, la peluche est soumise à différentes positions (déformation et changement d'échelle de l'objet suivi) et à de fortes variations d'illuminant. Le bloc que nous utilisons est un mauvais choix dans un tel cas car il y a beaucoup trop d'informations provenant du fond de la scène dedans. Néanmoins, notre critère apparaît assez robuste et suit avec succès la peluche même dans les dernières images où l'IVT échoue à cause d'une combinaison de poses extrêmes et de changement d'illuminant. Contrairement à l'IVT, aucun prétraitement pour enlever le bruit n'est fait.

Pour conclure les tests, nous appliquons notre filtre sur la séquence du poisson convertie en niveau de gris (figure 3.10). Cet algorithme est capable de suivre le poisson bleu comme dans la séquence couleur sans le perdre mais nous pouvons observer un glissement de la fenêtre de suivi de l'oeil vers la nageoire dorsale. En se référant aux résultats obtenus sur les images couleur, nous constatons que la structure et l'information de luminance ne sont pas suffisantes alors qu'une combinaison avec la

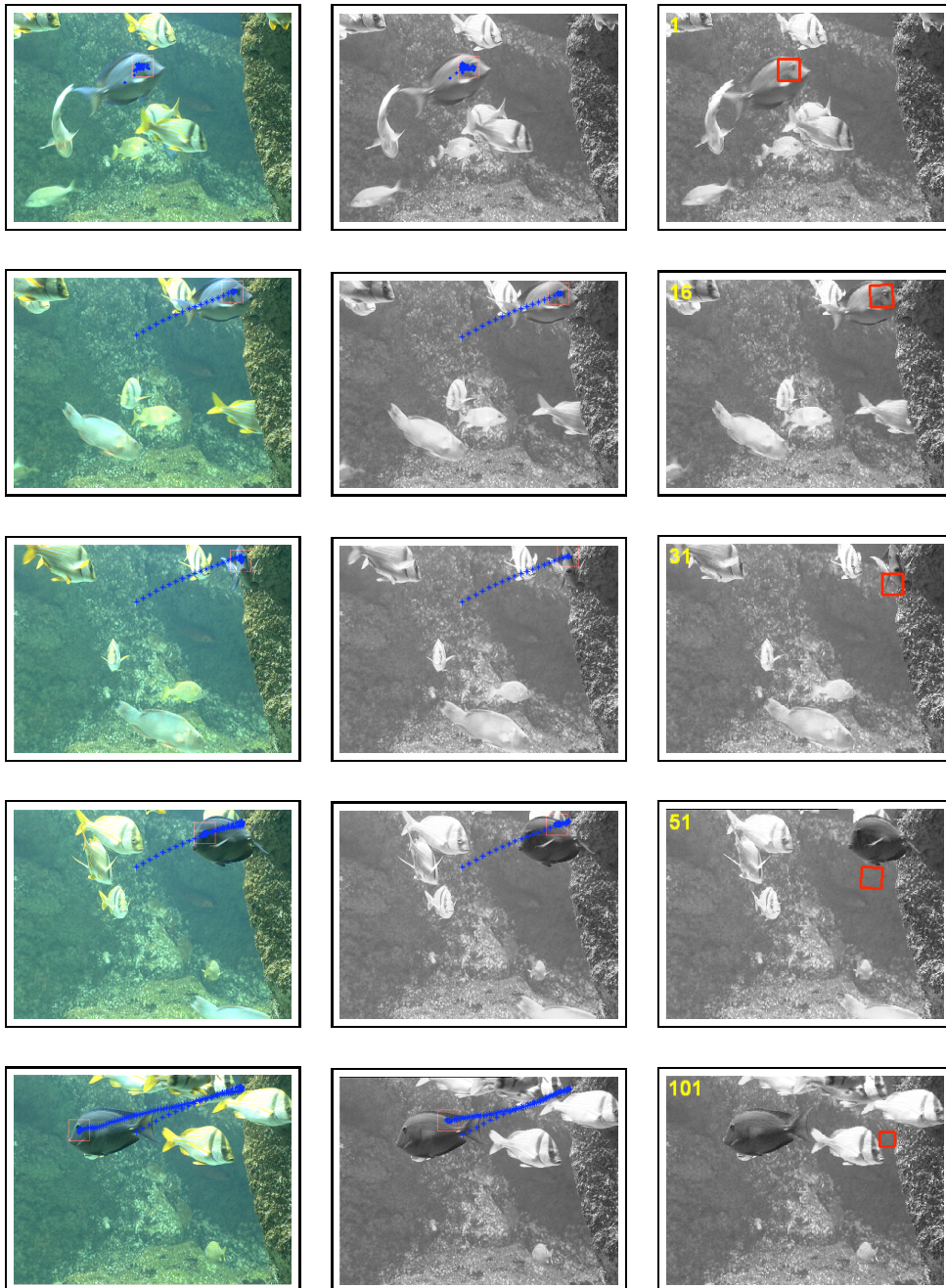


FIGURE 3.10 – Résultats de suivi pour les images #1, #16, #31, #51 et #101 de la séquence Aqu@thèque. Première colonne : résultats pour les images couleur utilisant notre méthode. Deuxième colonne : résultats pour les images en niveau de gris utilisant notre méthode. Troisième colonne : résultats pour les images en niveau de gris utilisant la méthode IVT.

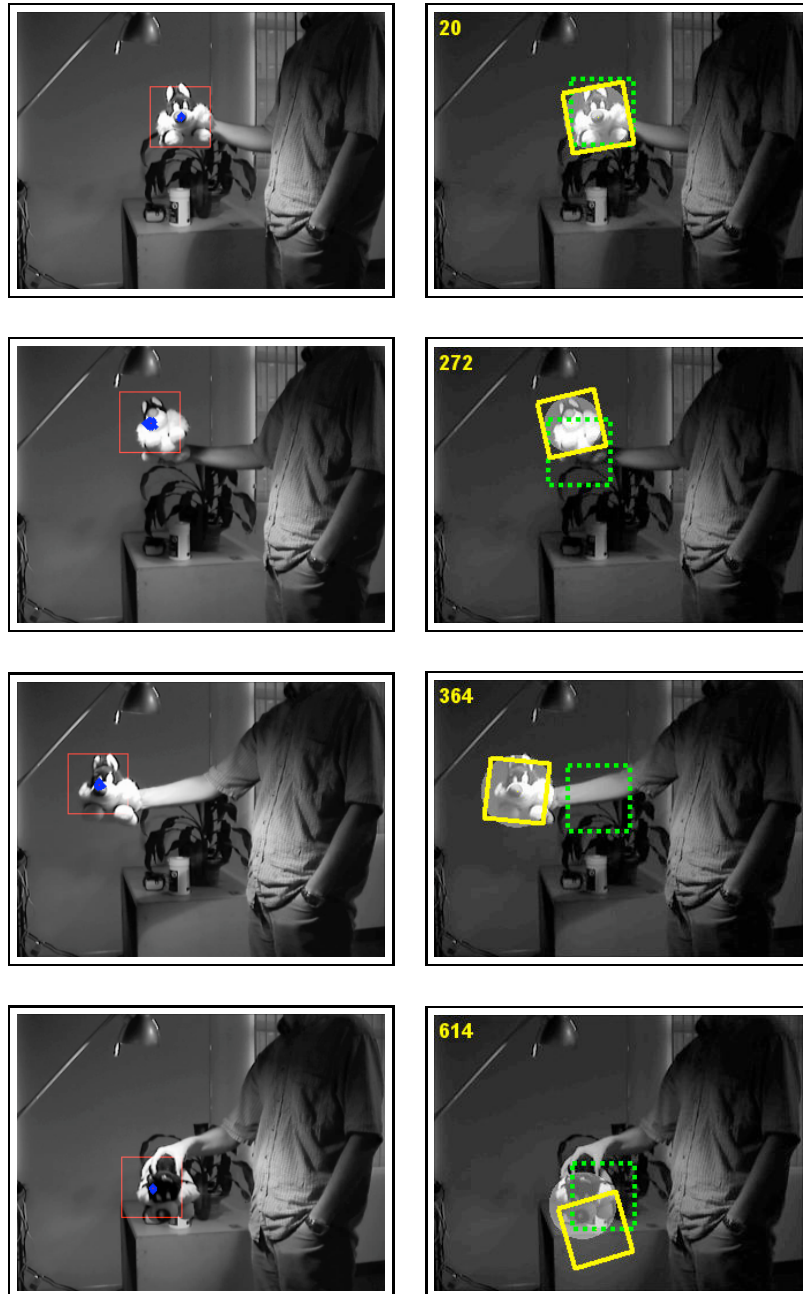


FIGURE 3.11 – Première colonne : notre méthode. Deuxième colonne : la méthode WSL [Jepson 2003] (ellipse en surbrillance), la méthode Mean Shift [Comaniciu 2000] (carré vert pointillé) et la méthode IVT [Ross 2007] (carré jaune).

couleur est pertinente. Sur la même séquence en niveau de gris, l'IVT échoue quand le poisson fait demi-tour (figure 3.10 image #31).

3.4 Détection de points d'intérêt

La détection de points d'intérêts est largement étudiée depuis 20 ans. Le lecteur pourra se référer à [Tuytelaars 2008] pour un état de l'art et de plus amples détails sur les méthodes existantes. Dans cette section, nous voulons proposer un nouveau cadre pour la détection de points d'intérêt basé sur de la géométrie différentielle. En plongeant une image en niveau de gris comme une surface dans \mathbb{R}^3 (\mathbb{R}^5 pour une image couleur), nous pouvons utiliser les outils de géométrie différentielle tels que la métrique riemannienne de la variété. Ce travail préliminaire est proposé dans [Noble 1988] et un complément sur la relation avec le discriminant est donné dans un premier temps. Ensuite, une extension est proposée pour prendre en considération des données tridimensionnelles telles qu'un espace-échelle ou une séquence d'images. Enfin, un nouveau détecteur de points d'intérêt est proposé et de nombreuses comparaisons sont faites avec les méthodes existant dans l'état de l'art.

3.4.1 Détecteur de Harris et géométrie différentielle

3.4.1.1 Détecteur de Harris

Le détecteur de Harris est introduit pour la première fois dans [Harris 1988] par Harris et Stephens, c'est l'un des détecteurs de points d'intérêt le plus référencé dans la littérature. Plusieurs extensions ont été proposées telles que le Harris-Laplace [Mikolajczyk 2004] pour résoudre des problèmes d'invariance d'échelle ou le Harris-Affine [Mikolajczyk 2004] pour résoudre des problèmes de transformation affine. Le détecteur de Harris est basé sur le lissage du tenseur de structure d'une image en niveau de gris I , aussi appelé matrice d'autocorrélation. Pour détecter des points d'intérêt, le comportement des valeurs propres de cette matrice est étudié ce qui donne la distribution du gradient d'un point $p \in \mathbb{R}^2$ dans un voisinage local. Rappelons la définition tenseur de structure :

$$\tilde{A} = g(\sigma_I) * \nabla I \nabla I^T = g(\sigma_I) * \begin{pmatrix} I_x^2(p, \sigma_D) & I_x(p, \sigma_D) I_y(p, \sigma_D) \\ I_x(p, \sigma_D) I_y(p, \sigma_D) & I_y^2(p, \sigma_D) \end{pmatrix}$$

où $I_x(p, \sigma_D) = \frac{\partial g_{\sigma_D}}{\partial x} * I(p)$, $I_y(p, \sigma_D) = \frac{\partial g_{\sigma_D}}{\partial y} * I(p)$ et $g_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$.

Les paramètres σ_D et σ_I sont respectivement les échelles de différentiation et d'intégration. Elles sont proportionnelles l'une à l'autre d'un facteur s [Mikolajczyk 2004] (voir travaux de [Lindeberg 1998]). Dans [Mikolajczyk 2004], la version normalisée en échelle de \tilde{A} est utilisée grâce à l'introduction de dérivées normalisées en échelle. Enfin, Harris et Stephens proposent une certaine fonction pour détecter des coins dans une image :

$$H = \det(\tilde{A}) - k \text{tr}^2(\tilde{A}) \quad (3.10)$$

Il est à noter que cette formulation permet d'éviter le calcul explicite des valeurs propres.

