# Contributions en segmentation statistique d'images et reconnaissance de formes 2D

Stéphane Derrode

▶ **To cite this version:**

Stéphane Derrode. Contributions en segmentation statistique d'images et reconnaissance de formes 2D. Traitement du signal et de l'image [eess.SP]. Université Paul Cézanne - Aix-Marseille III, 2008. tel-00325082

**HAL Id: tel-00325082**
**https://theses.hal.science/tel-00325082**

Submitted on 26 Sep 2008

**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 Paul Cézanne d'Aix-Marseille III**
**Faculté des Sciences et Techniques de Saint Jérôme**

# HABILITATION À DIRIGER DES RECHERCHES

présentée et soutenue publiquement par

**Stéphane DERRODE**

le 29 avril 2008

Discipline : Traitement du Signal et de l'Image
École Doctorale : Physique et Sciences de la Matière (ED 352)

## Contributions en
## Segmentation Statistique d'Images
## et
## Reconnaissance de Formes 2D

**Jury**

**Rapporteurs**

| | |
|---|---|
| J.-P. HATON | Professeur, Institut Universitaire de France |
| M. ACHEROY | Professeur, École Royale Militaire, Belgique |
| E. MOREAU | Professeur, Université du Sud Toulon Var |

**Examinateurs**

| | |
|---|---|
| W. PIECZYNSKI | Professeur, Télécom SudParis |
| J. BLANC-TALON | Direction Générale de l'Armement |
| S. BOURENNANE | Professeur, École Centrale Marseille |

# Remerciements

Les études exposées dans ce mémoire ont été, pour la plupart, réalisées au sein de l'équipe GSM de l'Institut Fresnel, dirigé par *Hugues Giovannini*. Elles n'auraient pu aboutir sans le concours d'un grand nombre de personnes auxquelles je voudrais exprimer ma profonde reconnaissance.

En premier lieu, *Salah Bourennane*, responsable du Groupe Signaux Multidimensionnels, pour la confiance qu'il m'a accordée, les conseils toujours opportuns et l'enthousiasme dont il fait preuve. Qu'il sache combien je lui suis reconnaissant de présenter mes travaux. Je remercie également tous les membres de GSM, *Mireille Guillaume*, *Mouloud Adel*, *Ali Khalighi*, *Thierry Gaidon* ainsi que les thésards, passés et futurs : *Damien*, *Cyril*, *Jean-Michel*, *Julien*, *Nadine*, *Alexis*, *Damien*, *Dong*, *Fan*, *William*, *Simon* et *Florian*.

D'autre part, je souhaite remercier tout particulièrement, *Wojciech Pieczynski*, responsable du département CITI de Télécom SudParis, pour l'honneur qu'il m'a accordé en présidant le jury. Je remercie vivement *Jean-Paul Haton*, Membre de l'Institut Universitaire de France, *Marc Acheroy*, Professeur à l'École Royale Militaire de Belgique, et *Eric Moreau*, Professeur à l'Université du Sud Toulon Var, d'avoir rapporté cette habilitation. Je remercie également *Jacques Blanc-Talon*, responsable de domaine scientifique à la Mission pour la Recherche et l'Innovation Scientifique (Direction Générale à l'Armement), d'avoir accepté d'être membre de mon jury.

Également, je voudrais remercier toutes les personnes avec qui j'ai eu la chance d'avoir d'enrichissants échanges scientifiques ; en particulier, mon ancien directeur de thèse, *Faouzi Ghorbel* de l'ENSI de Tunis, *Grégoire Mercier* de l'ENST-Bretagne, *Jordi Inglada* du CNES et *Emmanuel Trouvé* de l'Université de Savoie.

Je ne saurait oublier les amis, *Caroline*, *Muriel*, *Nicolas* et *Marc-Emmanuel*.

*à Lamia*
*et Sarah*

# Table des matières

**Remerciements** **iii**

**Résumé** **1**

**Abstract** **3**

**1 Synthèse des Travaux** **5**
1.1 Segmentation Statistique d'Images . . . . . 5
1.2 Détection et Reconnaissance de Formes 2D . . . . . 9
1.3 Plan du manuscript . . . . . 11

**2 Segmentation Statistique d'Images** **13**
2.1 Segmentation non supervisée d'images par CMC . . . . . 15
2.1.1 Modélisation probabiliste, parcours de Hilbert et décision bayésienne . . . . 15
2.1.2 Chaîne de Markov cachée . . . . . 17
2.1.3 Estimation non supervisée des paramètres par ECI . . . . . 20
2.2 Segmentation multicomposantes par CMC . . . . . 25
2.2.1 Extension vectorielle des CMCs . . . . . 27
2.2.2 Modèles spécifiés par analyse multivariée . . . . . 28
2.2.3 Modèles spécifiés par copules . . . . . 31
2.3 Deux nouvelles topologies pour les CMCs . . . . . 36
2.3.1 Modèle des CMCs d'ordre supérieur . . . . . 36
2.3.2 Modèle des CMCs couples . . . . . 41
2.4 Segmentation imprécise par CMC floue (thèse de C. Carincotte) . . . . . 48
2.4.1 Chaîne de Markov floue . . . . . 50
2.4.2 Estimation des paramètres du modèle des CMCs floue . . . . . 50
2.4.3 Détection floue de changements . . . . . 54
2.5 Conclusion & Perspectives . . . . . 56

**3 Détection et Reconnaissance de Formes 2D** **59**
3.1 Transformée de Fourier-Mellin et reconnaissance de formes (thèse) . . . . . 60

3.1.1 Trois approximations numériques de la TFMA . . . . . 60
3.1.2 Extraction de familles complètes d'invariants par similitudes . . . . . 62
3.1.3 Estimation des paramètres et détection des symétries . . . . . 64
3.2 Moments complexes et indexation de bases d'images (thèse de R. Mezhoud) . . . . . 66
3.2.1 Moments géométriques et invariants de Hu . . . . . 67
3.2.2 Invariants complets des moments complexes . . . . . 68
3.2.3 Reconstruction à partir des moments complexes . . . . . 70
3.2.4 Indexation de bases d'objets à niveaux de gris . . . . . 70
3.3 *A priori* géométrique pour les *snakes* et détection d'objet (master recherche de M.-A. Chermi) . . . . . 74
3.3.1 Le modèle original des *snakes* . . . . . 75
3.3.2 Introduction d'un *a priori* de forme dans les *snakes* . . . . . 76
3.3.3 Illustration du comportement du *snake* avec *a priori* . . . . . 78
3.3.4 Application à des séquences d'images . . . . . 79
3.4 Mélange non-gaussien et poursuite d'objets dans une vidéo (thèse de W. Ketchantang) 81
3.4.1 Suivi colorimétrique d'un objet . . . . . 82
3.4.2 Reconnaissance de l'iris pour l'authentification des personnes . . . . . 85
3.5 Conclusion & Perspectives . . . . . 88

**4 Conclusion Générale et Projets de Recherche 91**
4.1 Synthèse des résultats et projets en cours . . . . . 91
4.2 Projets de recherche . . . . . 93

**Listes des Figures 95**

**Sigles & Abréviations 99**

**Bibliographie 101**

# Résumé

Ce mémoire retrace les activités de recherche que j'ai développées depuis 9 années dont 7 passées au sein de l'équipe Groupe Signaux Multidimensionnels de l'Institut Fresnel et à l'École Centrale Marseille. Les travaux que je présente explorent certains aspects de la *segmentation statistique d'images* pour des applications en imagerie spatiale et de la *description invariante de formes 2D* pour la reconnaissance d'objets en imagerie vidéo.

Plus précisément, la première partie de ce document expose plusieurs extensions du modèle des chaînes de Markov cachées (CMC). Ces extensions portent sur des modifications soit de la modélisation des données observées avec le modèle de *chaîne de Markov vectorielle* et des données cachées avec le modèle de *chaîne de Markov floue*, soit de la topologie de la chaîne –et donc des hypothèses de dépendance statistique sous-jacentes–, aboutissant aux modèles appelés *chaîne de Markov d'ordre supérieur* et *chaîne de Markov couple*. Ces modèles sont évalués dans le cadre de la segmentation d'images radar et de la détection de changements lors de catastrophes naturelles.