3.4.1.2 Géométrie différentielle

A. Noble [Noble 1988] donne une interprétation du détecteur de Harris dans le cadre de la géométrie différentielle. Considérons une image en niveau de gris I lissée par un noyau gaussien d'écart-type σ_D comme une surface M plongée dans \mathbb{R}^3 muni de la métrique :

$$m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mu \end{pmatrix}$$

Plus précisément, la paramétrisation est donnée par :

$$\begin{aligned} \psi : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longrightarrow (x, y, I(x, y, \sigma_D)) \end{aligned}$$

La métrique m induit une métrique riemannienne :

$$ds^2 = E dx^2 + 2F dx dy + G dy^2 = dx^2 + dy^2 + \mu dI^2$$

sur M appelée la première forme fondamentale de la surface. La représentation matricielle de ds^2 dans la base (ψ_x, ψ_y) de l'espace tangent $T_p M$, pour $p \in M$ est donnée par :

$$I_p = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$$

où $E = \langle \psi_x, \psi_x \rangle_m$, $F = \langle \psi_x, \psi_y \rangle_m$, $G = \langle \psi_y, \psi_y \rangle_m$ et $\langle \cdot, \cdot \rangle_m$ est le produit scalaire induit sur $T_p M$ par m . En calculant explicitement E, F et G , nous obtenons :

$$I_p = \begin{pmatrix} 1 + \mu I_x^2(p, \sigma_D) & \mu I_x(p, \sigma_D) I_y(p, \sigma_D) \\ \mu I_x(p, \sigma_D) I_y(p, \sigma_D) & 1 + \mu I_y^2(p, \sigma_D) \end{pmatrix} \quad (3.11)$$

Cette expression est similaire à celle du tenseur de structure quand $\mu = 1$ dans m . Nous posons :

$$A = g(\sigma_I) * I_p$$

En choisissant μ comme fonction de l'écart-type ($\mu = \sigma_D^2$) donne une version normalisée en échelle similaire à [Mikolajczyk 2004].

Le discriminant

$$\Delta = \text{tr}^2(A) - 4 \det(A)$$

du polynôme caractéristique

$$p_A(\lambda) = \lambda^2 - \text{tr}(A)\lambda + \det(A)$$



FIGURE 3.12 – Deux images successives de la séquence Graffiti avec les points d'intérêt détectés (croix vertes).

de A n'est pas négatif puisque A est symétrique. A noter que $\tilde{\Delta} = -\frac{1}{4}\Delta$ est précisément le détecteur de Harris pour $k = 1/4$ dans l'équation 3.10. Plus généralement, nous posons :

$$H_k = -k\Delta_k$$

où $\Delta_k = \text{tr}^2(A) - \frac{1}{k}\det(A)$. Il apparaît que le signe de H_k est l'opposé du signe du discriminant relaxé Δ_k . Puisque $0 < k < 1/4$, ce dernier peut maintenant prendre des valeurs négatives.

3.4.2 Géométrie différentielle dans un espace-échelle

Ce paragraphe étend l'approche précédente à une représentation dans un espace-échelle. Soit une image en niveau de gris I , lissée avec un noyau gaussien d'écart-type s , vue comme une 3-variété plongée dans \mathbb{R}^4 muni de la métrique :

$$m = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \nu & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix}$$

Plus précisément, la paramétrisation est donnée par :

$$\begin{aligned} \psi : \mathbb{R}^3 &\longrightarrow \mathbb{R}^4 \\ (x, y, s) &\longrightarrow (x, y, s, I(x, y, s)) \end{aligned}$$

La métrique m induit une métrique riemannienne qui est donnée dans la base (ψ_x, ψ_y, ψ_s) de l'espace tangent T_pM , pour $p \in M$ par :

$$I_p = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

où

$$\begin{aligned} a &= \langle \psi_x, \psi_x \rangle_m, \quad b = \langle \psi_x, \psi_y \rangle_m, \quad c = \langle \psi_x, \psi_s \rangle_m \\ d &= \langle \psi_y, \psi_y \rangle_m, \quad e = \langle \psi_y, \psi_s \rangle_m, \quad f = \langle \psi_s, \psi_s \rangle_m \end{aligned}$$

et $\langle \cdot, \cdot \rangle_m$ est le produit scalaire induit sur $T_p M$ par m . Le calcul explicite des coefficients donne :

$$I_p = \begin{pmatrix} 1 + \mu I_x^2 & \mu I_x I_y & \mu I_x I_s \\ \mu I_x I_y & 1 + \mu I_y^2 & \mu I_y I_s \\ \mu I_x I_s & \mu I_y I_s & \nu + \mu I_s^2 \end{pmatrix} \quad (3.12)$$

La version lissée est donnée par $A = g(\sigma_I, \tau_I) * I_p$ où g est un noyau gaussien 3D. Nous procédons comme précédemment, le polynôme caractéristique de A est :

$$p_A(\lambda) = -\lambda^3 + \text{tr}(A)\lambda^2 + \frac{1}{2}(\text{tr}(A^2) - \text{tr}^2(A))\lambda + \det(A)$$

Et les racines de p_A sont implémentées en utilisant les formules de Cardan pour les équations du troisième degré. Rappelons que le discriminant d'un polynôme $P(x) = ax^3 + bx^2 + cx + d = 0$ de degré 3 (avec a, b, c et d réel) est

$$\Delta = q^2 + \frac{4}{27}p^3, \quad p = -\frac{b^2}{3a^2} + \frac{c}{a}, \quad q = \frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a} \right) + \frac{d}{a}$$

Rappelons également que :

- $\Delta > 0$ implique que P possède une racine réelle et deux racines complexes conjuguées.
- $\Delta = 0$ implique que P possède deux racines réelles (dont l'une est d'ordre 2).
- $\Delta < 0$ implique que P possède trois racines réelles.

Dans le cas présent $a = -1$, $b = \text{tr}(A)$, $c = \frac{1}{2}(\text{tr}(A^2) - \text{tr}^2(A))$ et $d = \det(A)$. Relaxant le coefficient $4/27$, nous introduisons :

$$H_k = -(q^2 + kp^3)$$

La gamme de valeurs positives peut être élevée et, pour faciliter le seuillage dans l'algorithme proposé, nous considérons :

$$H_{3D} = \begin{cases} \sqrt{H_k} = \sqrt{-(q^2 + kp^3)} & \text{si } H_k > 0 \\ H_k & \text{sinon} \end{cases} \quad (3.13)$$

Les expérimentations montrent que pour un bon k , H_{3D} se comporte comme le détecteur de Harris : il est positif pour les coins, négatif pour les contours et proche de zéro pour les zones homogènes (voir figure 3.13). Dans la suite, nous choisissons $k = 0.148$.

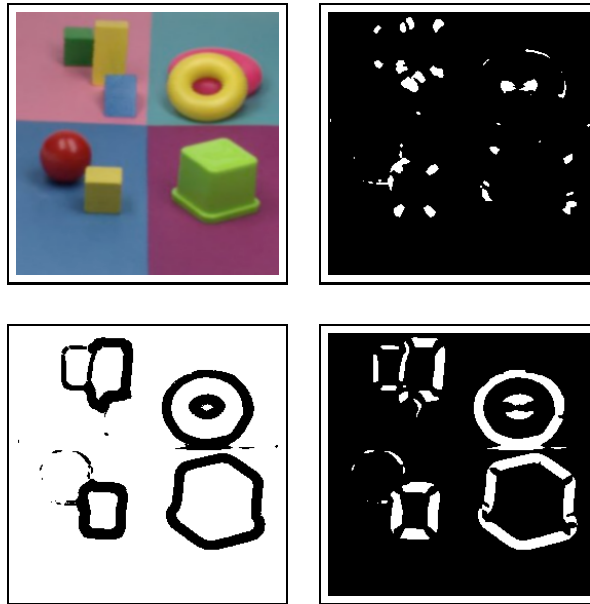


FIGURE 3.13 – De gauche à droite. Première ligne : image originale, $H_\alpha > 0$. Deuxième ligne : $H_\alpha \approx 0$, $H_\alpha < 0$ (les zones blanches correspondent aux structures détectées)

3.4.3 Résultats

3.4.3.1 Implémentation

L'espace-échelle d'une image est implémenté comme suit :

- une échelle de départ s_0 est choisie ($s_0 = 0.5$)
- les niveaux d'échelle sont donnés par $s_n = \xi^n s_0$ (avec $\xi = \sqrt{2}$ voir [Lindeberg 1998, Lowe 1999, Mikolajczyk 2004])
- les dérivées sont implémentées par convolution de l'image en niveau de gris $I(x, y, s)$ avec les dérivées gaussiennes 3D séparables d'écart-type $\sigma_D = 1$ et $\tau_D = 1$.
- à l'échelle initiale s_0 , les points d'intérêts sont extraits en détectant les maximaux locaux du H_{3D} dans le 8-voisinage du pixel (i, j, s_0) .
- les points d'intérêt plus grand que 1% de la valeur maximale (comme pour les détecteur de Harris) sont sélectionnés.

Un autre critère pour sélectionner les points est introduit : seuls les points dont la valeur du le discriminant est décroissante à travers les échelles sont finalement gardés. En effet, un "vrai coin" tend à devenir une zone homogène quand l'échelle grandit, et donc la valeur du discriminant doit diminuer. En pratique, nous regardons le signe de la dérivée de H_{3D} par rapport à l'échelle.

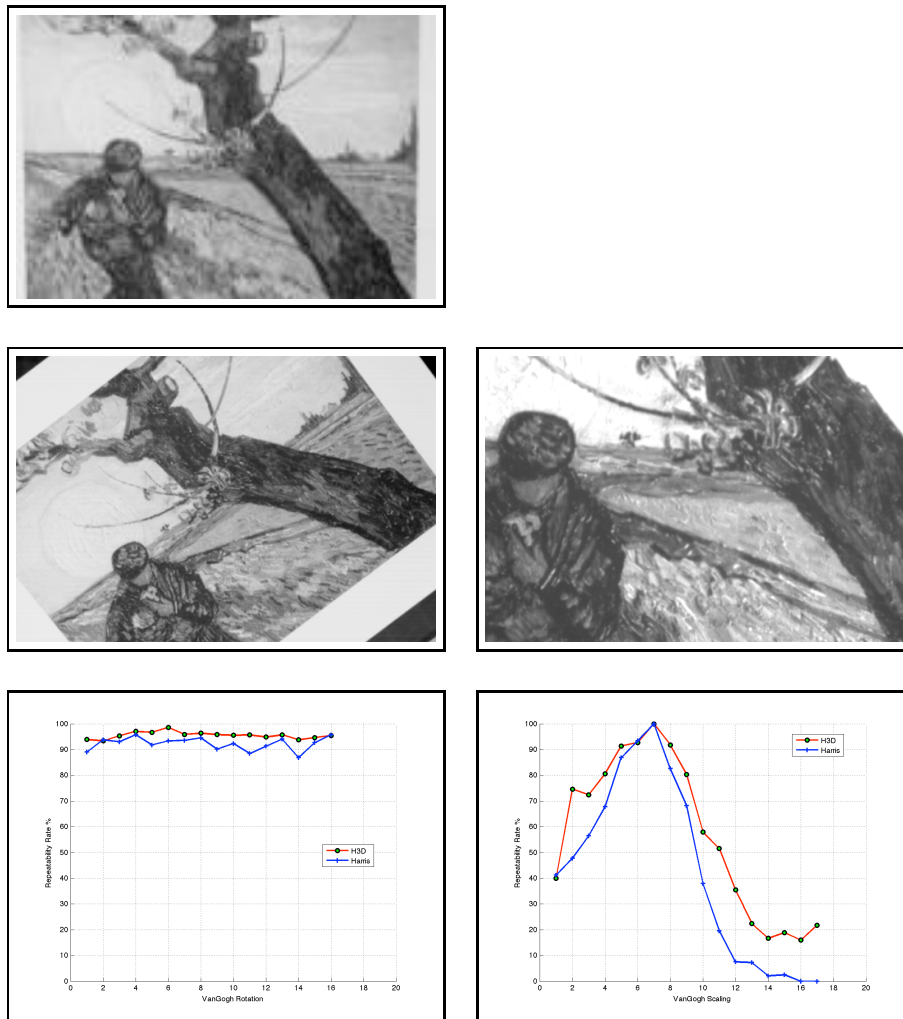


FIGURE 3.14 – Images de la séquence VanGogh. De gauche à droite. Première ligne : image originale. Deuxième ligne : rotation, zoom. Troisième ligne : répétabilité pour la rotation et le zoom.

3.4.3.2 Méthode d'évaluation

De nombreux de travaux ont été fait dans l'évaluation des détecteurs de points d'intérêt, le plus populaire est peut-être le travail exhaustif de K. Mikolajczyk et C. Schmid dans [Mikolajczyk 2005, Schmid 2000]. Nous utiliserons plus particulièrement la répétabilité R comparée avec d'autres détecteurs. Comme dans [Schmid 2000], le taux de répétabilité entre une paire d'images (I_1, I_2) possédant (n_1, n_2) points d'intérêt dans la partie commune de l'image et respectant la transformation relative T est donnée par la relation :

$$R(I_1, I_2, T) = \frac{\#\text{correspondences}}{\min(n_1, n_2)}$$

où deux points d'intérêt $s_1 = (x_1, y_1) \in I_1$ et $s_2 = (x_2, y_2) \in I_2$ correspondent ssi $D(Ts_1, s_2) < r_0$ (r_0 est le rayon du voisinage au point Ts_1). Les illustrations sont données dans Fig. 3.14.

Le critère de répétabilité est modifié [Willis 2009] en exigeant que deux points d'intérêt, $s_1 = (x_1, y_1) \in I_1$ et $s_2 = (x_2, y_2) \in I_2$ correspondent ssi $\min D(Ts_1, s_i) = s_2$ et $D(Ts_1, s_2) < r_0$ et $\min D(s_i, T^{-1}s_2) = s_1$ et $D(s_1, T^{-1}s_2) < r_0$. Ce critère requiert que les points d'intérêt correspondants soient les plus proches voisins l'un de l'autre sous transformation directe et inverse, et que leur distance point à point est plus petite qu'un rayon r_0 . L'évaluation sur des images de la figure 3.15 est faite avec ce critère ainsi qu'avec deux rayons différents $r_0 = 1.5$ et $r_0 = 5$.

3.4.3.3 Discussion

La figure 3.12 donne un exemple de points d'intérêt détectés avec le détecteur H_{3D} . Afin d'évaluer le détecteur proposé, nous procédons en deux étapes. D'abord nous le comparons avec le détecteur de Harris, qui d'après le travail de C. Schmid *et al.* [Schmid 2000] est le plus répétable, sur les séquence d'images VanGogh³. Les résultats dans la figure 3.14 montrent la répétabilité du détecteur H_{3D} et celle du détecteur de Harris classique lors de rotations et de changements d'échelle. Puisque notre détecteur est construit sur le même modèle que le détecteur de Harris, nous pouvons espérer obtenir des résultats similaires et potentiellement meilleurs. Les premiers résultats donne une réponse partielle à cette question et montrent qu'il possède la même robustesse aux rotations et changements d'échelle.

La deuxième étape compare notre méthode avec différents détecteurs de points/régions d'intérêt sur une base de données⁴ bien connue de [Mikolajczyk 2005] et illustrée dans la figure 3.15. Les détecteurs de Harris, déterminant de la hessienne (DoH) et laplacien de gaussiennes (LoG) sont implémentés suivant le code⁵ mis en ligne par I. Laptev avec les paramètres par défaut (échelle de dérivation $\sigma_D = 2$ et échelle d'intégration $\sigma_I = \sqrt{2}\sigma_D$), la sélection des points les plus pertinents est faite par seuillage du détecteur de Harris avec 1% de la valeur maximale et un maximum

3. <http://lear.inrialpes.fr/people/mikolajczyk/Database/index.html>

4. <http://www.robots.ox.ac.uk/vgg/research/affine/index.html>

5. <http://www.nada.kth.se/laptev/code.html>

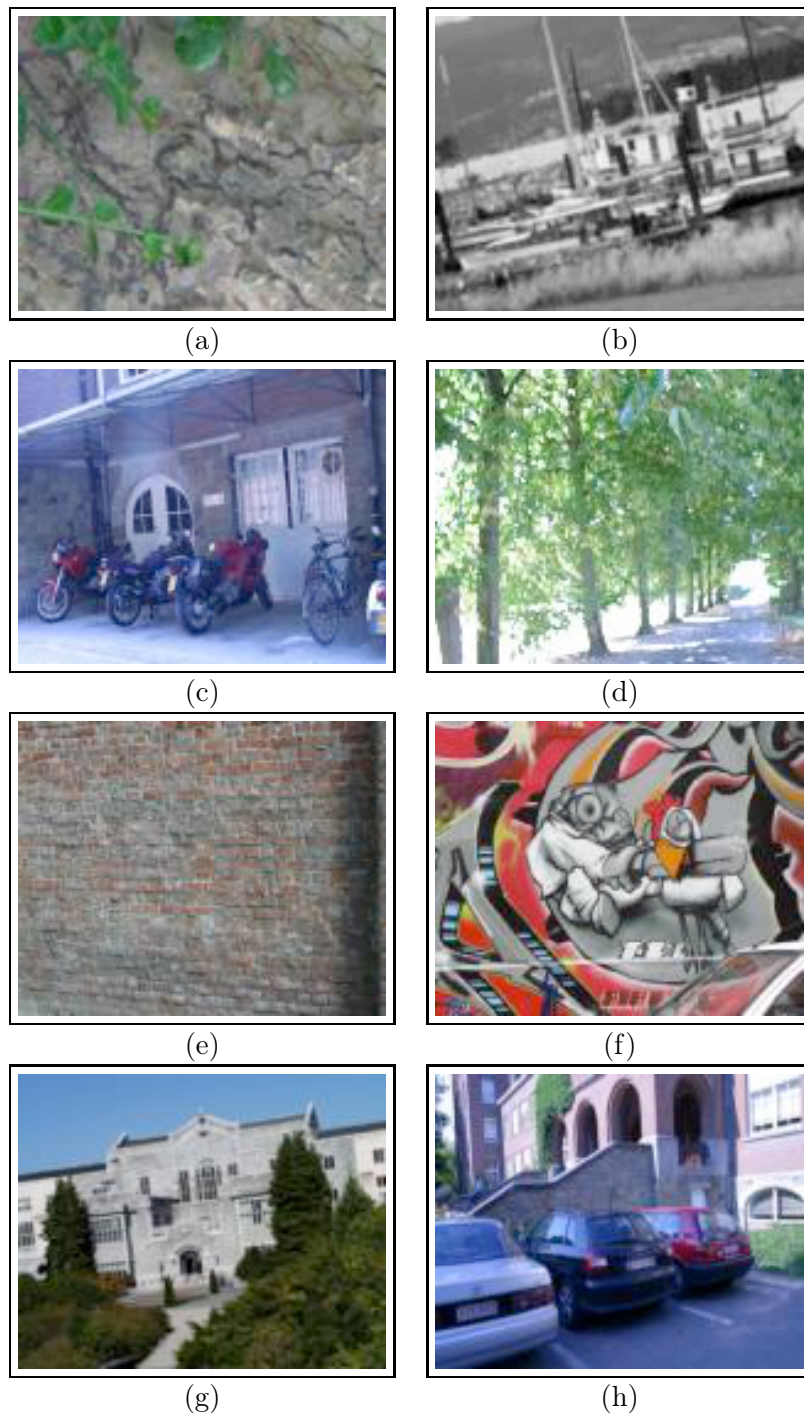


FIGURE 3.15 – Images de la base de données [Mikolajczyk 2005]. (a)(b) sont des transformations en échelle et rotation, (c)(d) sont des changements de point de vue, (e)(f) sont des images floutées, (g) subit différentes compressions JPEG et (h) des variations de luminosité.

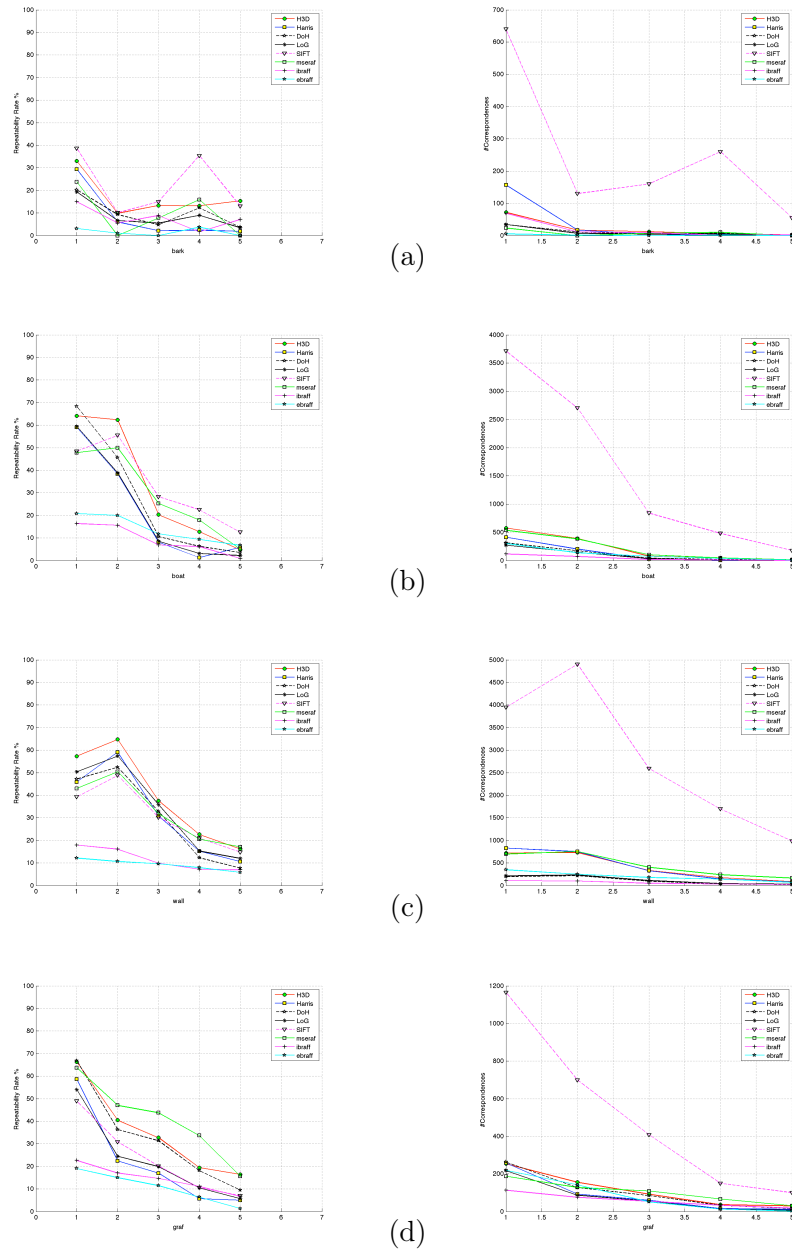


FIGURE 3.16 – Résultats pour les séquences (a), (b), (c), (d) avec $r_0 = 1.5$. Première colonne, taux de répétabilité en pourcentage. Deuxième colonne, nombre de correspondances.

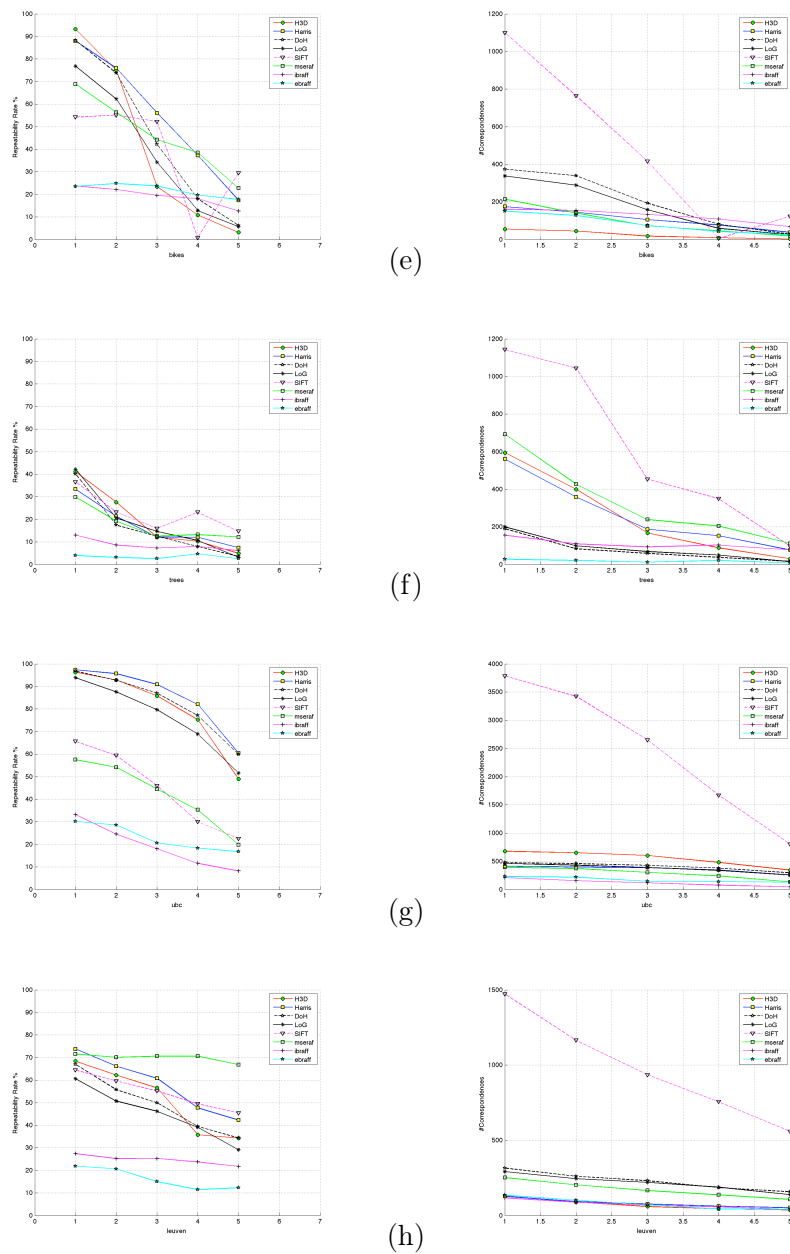


FIGURE 3.17 – Résultats pour les séquences (e), (f), (g), (h) avec $r_0 = 1.5$. Première colonne, taux de répétabilité en pourcentage. Deuxième colonne, nombre de correspondances.

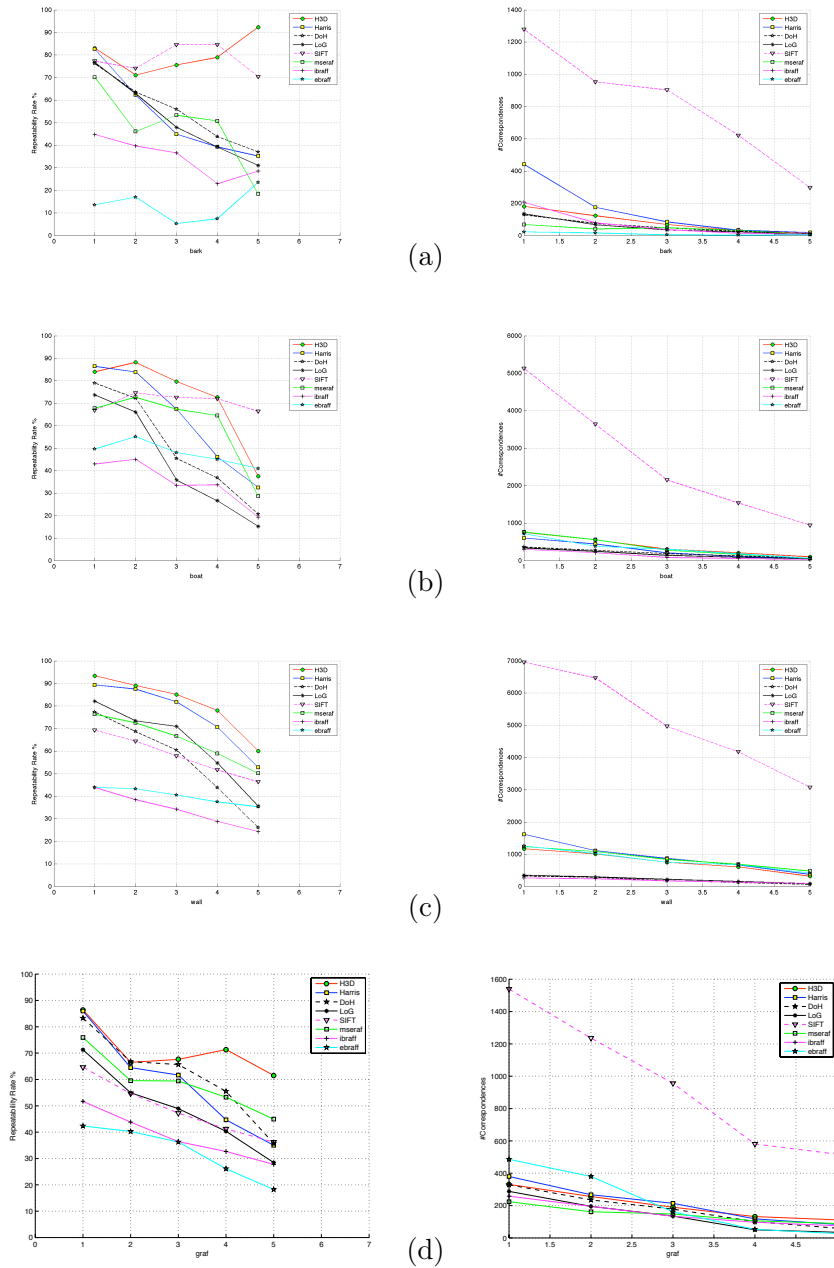
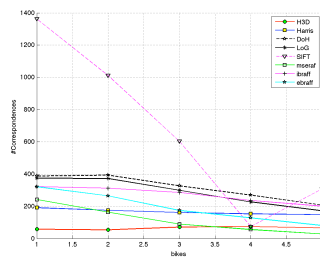
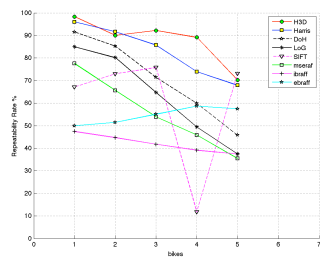
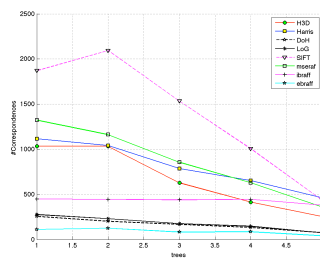
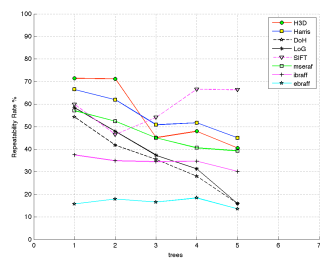


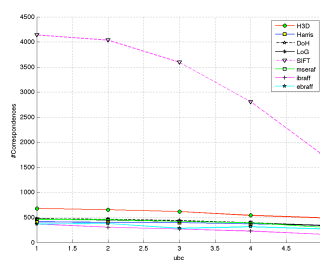
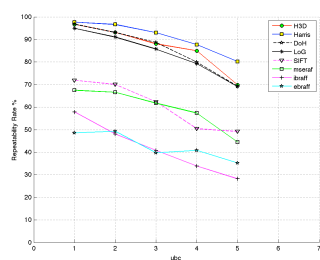
FIGURE 3.18 – Résultats pour les séquences (a), (b), (c), (d) avec $r_0 = 5$. Première colonne, taux de répétabilité en pourcentage. Deuxième colonne, nombre de correspondances.



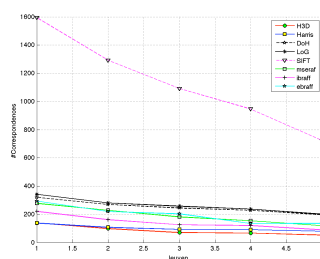
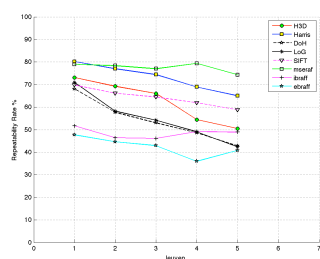
(e)



(f)



(g)



(h)

FIGURE 3.19 – Résultats pour les séquences (a), (b), (c), (d) avec $r_0 = 5$. Première colonne, taux de répétabilité en pourcentage. Deuxième colonne, nombre de correspondances.

de 1000 points pour les deux autres. La méthode SIFT [Lowe 2004] est accessible sur le site internet des auteurs et est utilisée avec les paramètres par défaut. Les résultats des méthodes MSER [Matas 2002], IBR et EBR [Tuytelaars 2004] sont ceux obtenus dans [Mikolajczyk 2005].

Rotation et changement d'échelle :

La première expérimentation est faite sur les images VanGogh soumises à rotation et changement d'échelle. La figure 3.14 montre l'objet de référence (première ligne, gauche), une rotation de l'objet (première ligne, milieu) et un changement d'échelle de l'objet (première ligne, droite). A la deuxième ligne, le taux de répétabilité pour chaque séquence est montré et comme nous pouvons le constater, notre détecteur apparait aussi répétable que le détecteur de Harris dans les deux cas.

Changement de point de vue :

Les effets du changement de point de vue pour la scène du graffiti et pour les images contenant des textures répétées dans la figure 3.15(c,d) sont affichées dans les figures 3.16(c,d) et 3.18(c,d). Les figures 3.16 et 3.18 montrent sur la première ligne la répétabilité et sur la deuxième ligne le nombre de correspondance absolu pour un rayon de r_0 égal à 1.5 et 5 pixels respectivement. Notre méthode donne de bons résultats et est meilleure que le détecteur de Harris. Nous pouvons voir qu'en augmentant le rayon, notre détecteur est plus performant que les autres.

Changement d'échelle :

Les figures (3.16, 3.18)(a,b) montrent les résultats pour la scène texturée et la scène structurée de la figure 3.15(a,b). Les résultats montrent que la méthode SIFT donne de meilleurs résultats que notre détecteur. Mais notre détecteur est toujours meilleur que le détecteur de Harris.

Blur :

Les figures 3.17(e,f) et 3.19(e,f) montrent les résultats pour la scène texturée et la scène structurée de la figure 3.15(e,f). Puisque notre méthode n'est pas invariante par échelle, nous observons que la répétabilité du H_{3D} chute quand le floutage devient grand pour $r_0 = 1.5$. Cela montre que notre méthode n'a pas une bonne localisation et ceci est dû au fait que nous ne considérons pas assez l'échelle dans notre algorithme. Le détecteur de Harris donne de bons résultats grâce au choix de σ_D et σ_I qui compense l'effet de flou et donc la délocalisation des points d'intérêt. Pour $r_0 = 5$, cela confirme que notre détecteur est en fait répétable comme le détecteur de Harris dans ce cas mais le H_{3D} souffre de la délocalisation de ses points d'intérêt.

Artéfacts JPEG :

Les figures 3.17(g) et 3.19(g) montrent les résultats pour une séquence subsistante de fortes compression JPEG de la figure 3.15(g). Le détecteur de Harris est le meilleur mais il y a un fort pourcentage pour les quatre premières méthodes.

Changement d'illuminant :

Les figures 3.17(h) et 3.19(h) montre des résultats dans le cas de changement d'illuminant dans les images de la figure 3.15(h). Toutes les courbes sont presque horizontales, montrant ainsi une bonne robustesse aux changements d'illuminant bien que les méthodes MSER et Harris obtiennent les meilleurs résultats.

Les expérimentations montrent que le détecteur H_{3D} est performant dans différentes conditions. Il est important de noter que nous n'utilisons pas toutes les possibilités offertes par la représentation d'espace-échelle. De futures recherches pourront être menées en utilisant plus en profondeur certaines caractéristiques des surfaces comme la deuxième forme fondamentale afin de sélectionner plus pertinemment les points d'intérêt.

3.4.4 Complément 1 : H_{3D} couleur

Jusqu'à maintenant, nous avons travaillé avec des images en niveau de gris mais grâce au cadre de la géométrie différentielle, il est très facile d'étendre tous ces résultats aux images couleur puisqu'il suffit de changer la paramétrisation ψ .

Soit $I(x, y) = (I_1(x, y), I_2(x, y), I_3(x, y))$ une image couleur dans RVB vue comme une 3-variété plongée dans \mathbb{R}^6 muni de la métrique :

$$m = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

La paramétrisation de la surface est donnée par :

$$\begin{aligned} \psi : \mathbb{R}^3 &\longrightarrow \mathbb{R}^6 \\ (x, y, s) &\longrightarrow (x, y, s, R(x, y, s), V(x, y, s), B(x, y, s)) \end{aligned}$$

La matrice de la première forme fondamentale de cette surface est :

$$I_p = \begin{pmatrix} 1 + \sum_{i=1}^3 \lambda_i I_{ix}^2 & \sum_{i=1}^3 \lambda_i I_{ix} I_{iy} & \sum_{i=1}^3 \lambda_i I_{ix} I_{is} \\ \sum_{i=1}^3 \lambda_i I_{ix} I_{iy} & 1 + \sum_{i=1}^3 \lambda_i I_{iy}^2 & \sum_{i=1}^3 \lambda_i I_{iy} I_{is} \\ \sum_{i=1}^3 \lambda_i I_{ix} I_{is} & \sum_{i=1}^3 \lambda_i I_{iy} I_{is} & \nu + \sum_{i=1}^3 \lambda_i I_{is}^2 \end{pmatrix} \quad (3.14)$$

Par suite, il suffit d'appliquer le détecteur H_{3D} en utilisant la trace et le déterminant de cette matrice.

3.4.5 Complément 2 : H_{3D} et signal monogène couleur

Dans ce paragraphe, nous travaillons avec la fonction d'angle φ_V dans le but de détecter des points d'intérêt dans une couleur choisie. Nous reprenons l'approche développée dans le paragraphe 3.4.2 et nous paramétrisons cette fois-ci la surface par :

$$\begin{aligned} \psi : \mathbb{R}^3 &\longrightarrow \mathbb{R}^4 \\ (x, y, s) &\longrightarrow (x, y, s, \varphi_V(x, y, s)) \end{aligned}$$

La matrice de la première forme fondamentale s'écrit alors :

$$I_p = \begin{pmatrix} 1 + (\varphi_{V,x})^2 & \varphi_{V,x}\varphi_{V,y} & \varphi_{V,x}\varphi_{V,\tau} \\ \varphi_{V,x}\varphi_{V,y} & 1 + (\varphi_{V,y})^2 & \varphi_{V,y}\varphi_{V,\tau} \\ \varphi_{V,x}\varphi_{V,\tau} & \varphi_{V,y}\varphi_{V,\tau} & 1 + (\varphi_{V,\tau})^2 \end{pmatrix}$$

Nous pouvons alors utiliser le détecteur H_{3D} associé au polynôme caractéristique de cette matrice.

Nos premiers résultats utilisant cette approche montrent que nous détectons bien les points d'intérêt dans une couleur choisie. La figure 3.20 montre une comparaison de notre approche avec deux méthodes basées sur la couleur décrites dans [van de Weijer 2006]. La figure 3.21 donne un aperçu de notre détecteur sur deux des images de la base de données [Mikolajczyk 2005].

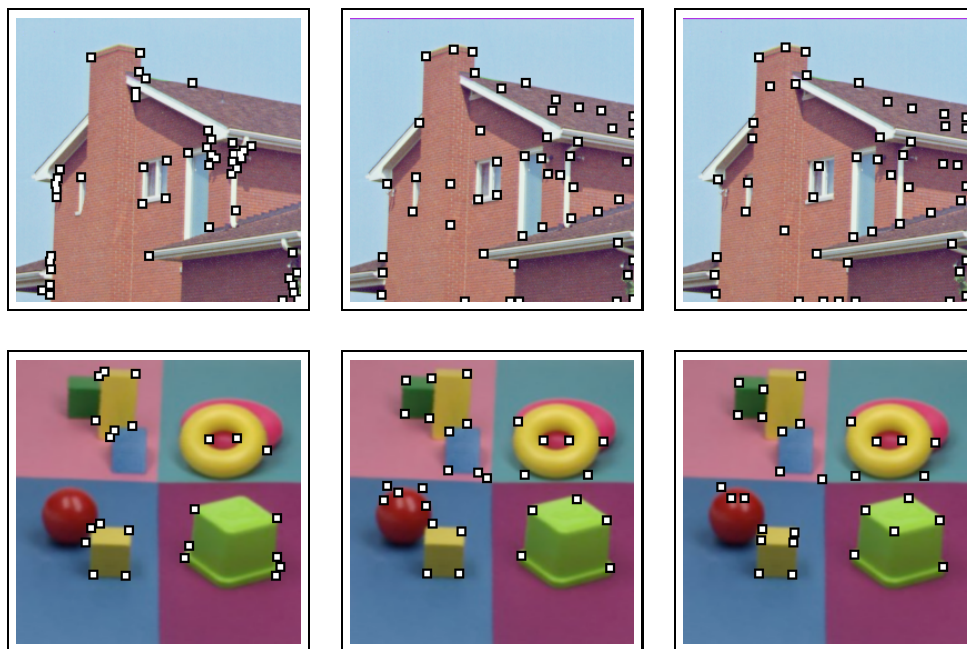


FIGURE 3.20 – De gauche à droite : détecteurs H_{3D} , Color Boost Harris et Color Harris appliqués à l'image de maison (première ligne) et à l'image de jouets (deuxième ligne). La couleur choisie est le rouge pour la maison et le jaune pour les jouets.

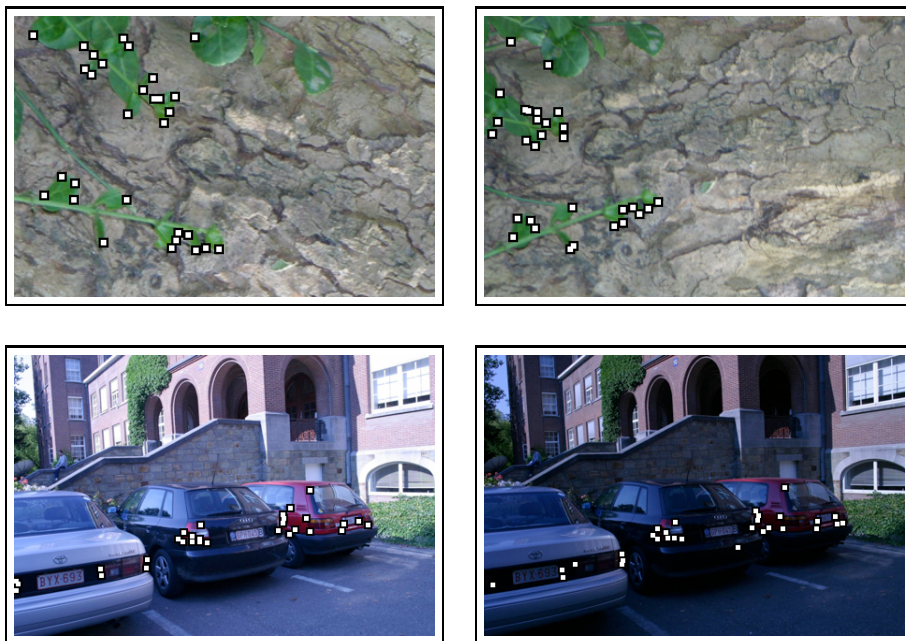


FIGURE 3.21 – Notre méthode appliquée sur les images 1 et 2 de la séquence "Bark" (première ligne) et sur les images 1 et 4 de la séquence "Leuven" (deuxième ligne) provenant de [Mikolajczyk 2005]. La couleur choisie est le vert pour "Bark" et le rouge pour "Leuven"

Conclusion

Les travaux décrits dans cette thèse portent sur la définition et la mise en application des propriétés du signal monogène couleur. Nous avons tiré profit d'une "double utilisation" du cadre mathématique des algèbres de Clifford : tout d'abord pour modéliser la géométrie de l'espace couleur et ensuite pour obtenir une généralisation du signal analytique. La construction proposée, nous l'avons vu, soulève de façon cruciale la question de la définition de la phase d'une image couleur. Nous avons apporté certains éléments de compréhension mais il n'en demeure pas moins que cette question est une des perspectives très motivantes pour la poursuite de l'étude du signal monogène couleur.

Tout comme la nature holomorphe du signal analytique est finalement peu exploitée en théorie du signal, on peut s'interroger sur les propriétés du signal monogène couleur au regard du caractère monogène. Nous n'avons pas du tout abordé ce problème, certainement ardu. Une des pistes possibles qui mérite d'être explorée consiste à analyser le signal via une construction alternative, que nous n'avons pas décrite dans le mémoire, utilisant la formule intégrale de Cauchy et les techniques de l'analyse de Clifford [Fleischmann 2008].

De travaux récents [Batard 2010] ont abouti à la définition d'une transformée de Fourier Clifford pour des fonctions à valeurs dans la partie vectorielles d'une algèbre de Clifford de type $\mathbb{R}_{n,0}$. Cette transformée a notamment été utilisée pour le filtrage d'images couleur et la définition de descripteurs de Fourier généralisés couleur [Menesson 2010]. Il est pour le moins tentant d'appliquer cette transformée sur le signal monogène couleur (qui est à valeurs dans la partie vectorielle de $\mathbb{R}_{5,0}$) afin d'analyser les fréquences des structures locales et des couleurs. Un travail théorique peut être envisagé pour mettre à jour des propriétés spectrales analogues à celle du signal analytique (voir dans cette optique [Felsberg 2001]).

En ce qui concerne les applications, nous nous sommes bornés à donner des exemples significatifs. Bien évidemment, ces exemples peuvent être repris de façon plus pertinente dans des contextes applicatifs précis : amélioration du terme de régularisation pour le flot optique, choix adapté de la métrique pour la segmentation ou de la vraisemblance dans le filtrage particulière pour le suivi.

D'une façon générale, les travaux menés s'inscrivent dans la problématique qui consiste à fournir des modèles mathématiques très génériques pour l'étude des images multicanaux ou multidimensionnelles dans l'esprit de [Batard 2009, Batard 2008]. La construction que l'on a proposée peut facilement être étendue à ce type d'images. Cependant, la compréhension des phénomènes, en

termes de théorie du signal, est loin d'être évidente. Il n'en demeure pas moins que la mise en place de tels modèles semble être un enjeu très actuel du traitement d'images.

Pour finir, je souhaiterais remercier la région Poitou-Charentes ainsi que l'Office of Naval Research (NOR Grant N00014-09-1-0493) pour leur support financier durant cette thèse.

Espace-échelle de Poisson

Dans cette annexe, nous reprenons les travaux de M. Felsberg et G. Sommer [Felsberg 2004] portant sur l'espace-échelle de Poisson.

A.1 Définition

Quand nous parlons d'espace-échelle, il est naturel de penser immédiatement à l'espace-échelle gaussien car c'est celui que la majorité des chercheurs en traitement d'images utilisent. Pendant longtemps, il a même été considéré comme l'unique espace-échelle linéaire (voir [Babaud 1986][Iijima 1962][Lindeberg 1994]) jusqu'aux travaux de M. Felsberg. La question de la définition d'un espace-échelle se pose donc. D'un point de vue théorique un espace-échelle vérifie une certaine axiomatique basée sur la théorie des semi-groupes [Weickert 1999][Iijima 1959]. Cette axiomatique est basée sur l'hypothèse qu'il y a une transformation Φ qui transforme une image $f(X)$ ($X = (x, y) \in \mathbb{R}^2$) en une version floutée $f_s(X)$. La transformation a la forme générale :

$$f_s(X) = \Phi(f(X'), X, s) = \int_{\mathbb{R}^2} \phi(f(X'), X, X', s) dX'$$

où $s \in \mathbb{R}^+$. Si cette transformation vérifie certains axiomes, elle établit un espace-échelle linéaire invariant en échelle et rotation.

Définition 2. (*Espace-échelle linéaire invariant en échelle et rotation*) Une transformation définit un espace-échelle linéaire invariant en échelle et rotation si elle vérifie les axiomes suivants :

- A_1 . Linéarité (par rapport à la multiplication).
- A_2 . Invariant par translation.
- A_3 . Propriété de semi-groupe.
- A_4 . Invariant en échelle et rotation.
- A_5 . Positivité

D'après [Weickert 1999] (respectivement [Iijima 1962]) nous détaillons plus formellement les axiomes précédents :

- A_1 . signifie que $\Phi(\lambda f(X'), X, s) = \lambda \Phi(f(X'), X, s)$ pour tout $\lambda \in \mathbb{R}$.
- A_2 . signifie que $\Phi(f(X' - X_0), X, s) = \Phi(f(X'), X - X_0, s)$ pour tout $X_0 \in \mathbb{R}^2$.
- A_3 . signifie qu'il existe une application $S : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ telle que $\Phi(\Phi(f(X''), X', s_1), X, s_2) = \Phi(f(X''), X, S(s_1, s_2))$ pour tout $s_1, s_2 \in \mathbb{R}^+$.

A_4 . signifie qu'il existe une application $T : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ telle que $\Phi(f(aRX'), X, s) = \Phi(f(X'), aRX, T(s, a))$ pour tout $a \in \mathbb{R}^+$ et tout $R \in SO(2)$.

A_5 . signifie que $\Phi(f(X'), X, s) > 0$ pour tout $f > 0$ et tout $s \in \mathbb{R}^+$.

Dans la littérature, il existe d'autres axiomes complémentaires mais qui seraient inutiles dans cette annexe, plus de détails peuvent être trouvés dans [Felsberg 2004][Weickert 1999].

A.2 Espace-échelle de Poisson

Définition 3. *Le noyau de Poisson 2D est donnée par [Stein 1971] :*

$$p(x, y, s) = \frac{s}{2\pi(x^2 + y^2 + s^2)^{3/2}} \quad (\text{A.1})$$

Le noyau de Poisson établit un espace-échelle linéaire et isotrope puisqu'il satisfait les axiomes de la définition 2 (voir également [Felsberg 2001][Duits 2004]) :

Théorème 4. *(Espace-échelle de Poisson) Le noyau de Poisson A.1 établit un espace-échelle linéaire au sens de la définition 2*

Démonstration. La preuve est donnée dans [Felsberg 2004]. □

Noyaux et transformée de Fourier dans $\mathbb{R}_{2,0}$

B.1 Transformée de Fourier

Soient f et g des fonctions vectorielles intégrables sur $\Omega \subset \mathbb{R}$, alors la transformée de Fourier dans $\mathbb{R}_{2,0}$ et son inverse sont définis par :

$$\mathcal{F}_1\{f(x)\}(u) = \int_{\Omega} f(x)e^{-2\pi I_2 x u} dx$$

$$\mathcal{F}_1^{-1}\{g(u)\}(x) = \int_{\Omega} g(u)e^{2\pi I_2 x u} du$$

avec $I_2 = e_1 e_2$ pseudo-scalaire de $\mathbb{R}_{2,0}$. Puisque l'espace engendré par $\{1, I_2\}$ est isomorphe à \mathbb{C} , cette transformée de Fourier agit comme la transformée de Fourier classique (voir [Felsberg 2001] pour les propriétés qui en découlent).

B.2 Noyau de Hilbert

La transformée de Fourier du noyau de Hilbert est donnée par :

$$\mathcal{F}_1 \left\{ \frac{e_1 e_2}{\pi x} \right\} (u) = \frac{u}{|u|}$$

En utilisant le résultat classique :

$$\mathcal{F}_1 \left\{ \frac{1}{\pi x} \right\} (u) = -I_2 \text{signe}(u) = -I_2 \frac{u}{|u|}$$

Nous multiplions par $e_1 e_2$ à gauche et nous obtenons le résultat.

B.3 Noyau de Poisson

La transformée de Fourier du noyau de Poisson 1D est donnée pour $y > 0$ par :

$$\mathcal{F}_1 \left\{ \frac{y}{\pi(x^2 + y^2)} \right\} (u) = e^{-2\pi|u|y}$$

En utilisant l'intégrale n°18 de la table d'intégration 19 dans [Bronstein 1993] :

$$\int_0^{\infty} \frac{\cos(ax)}{1+x^2} dx = \frac{\pi}{2} e^{-|a|}$$

Et en effectuant la substitution $x' = bx$ où $b > 0$, nous obtenons :

$$\int_0^{\infty} \frac{b^2 \cos(ab^{-1}x')}{b^2 + x'^2} b^{-1} dx' = \frac{\pi}{2} e^{-|a|}$$

On pose $a' = ab^{-1}$, ce qui donne :

$$\int_0^{\infty} \frac{b}{b^2 + x'^2} \cos(a'x') dx' = \frac{\pi}{2} e^{-|a|} \stackrel{b>0}{=} \frac{\pi}{2} e^{-|a'|b}$$

qui est la transformée Fourier en cosinus de $\frac{b}{b^2 + x^2}$ (a' étant la fréquence de rotation).

Enfin, on pose $b = y$ et $a' = 2\pi u$

$$\mathcal{F}_1 \left\{ \frac{y}{\pi(x^2 + y^2)} \right\} (u) = \frac{2}{\pi} \int_0^{\infty} \frac{y}{x^2 + y^2} \cos(2\pi xu) dx = e^{-2\pi|u|y}$$

B.4 Noyau de Poisson conjugué

La transformée de Fourier du noyau de Poisson 1D conjugué est donnée pour $y > 0$ par :

$$\mathcal{F}_1 \left\{ \frac{x}{\pi(x^2 + y^2)} I_2 \right\} (u) = \frac{u}{|u|} e^{-2\pi|u|y}$$

En utilisant l'intégrale n°17 de la table d'intégration 19 dans [Bronstein 1993] :

$$\int_0^{\infty} \frac{x \sin(bx)}{a^2 + x^2} dx = \text{signe}(b) \frac{\pi}{2} e^{-|ab|}$$

Et en posant $b = 2\pi u$ et $a = y$, nous obtenons la transformée de Fourier en sinus du noyau de Poisson 1D conjugué :

$$\mathcal{F}_1 \left\{ \frac{x}{\pi(x^2 + y^2)} I_2 \right\} (u) = \text{signe}(u) e^{-2\pi|uy|} \stackrel{y>0}{=} \frac{u}{|u|} e^{-2\pi|u|y}$$

Noyaux et transformée de Fourier dans $\mathbb{R}_{3,0}$

C.1 Transformée de Fourier

Soient f et g des fonctions vectorielles intégrables sur $\Omega \subset \mathbb{R}^2$, alors la transformée de Fourier dans $\mathbb{R}_{3,0}$ et son inverse sont définis par :

$$\mathcal{F}_2\{f(x, y)\}(u, v) = \int_{\Omega} f(x, y) e^{-2\pi I_3(xu+yv)} dx dy$$

$$\mathcal{F}_2^{-1}\{g(u, v)\}(x, y) = \int_{\Omega} g(u, v) e^{2\pi I_3(xu+yv)} du dv$$

avec $I_3 = e_1 e_2 e_3$ pseudo-scalaire de $\mathbb{R}_{3,0}$. Puisque l'espace engendré par $\{1, I_3\}$ est isomorphe à \mathbb{C} , cette transformée de Fourier agit comme la transformée de Fourier classique (voir [Felsberg 2001][Ebling 2005] pour les propriétés qui en découlent).

C.2 Noyau de Riesz

La transformée de Fourier du noyau de Riesz est donnée par :

$$\mathcal{F}_2 \left\{ \frac{x e_1 + y e_2}{2\pi(x^2 + y^2)^{3/2}} e_3 \right\} (u, v) = \frac{u e_1 + v e_2}{\sqrt{u^2 + v^2}} I_2^{-1}$$

D'après la table des transformations d'Hankel dans [Bracewell 1995] :

$$\mathcal{F}_2 \left\{ \frac{1}{\sqrt{x^2 + y^2}} \right\} (u, v) = \frac{1}{\sqrt{u^2 + v^2}}$$

En appliquant le théorème de dérivation de la transformée de Fourier 2D :

$$\mathcal{F}_2 \left\{ -\frac{x e_1 + y e_2}{(x^2 + y^2)^{3/2}} \right\} (u, v) = -2\pi \frac{u e_1 + v e_2}{\sqrt{u^2 + v^2}} I_3^{-1}$$

Enfin, en multipliant par $-\frac{e_3}{2\pi}$ à droite, la transformée de Fourier du noyau de Riesz est obtenu.

C.3 Noyau de Poisson

La transformée de Fourier du noyau de Poisson 2D est donnée pour $s > 0$ par :

$$\mathcal{F}_2 \left\{ \frac{s}{2\pi(x^2 + y^2 + s^2)^{3/2}} \right\} (u, v) = e^{-2\pi\sqrt{u^2+v^2}s}$$

D'après la table des transformations d'Hankel dans [Bracewell 1995] :

$$\mathcal{F}_2 \left\{ (1 + x^2 + y^2)^{-3/2} \right\} (u, v) = 2\pi e^{-2\pi\sqrt{u^2+v^2}}$$

En effectuant la substitution $x' = sx$ et $y' = sy$ où $s > 0$ et en appliquant le théorème affine [Bracewell 1995], nous obtenons :

$$\mathcal{F}_2 \left\{ s^3(s^2 + x'^2 + y'^2)^{-3/2} \right\} (u, v) = s^2 2\pi e^{-2\pi\sqrt{u^2+v^2}s}$$

En multipliant par $(s^2 2\pi)^{-1}$, la transformée de Fourier du noyau de Poisson est obtenue.

C.4 Noyau de Poisson conjugué

La transformée de Fourier du noyau de Poisson 2D conjugué est donnée pour $s > 0$ par :

$$\mathcal{F}_2 \left\{ \frac{xe_2 + ye_2}{2\pi(x^2 + y^2 + s^2)^{3/2}} e_3 \right\} (u, v) = \frac{ue_1 + ve_2}{\sqrt{u^2 + v^2}} I_2^{-1} e^{-2\pi\sqrt{u^2+v^2}s}$$

D'après la table des transformations de Hankel dans [Bracewell 1995] :

$$\mathcal{F}_2 \left\{ (1 + x^2 + y^2)^{-1/2} \right\} (u, v) = (u^2 + v^2)^{-1/2} e^{-2\pi\sqrt{u^2+v^2}}$$

et en substituant $x' = sx$ et $y' = sy$ ($s > 0$)

$$\mathcal{F}_2 \left\{ s(s^2 + x'^2 + y'^2)^{-1/2} \right\} (u, v) = s^2 (s\sqrt{u^2 + v^2})^{-1} e^{-2\pi\sqrt{u^2+v^2}s}$$

d'où

$$\mathcal{F}_2 \left\{ (s^2 + x^2 + y^2)^{-1/2} \right\} (u, v) = (u^2 + v^2)^{-1/2} e^{-2\pi\sqrt{u^2+v^2}s}$$

Le théorème de dérivation nous donne :

$$\mathcal{F}_2 \left\{ -(xe_1 + ye_2)(s^2 + x^2 + y^2)^{-3/2} \right\} (u, v) = -2\pi(ue_1 + ve_2) I_3^{-1} (u^2 + v^2)^{-1/2} e^{-2\pi\sqrt{u^2+v^2}s}$$

En multipliant par $-\frac{e_3}{2\pi}$ par la droite, la transformée de Fourier du noyau de Poisson conjugué est obtenue.

Noyaux et transformée de Fourier dans $\mathbb{R}_{7,0}$.

D.1 Transformée de Fourier

Pour une question de clarté, nous notons $X = (x_1, x_2)$, $U = (u_1, u_2)$. Soient f et g des fonctions vectorielles intégrables sur $\Omega \subset \mathbb{R}^2$, alors la transformée de Fourier et son inverse sont définis par :

$$\mathcal{F}\{f(X)\}(U) = \int_{\Omega} f(X) e^{-2\pi I_7 X \cdot U} dX$$

$$\mathcal{F}^{-1}\{g(U)\}(X) = \int_{\Omega} g(U) e^{2\pi I_7 X \cdot U} dU$$

avec $I_7 = e_1 e_2 e_3 e_4 e_5 e_6 e_7$ pseudo-scalaire de $\mathbb{R}_{7,0}$. Puisque I_7 commute avec tous les éléments de l'algèbre et satisfait $I_7^2 = -1$ alors cette transformée de Fourier peut être assimilée à la transformée de Fourier usuelle utilisant la notation complexe. Un exemple similaire est traité dans [Ebling 2005].

D.2 Transformée de Riesz

La transformée de Fourier du noyau de Riesz est donnée par :

$$\mathcal{F} \left\{ \frac{x_1 e_1 + x_2 e_2}{2\pi(x_1^2 + x_2^2)^{3/2}} e_i \right\} (u_1, u_2) = \frac{u_1 e_1 + u_2 e_2}{(u_1^2 + u_2^2)^{1/2}} I_7 e_i$$

En utilisant la table des transformées de Hankel dans [Bracewell 1995] :

$$\mathcal{F}_2 \{(x_1^2 + x_2^2)^{-1/2}\}(U) = (u_1^2 + u_2^2)^{-1/2}$$

Nous appliquons le théorème de dérivation de la transformée de Fourier :

$$\mathcal{F}_2 \left\{ \frac{x_1 e_1 + x_2 e_2}{(x_1^2 + x_2^2)^{3/2}} \right\} (U) = 2\pi I_7 \frac{u_1 e_1 + u_2 e_2}{(u_1^2 + u_2^2)^{1/2}}$$

et en multipliant par $\frac{e_i}{2\pi}$ à droite, la transformée de Fourier du noyau de Riesz est obtenu.

D.3 Noyau de Poisson

La transformée de Fourier du noyau de Poisson est donnée par :

$$\mathcal{F}\left\{\frac{x_i}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}}\right\}(U) = e^{-2\pi(u_1^2 + u_2^2)^{1/2}x_i}$$

pour $x_i > 0$.

En utilisant la table des transformées de Hankel dans [Bracewell 1995] :

$$\mathcal{F}\{(1 + x_1^2 + x_2^2)^{-3/2}\}(U) = 2\pi e^{-2\pi(u_1^2 + u_2^2)^{1/2}}$$

Nous substituons $r' = x_i(x_1^2 + x_2^2)^{1/2}$, $x_i > 0$ et nous appliquons le théorème affine [Bracewell 1995] ce qui nous amène à :

$$\mathcal{F}\{x_i^3(x_i^2 + r'^2)^{-3/2}\}(U) = x_i^2 2\pi e^{-2\pi(u_1^2 + u_2^2)^{1/2}x_i}$$

En posant $r' = |X|$ et en multipliant par $(x_i^2 2\pi)^{-1}$, la transformée de Fourier du noyau de Poisson est obtenu.

D.4 Noyau de Poisson conjugué

La transformée de Fourier du noyau de Poisson conjugué est donnée par :

$$\begin{aligned} \mathcal{F}\left\{\frac{x_1 e_1 + x_2 e_2}{2\pi(x_1^2 + x_2^2 + x_i^2)^{3/2}} e_i\right\}(U) \\ = \frac{u_1 e_1 + u_2 e_2}{(u_1^2 + u_2^2)^{1/2}} e^{-2\pi(u_1^2 + u_2^2)^{1/2}x_i} I_7 e_i \end{aligned}$$

pour $x_i > 0$.

En utilisant la table des transformées de Hankel dans [Bracewell 1995] :

$$\mathcal{F}\{(1 + x_1^2 + x_2^2)^{-1/2}\}(U) = \frac{e^{-2\pi(u_1^2 + u_2^2)^{1/2}}}{(u_1^2 + u_2^2)^{1/2}}$$

et en posant $r' = x_i(x_1^2 + x_2^2)^{1/2}$ ($x_i > 0$), nous obtenons :

$$\begin{aligned} \mathcal{F}\{x_i(x_i^2 + r'^2)^{-1/2}\}(U) \\ = x_i^2 ((u_1^2 + u_2^2)^{1/2} x_i)^{-1} e^{-2\pi(u_1^2 + u_2^2)^{1/2}x_i} \end{aligned}$$

d'où

$$\mathcal{F}\{(x_i^2 + |X|^2)^{-1/2}\}(U) = |U|^{-1} e^{-2\pi|U|x_i}$$

Le théorème de dérivation nous donne :

$$\begin{aligned} \mathcal{F}\{-(x_1 e_1 + x_2 e_2)(x_i^2 + |X|^2)^{-3/2}\}(U) \\ = -2\pi(u_1 e_1 + u_2 e_2) I_7 |U|^{-1} e^{-2\pi|U|x_i} \end{aligned}$$

et en multipliant par $-\frac{e_i}{2\pi}$ à droite, la transformée de Fourier du noyau de Poisson conjugué est obtenue.

Analyse de la complexité

Les algorithmes décrits dans cette thèse ont été mis en œuvre sans optimisation particulière, aussi cette annexe donne une analyse de leur complexité mais pas de temps d'exécution.

E.1 Signal Monogène Couleur

À échelle τ fixée, la complexité, en temps et espace mémoire, du signal monogène couleur peut être estimée, à partir de l'équation 2.4.3. Dans la suite, nous considérons une image couleur de taille $3 \times M \times M$ et un masque de convolution $K \times K$, typiquement fixé à 11×11 .

f_c est la somme de 5 termes et sa complexité globale en temps est donnée par le terme de complexité maximale, c'est-à-dire A_1 ou A_2 .

Plus précisément, deux convolutions successives avec des masques non séparables sont nécessaires. La complexité est alors celle de la convolution, en $O(K^2M^2)$ si elle est effectuée dans le domaine spatial, et en $O(M \log M)$ si elle est effectuée dans le domaine de Fourier. Le choix entre l'une ou l'autre de ces deux méthodes dépend de la taille du masque, une grande taille de masque conduisant à préférer le domaine de Fourier. Le signal monogène couleur est stocké sous la forme de 5 images de tailles $M \times M$ et la complexité en espace mémoire est $O(M^2)$. Les deux complexités en temps et espace-mémoire sont identiques si elles sont estimées à partir de l'équation 2.4.3 (signal monogène dans $\mathbb{R}_{5,0}$) ou de l'équation 2.13 (signal monogène dans $\mathbb{R}_{9,0}$).

E.2 Représentation Polaire

La représentation polaire d'une image couleur, à partir du signal monogène associé f_c et le vecteur d'analyse V , est calculée par le produit géométrique de ces deux quantités. Dans le cas le plus général ce produit nécessite $2^{2.5}$ ou $2^{2.9}$ opérations selon que le calcul est fait dans $\mathbb{R}_{5,0}$ ou $\mathbb{R}_{9,0}$. Cependant, l'équation 2.4.3 montre que $f_c V$ est un 2-verseur qui n'est donc constitué que d'une composante scalaire et de parties bivectorielles. Par exemple pour $\mathbb{R}_{5,0}$ ce multivecteur possède exactement 11 éléments. Le calcul de la représentation polaire s'effectue donc en temps constant ($O(1)$) pour chaque pixel et en $O(M^2)$ pour l'image. Le stockage de chaque 2-verseur $\chi(V)$, identifié à un nombre complexe (équation 2.4.3), nécessite exactement 2 valeurs réelles (φ_V and A_V) et la complexité en espace-mémoire est en $O(M^2)$ pour l'image.

E.3 Segmentation Multi-échelle

Le procédé de segmentation d'une image couleur $3 \times M \times M$, avec un masque de convolution $K \times K$, nécessite le choix d'un nombre T d'échelles τ . L'étape la plus coûteuse de l'algorithme est le calcul de $\psi(x, y, \tau)$ en chaque pixel et à chaque échelle τ . Pour cela, la représentation polaire est faite pour chacune des T échelles. Au final, la complexité en temps pour l'image est en $O(T K^2 M^2)$ et la complexité en espace-mémoire est en $O(T M^2)$. Les autres étapes de la segmentation sont de coûts équivalents ou moindres.

E.4 Flot Optique Couleur

La complexité du flot optique est celle de l'algorithme original de Bruhn *et al.* [Bruhn 2005] puisque le même algorithme, avec une contrainte modifiée, est appliqué sur l'image φ_V . Les auteurs indiquent un quasi temps réel avec une approche multi-grille.

Algorithme de Hestenes et Sobczyk

L'algèbre de Lie de $\text{Spin}(n)$ est $\mathbb{R}_{n,0}^2$ muni du crochet de Lie :

$$A \times B = AB - BA$$

Comme l'application exponentielle est surjective ($\text{Spin}(n)$ est un groupe de Lie compact connexe), tout spineur peut s'écrire

$$S = \sum_{i=0}^{\infty} \frac{1}{i!} A^i$$

pour un bivecteur A .

D'après Hestenes et Sobczyk [Hestenes 1984], nous savons que tout A dans $\mathbb{R}_{n,0}^2$ est de la forme

$$A = A_1 + A_2 + \cdots + A_m$$

où $m \leq n/2$ et

$$A_j = \|A_j\| a_j b_j, \quad j \in \{1, \dots, m\}$$

avec

$$\{a_1, \dots, a_m, b_1, \dots, b_m\}$$

un ensemble de vecteurs orthonormaux. Alors

$$A_j A_k = A_k A_j = A_k \wedge A_j$$

pour tout $j \neq k$ et

$$A_k^2 = -\|A_k\|^2 < 0$$

Cela signifie que les plans codés par A_k et A_j sont orthogonaux et implique que

$$e^{A_1 + A_2 + \cdots + A_m} = e^{A_{\sigma(1)}} e^{A_{\sigma(2)}} \dots e^{A_{\sigma(m)}}$$

pour tout σ dans le groupe des permutations $\mathfrak{S}(m)$. Mais comme A_i^2 est négatif, nous avons

$$e^{A_i} = \cos(\|A_i\|) + \sin(\|A_i\|) \frac{A_i}{\|A_i\|}$$

La rotation correspondante

$$R_i : x \mapsto e^{-A_i} x e^{A_i}$$

agit dans le plan orienté défini par A_i comme plan de rotation d'angle $2\|A_i\|$. Les vecteurs orthogonaux à A_i sont invariants sous R_i .

Il apparait que tout élément R de $SO(n)$ est une composition de simples rotations commutatives, dans le sens où elles n'ont qu'un plan invariant. Les vecteurs invariants par R sont ceux du sous-espace orthogonal à A . Si $m = n/2$, c'est trivial. La composition précédente n'est pas unique si $\|A_k\| = \|A_j\|$ pour $j \neq k$. Dans ce cas, une infinité de plans sont invariants à gauche par R .

Bibliographie

- [Arulampalam 2002] S. Arulampalam, S. Maskell, N. Gordon et T. Clapp. *A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking*. IEEE Transactions on Signal Processing, vol. 50, no. 2, pages 174–188, 2002. [69](#), [70](#)
- [Babaud 1986] J. Babaud, A.P. Witkin, M. Baudin et R.O. Duda. *Uniqueness of the Gaussian kernel for scale-space filtering*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 1, pages 26–33, 1986. [93](#)
- [Baker 2007] S. Baker, D. Scharstein et J.P. Lewis. *A database and evaluation methodology for optical flow*. In IEEE Proc. Int. Conf. of Computer Vision (ICCV'07), pages 1–8, Rio de Janeiro, Brazil, oct 2007. [50](#)
- [Bar-Shalom 1988] Y. Bar-Shalom et T.E. Fortmann. *Tracking and data association*. Academic Press, 1988. [68](#)
- [Batard 2008] T. Batard et M. Berthier. *Clifford algebras bundles to multidimensional image segmentation*. To appear in Advances in Applied Clifford Algebras. Special Issue of the Proc. of the 8th International Conference on Clifford Algebras and their Applications in Mathematical Physics ICCA8, 2008. [91](#)
- [Batard 2009] T. Batard, C. Saint-Jean et M. Berthier. *A Metric Approach to nD Images Edge Detection with Clifford Algebras*. Journal of Mathematical Imaging and Vision, vol. 33(3), pages 296–312, 2009. [5](#), [91](#)
- [Batard 2010] T. Batard, M. Berthier et C. Saint-Jean. *Clifford Fourier Transform for Color Image Processing*. In E. Bayro-Corrochano et G. Scheuermann Eds, editeurs, Geometric Algebra Computing in Engineering and Computer Science, chapitre 8, pages 135–161. Springer Verlag, 2010. [91](#)
- [Bichot 2005] E. Bichot, L. Mascarilla et P. Courtellemont. *Motion based Segmentation to improve Tracking of Non Rigid Objects*. In Proceedings of the 3rd International Conference on Sciences of Electronic, Technologies of Information and Telecommunications (SETIT'05), Sousse, Tunisie, march 2005. [66](#)
- [Bracewell 1995] R. N. Bracewell. Two-dimensional imaging. Prentice Hall Signal Processing Series, 1995. [97](#), [98](#), [99](#), [100](#)
- [Bronstein 1993] I. Bronstein, K. Semendjajew, G. Musiol et H. Mühlig. Taschenbuch der mathematik. Verlag Harri Deutsch, 1993. [96](#)
- [Bruhn 2005] A. Bruhn et J. Weickert. *Lukas/Kanade meets Horn/Schunk : Combining Local and Global Optic Flow*,. Int. Journal of Computer Vision, vol. 62, no. 3, pages 211–231, 2005. [49](#), [50](#), [51](#), [53](#), [102](#)
- [Bülöw 2001] T. Bülöw et G. Sommer. *Hypercomplex signals - a novel extension of the analytic signal to the multidimensional case*. IEEE Transactions on Signal Processing, vol. 49, no. 11, pages 2844–2852, 2001. [23](#), [28](#)

- [Canny 1986] J. Canny. *A Computational Approach To Edge Detection*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 8, pages 679–714, 1986. 61
- [Cappé 2007] O. Cappé, S. J. Godsill et E. Moulines. *An overview of existing methods and recent advances in sequential Monte Carlo*. In Proceedings of the IEEE, volume 95, pages 899 – 924, 2007. 65
- [Comaniciu 2000] D. Comaniciu, V. Ramesh et P. Meer. *Real-Time Tracking of Non-Rigid Objects using Mean Shift*. In Computer Vision and Pattern Recognition (CVPR'00), volume 2, pages 142–149, Hilton Head Island, South California, 2000. 65, 73
- [Cumani 1991] A. Cumani. *Edge Detection in Multispectral Images*. CVGIP : Graphical Models and Image Processing, vol. 53, no. 1, 1991. 56
- [Demarcq 2009a] G. Demarcq, L. Mascarilla et P. Courtellemont. *The Color Monogenic Signal : a New Framework for Color Image Processing. Application to Color Optical Flow*. In IEEE Proc. Int. Conf. on Image Processing (ICIP'09), 2009. 51
- [Demarcq 2009b] G. Demarcq, L. Mascarilla et P. Courtellemont. *A Metric and Multiscale Color Segmentation using the Color Monogenic Signal*. Lecture Notes in Computer Science, CAIP'09, vol. 5702, pages 906–913, 2009. 57, 63, 65
- [DiZenko 1986] S. DiZenko. *A note on the gradient of a multi-image*. Computer Vision Graphics and Image Processing, vol. 33, no. 1, pages 116–125, 1986. 55
- [Dockstader 2001] S. L. Dockstader et A. M. Tekalp. *On the Tracking of Articulated and Occluded Video Object Motion*. Real-Time Imaging, vol. 7, no. 5, pages 415–432, 2001. 65
- [Duits 2004] R. Duits, L.M.J Florack, J. de Graaf et B.M. ter Haar Romeny. *On the axioms of scale-space theory*. Journal of Mathematical Imaging and Vision, vol. 3, pages 267–298, 2004. 94
- [Ebling 2005] J. Ebling et G. Scheuermann. *Clifford Fourier Transform on Vector Fields*. IEEE Trans. Vis. Comput. Graph., vol. 11, no. 4, pages 469–479, 2005. 97, 99
- [Fairchild 2005] M. Fairchild. *Color appearance models*. John Wiley and Sons, 2005. 31
- [Felsberg 2001] M. Felsberg et G. Sommer. *The Monogenic Signal*. IEEE Trans. on Signal Processing, vol. 49, no. 12, pages 3136–3144, 2001. 5, 24, 41, 44, 45, 91, 94, 95, 97
- [Felsberg 2002] M. Felsberg. *Low-Level Image Processing with the Structure Multi-vector*. PhD thesis, Christian-Albrechts University, Kiel, 2002. 17, 24, 26, 28

- [Felsberg 2004] M. Felsberg et G. Sommer. *The Monogenic Scale-Space : A Unifying Approach to Phase-Based Image Processing in Scale-Space*. Journal of Mathematical Imaging and Vision, vol. 21, pages 5–26, 2004. 93, 94
- [Fleischmann 2008] O. Fleischmann. 2d signal analysis by generalized hilbert transforms. Master's thesis, Christian-Albrechts Universität zu Kiel, 2008. 91
- [Gabor 1946] D. Gabor. *Theory of communication*. Journal of the IEE (London), vol. 93, pages 429–457, 1946. 17
- [Gallier 2000] J.H. Gallier. Geometric methods and applications., volume 38. Springer-Verlag, TAM, 2000. 15
- [Gevers 2006] T. Gevers. Color feature detection : An overview. CRC Press, 2006. 57
- [Gilbert 1991] J.E. Gilbert et M.A.M. Murray. Clifford algebras and dirac operators in harmonic analysis. Cambridge University Press, 1991. 17, 19, 25
- [Gordon 1993] N. Gordon, D. Salmond et A. F. M. Smith. *Novel approach to non-linear and non-Gaussian Bayesian state estimation*. In Proceedings IEE-F, Radar and signal processing, volume 140, pages 107–113, 1993. 70
- [Hahn 1996] S.L. Hahn. Hilbert transforms in signal processing. Artech House in Boston, MA, 1996. 23
- [Harris 1988] C. Harris et M. Stephens. *A combined corner and edge detector*. In Alvey Vision Conference, pages 147–151, 1988. 74
- [Harris 1993] C. Harris. *Tracking with rigid models*. In Active Vision, pages 59–73. MIT Press, 1993. 65
- [Hestenes 1984] D. Hestenes et G. Sobczyk. Clifford algebra to geometric calculus. Reidel, 1984. 103
- [Horn 1981] B.K.P. Horn et B.G. Schunk. *Determining optical flow*. Artificial Intelligence, vol. 17, pages 185–203, 1981. 49
- [Hue 2003] C. Hue. *Méthodes séquentielles de Monte Carlo pour le filtrage non linéaire multi-objets dans un environnement bruité. Applications au pistage multi-cibles et à la trajectographie d'entités dans des séquences d'images 2D*. Thèse de doctorat, Université de Rennes, 2003. 68, 69
- [Iijima 1959] T. Iijima. *Basic Theory of pattern observation*. In Papers of Technical Group on Automata and Automatic Control, IECE, Japan, 1959. 93
- [Iijima 1962] T. Iijima. *Basic theory on normalization of pattern (In case of a typical one-dimensional pattern)*. Bulletin of the Electrotechnical Laboratory, vol. 26, pages 368–388, 1962. 93
- [Isard 1998a] M. Isard. *Visual Motion Analysis by Probabilistic Propagation of Conditional Density*. PhD thesis, Oxford University, 1998. 69
- [Isard 1998b] M. Isard et A. Blake. *Condensation - conditional density propagation for visual tracking*. International Journal of Computer Vision, vol. 29 (1), pages 5–28, 1998. 70

- [Isard 2001] M. Isard et J. MacCormick. *BraMBLe : A Bayesian multiple-blob tracker*. In Proceedings International Conference on Computer Vision (ICCV'01), volume 2, pages 34–41, Vancouver, BC, Canada, 2001. 68
- [Jepson 2003] A.D. Jepson, D.J. Fleet et T.F. El-Maraghi. *Robust Online Appearance Models for Visual Tracking*. IEEE Transaction on Pattern Recognition, vol. 25, no. 10, pages 1296–1311, 2003. 65, 73
- [KaewTrakulPong 2003] P. KaewTrakulPong et R. Bowden. *A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes*. In Image and Vision Computing, volume 21, pages 913–929, 2003. 65
- [Koller-Meier 2001] E. Koller-Meier et F. Ade. *Tracking multiple objects using the Condensation algorithm*. Journal of Robotics and Autonomous Systems, vol. 34, pages 93–105, 2001. 65
- [Lindeberg 1994] T. Lindeberg. *Scale-space theory in computer vision*. Kluwer Academic, 1994. 93
- [Lindeberg 1998] T. Lindeberg. *Feature detection with automatic scale selection*. International Journal of Computer Vision, vol. 30, no. 2, pages 79–116, 1998. 61, 74, 78
- [Lowe 1999] D. G. Lowe. *Object recognition from local scale-invariant features*. Proceedings of the International Conference on Computer Vision (ICCV'99), Toronto, vol. 2, pages 1150–1157, 1999. 78
- [Lowe 2004] D. Lowe. *Distinctive image features from scale-invariant keypoints*. International Journal of Computer Vision, vol. 2, no. 60, pages 91–110, 2004. 80
- [Lucas 1981] B. Lucas et T. Kanade. *An iterative image registration technique with an application to stereo vision*. in Proc. Seventh International Joint Conference on Artificial Intelligence (IJCAI'81), Vancouver, BC, Canada, pages 674–69, 1981. 48, 49
- [Magee 2004] D. R. Magee. *Tracking multiple vehicles using foreground, background and motion models*. Image and Vision Computing, vol. 22, no. 2, pages 143–155, 2004. 65
- [Martin 2001] D. Martin, C. Fowlkes, D. Tal et J. Malik. *A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics*. In proceedings of the Int. Conf. on Computer Vision, vol. vol. 2, pages 416–423, 2001. 57
- [Matas 2002] J. Matas, O. Chum, M. Urban et T. Pajdla. *Robust wide-baseline stereo from maximally stable extremal regions*. In British Machine Vision Conference (BMVC'02), pages 384–393, Cardiff, september 2002. 80
- [Mennesson 2010] J. Mennesson, C. Saint-Jean et L. Mascarilla. *De nouveaux descripteurs de Fourier géométriques pour l'analyse d'images couleur*. In Actes de la conférence RFIA 2010 Reconnaissance des Formes et Intelligence Artificielle, pages 599–606, Caen France, 01 2010. 91

- [Mikolajczyk 2004] K. Mikolajczyk et C. Schmid. *Scale and affine invariant interest point detectors*. International Journal of Computer Vision, vol. 1, no. 60, pages 63–86, 2004. 74, 75, 78
- [Mikolajczyk 2005] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir et L. Van Gool. *A comparison of affine region detectors*. International Journal of Computer Vision, vol. 65, no. 1/2, pages 43–72, 2005. 78, 80, 81, 88, 90
- [Moeslund 2001] T. B. Moeslund et E. Granum. *A Survey of Computer Vision-Based Human Motion Capture*. Computer Vision and Image Understanding (CVIU), vol. 81, no. 3, pages 231–268, 2001. 65
- [Monga 1990] O. Monga, R. Deriche, G. Malandain et J. Cocquerez. *Recursive filtering and edge closing : Two primary tools for 3D edge detection*. Lecture Notes in Computer Sciences, (ECCV'90), vol. 427, pages 56–65, 1990. 61
- [Noble 1988] J. A. Noble. *Finding corners*. Image and Vision Computing, vol. 6, no. 2, pages 121–128, 1988. 74, 75
- [Pai 2004] C.J. Pai, H.R. Tyan, Y.M. Liang, H.Y.M. Liao et S.W. Chen. *Pedestrian detection and tracking at crossroads*. Pattern Recognition, vol. 37, no. 5, pages 1025–1034, 2004. 65
- [Postnikov 1985] M. Postnikov. *Leçon de géométrie : groupes et algèbres de lie*. éditions MIR, 1985. 10, 11
- [Reid 1979] D. Reid. *An algorithm for tracking multiple targets*. IEEE Transactions on Automatic Control, vol. 24, no. 6, pages 843–854, 1979. 68
- [Rittscher 2000] J. Rittscher, J. Kato, S. Joga et A. Blake. *A Probabilistic Background Model for Tracking*. In Proc. European Conf. Computer Vision (ECCV'00), volume 2, pages 336–350, Dublin, Ireland, 2000. 65
- [Ross 2007] D. Ross, J. Lim, R-S. Lin et M-H. Yang. *Incremental Learning for Robust Visual Tracking*. International Journal of Computer Vision (IJCV), vol. Special Issue : Learning for Vision, 2007. 71, 73
- [Sangwine 2007] S. Sangwine et T.A. Ell. *Hypercomplex Fourier Transform of Color Images*. IEEE Transactions on Image Processing, vol. 16, no. 1, pages 22–35, 2007. 5
- [Sapiro 1997] G. Sapiro. *Color Snakes*. Computer Vision and Image Understanding (CVIU), vol. 68, no. 2, pages 407–416, 1997. 60
- [Schmid 2000] C. Schmid, R. Mohr et C. Bauckhage. *Evaluation of interest point detectors*. International Journal of Computer Vision (IJCV), vol. 37, no. 2, pages 151–172, 2000. 78, 80
- [Schulz 2001] D. Schulz, W. Burgard, D. Fox et A. Cremers. *Tracking multiple moving targets with a mobile robot using particle filters and statistical data association*. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'01), pages 1665–1670, Seoul, South Korea, 2001. 65, 68

- [Siebel 2003] N.T. Siebel. *Design and Implementation of People Tracking Algorithms for Visual Surveillance Applications*. PhD thesis, Department of Computer Science, The University of Reading, 2003. 65
- [Stark 1971] H. Stark. *An extension of the Hilbert transform product theorem*. Proc. IEEE, vol. 59, pages 1359–1360, 1971. 23
- [Stein 1971] E. Stein et G. Weiss. *Introduction to fourier analysis on euclidean spaces*. Princeton University Press, 1971. 94
- [Stein 1993] E. M. Stein. *Harmonic analysis*. Princeton University Press, 1993. 19, 25
- [Tschumperlé 2006] D. Tschumperlé. *Fast Anisotropic Smoothing of Multi-Valued Images using Curvature-Preserving PDE's*. International Journal of Computer Vision (IJCV), vol. 68, no. 1, pages 65–82, 2006. 60
- [Tuytelaars 2004] T. Tuytelaars et L. Van Gool. *Matching widely separated views based on affine invariant regions*. International Journal of Computer Vision, vol. 1, no. 59, pages 61–85, 2004. 80
- [Tuytelaars 2008] T. Tuytelaars et K. Mikolajczyk. *Local Invariant Feature Detectors : A Survey*. In Foundations and Trends in Computer Graphics and Vision (CGV), vol. 3, no. 1, pages 1–110, 2008. 74
- [van de Weijer 2006] Joost van de Weijer, Theo Gevers et Andrew D. Bagdanov. *Boosting Color Saliency in Image Feature Detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, pages 150–156, 2006. 88
- [Ville 1948] J. Ville. *Théorie et applications de la notion de signal analytique*. Cables et Transmissions, vol. 2A, pages 61–74, 1948. 17
- [web a] <http://mia.univ-larochelle.fr/content/view/168/131/lang,fr/>. 53
- [web b] <http://www.youtube.com/watch?v=RjrEQaG5jPM>. 53
- [Weickert 1999] J. Weickert, S. Ishikawa et A. Imiya. *Scale-Space as first been proposed in Japan*. Journal of Mathematical Imaging and Vision, vol. 10, pages 237–252, 1999. 93, 94
- [Wietzke 2009] L. Wietzke, G. Sommer et O. Fleischmann. *The Geometry of 2D signals*. In IEEE Proceedings of Computer Vision and Pattern Recognition CVPR'09), Miami, Florida, 2009. 41, 43
- [Wietzke 2010] L. Wietzke et G. Sommer. *The signal multi-vector*. Journal of Mathematical Imaging and Vision, vol. 37, no. 2, pages 132–150, 2010. 30
- [Willis 2009] A. Willis et Y. Sui. *An algebraic model for fast corner detection*. In International Conference on Computer Vision (ICCV'09), Kyoto, Japan, october 2009. 80
- [Zang 2007] D. Zang, L. Wietzke, C. Schmaltz et G. Sommer. *Dense optical flow estimation from the monogenic curvature tensor*. In Proc. Scale-Space and Variational Methods, vol. 4485, pages 239–250, 2007. 50, 51

Résumé : Dans cette thèse, une nouvelle représentation des images couleur basée sur une généralisation du signal analytique est introduite. En utilisant l'analogie entre les conditions de Cauchy-Riemann, qui définissent le caractère holomorphe d'une fonction, et l'équation de Dirac dans l'algèbre de Clifford $\mathbb{R}_{5,0}$, un système d'équations dont la solution est le signal monogène couleur est obtenu. Ce signal est notamment basé sur des noyaux de Riesz ainsi que de Poisson 2D, et une représentation polaire, basée sur un produit géométrique, peut lui être associée. Les applications envisagées reposent majoritairement sur cette représentation polaire et sur les informations de couleur et de structures locales s'y rattachant. Des problématiques liées au flot optique couleur, à la segmentation couleur multi-échelle, au suivi d'objets couleur et à la détection de points d'intérêt sont abordées. En ce qui concerne le flot optique, nous nous intéressons à l'extraction du mouvement d'objets d'une certaine couleur en remplaçant la contrainte de conservation de l'intensité par une contrainte de conservation d'angles. Pour la segmentation, une méthode de détection de contours basée sur de la géométrie différentielle et plus particulièrement sur la première forme fondamentale d'une surface, est proposée afin de déterminer les contours d'objets d'une couleur choisie. Pour le suivi d'objets, nous définissons un nouveau critère de similarité utilisant le produit géométrique que nous insérons dans un filtrage particulière. Enfin, nous resituons la définition du détecteur de Harris dans le cadre de la géométrie différentielle en faisant le lien entre ce dernier et une version "relaxée" du discriminant du polynôme caractéristique de la première forme fondamentale. Ensuite nous proposons une nouvelle version multi-échelle de ce détecteur en traitant le paramètre d'échelle comme une variable d'une variété de dimension 3.

Mots clés : Algèbre de Clifford, signal monogène couleur, flot optique couleur, segmentation couleur, suivi d'objets couleur, points d'intérêt.

The Color Monogenic Signal : theory and applications

Abstract : In this thesis, a novel framework for color image processing is introduced based on the generalization of the analytic signal. Using the analogy between the Cauchy-Riemann conditions and the Dirac equation in the Clifford algebra $\mathbb{R}_{5,0}$, a system of equations which leads to the color monogenic signal is obtained. This latter is based on the Riesz and 2D Poisson kernels, and a polar representation based on the geometric product can be associated to this signal. Some applications using color and local structure information provided by the polar representation are presented. Namely, color optical flow, color segmentation, color object tracking and points of interest are developed. Extraction of optical flow in a chosen color is obtained by replacing the brightness constancy assumption by an angle constancy. Edge detection is based on the first fundamental form from differential geometry in order to segment object in a predefined color. Object tracking application uses a new similarity criterion defined by geometric product of block of vectors. This latter is viewed as the likelihood measure of a particle filter. Last part of the thesis is devoted to the definition of the Harris detector in the framework of differential geometry and a link between this definition and a relaxed version of the characteristic polynomial discriminant of the first fundamental form is given. In this context, a new scale-space detector is provided as the result of handling the scale parameter as a variable in a 3-manifold.

Keywords : Clifford algebra, color monogenic signal, color optical flow, color segmentation, color object tracking, points of interest.