La seconde partie traite de la reconnaissance de formes 2D, avec pour thème centrale l'invariance géométrique. Dans un premier temps nous avons proposé de nouvelles familles complètes de descripteurs de forme invariants aux similitudes issues de la *transformée de Fourier-Mellin* et des *moments complexes*, pour des applications d'indexation de bases d'objets à niveaux de gris. La suite des travaux s'est orientée vers *la détection d'objets* avec l'intégration d'un *a priori* de forme invariant aux similitudes dans le modèle des *snakes* et *la poursuite d'objets d'intérêt* dans les séquences vidéo par un modèle de mélange de couleurs non gaussien.

Le document se conclut avec les perspectives que je compte donner à mes recherches, notamment les projets combinant segmentation d'images et reconnaissance de formes, dans le cadre des images très haute résolution des futurs capteurs optique et radar qui permettent d'accéder à des données sub-métriques.

**Mot-clés :** Segmentation d'images, Reconnaissance de formes, Télédétection, Chaîne de Markov cachées, Mélanges non gaussiens, Transformée de Fourier-Mellin, Contour actifs.

# Abstract

This dissertation follows 9 years of my research activities, including 7 years as an assistant professor at the *École Centrale Marseille* and as a researcher into the Multidimensional Signal Processing Group of *Institut Fresnel* (CNRS UMR 6133). Works which I present explore some aspects of the statistical segmentation of images for applications in space imagery and the invariant description of 2D shapes for object recognition in video imagery.

More precisely, the first part of the manuscript shows several extensions of the Hidden Markov Chain (HMC) model. First, these extensions deal with modifications of observed data with the *vectorial HMC* model and hidden data with the *fuzzy HMC* model. Second, we propose some modifications to the chain topology –and then on the independence assumptions the model relies on–, leading to the *higher-order HMC* and *pairwise HMC* models. These models are evaluated in the frame of earth observation, with Synthetic Aperture Radar image segmentation and bi-dates change detection applications for disaster management.

The second part is devoted to 2D shape recognition, with geometrical invariance as central topic. In a first stage, we propose new complete families of similarity-based invariant descriptors from *Fourier-Mellin transform* and from *complex moments*, with application to gray-level object database indexing. Then, our works turned to the detection of objects with the incorporation of a shape invariant *a priori* into the snakes model and the tracking of a selected object in color video by means of a non Gaussian mixture model of object colors.

This dissertation ends with the prospects of my research, especially those combining image segmentation with shape recognition, in the framework of the very high resolution images which allow sub-metric resolutions for both radar and optical data.

**Keywords :** Image segmentation, Shape recognition, Remote sensing, Hidden Markov chains, Non Gaussian mixtures, Fourier-Mellin transform, Active contours (snakes).

CHAPITRE 1

# Synthèse des Travaux

LE STAGE post-doctoral que j'ai réalisé suite à ma thèse a été l'occasion de découvrir le potentiel des modèles stochastiques, en particulier des modélisations markoviennes, pour la segmentation non supervisée d'images. C'est en effet au cours du projet intitulé « Segmentation d'Images Radar » (*SIR* [120]), réalisé au sein du département *ITI* [1] de l'ENST-BRETAGNE, que j'ai eu l'occasion d'appréhender le modèle des champs de Markov cachés et celui des Chaînes de Markov Cachées (CMCs), dans le cadre de la segmentation d'images Radar à Synthèse d'Ouverture (RSO). Ce travail initial a marqué le début d'une collaboration de recherche fructueuse avec Grégoire Mercier de l'ENST-BRETAGNE et Wojciech Pieczynski de l'INT, collaboration qui ne cesse de se renouveler à l'occasion de divers projets développés en commun (Projets *GET MOSS* [123], *GET DéCA* [125], *GDR-ISIS CACHANT* [68]...).

En parallèle, j'ai poursuivi mes travaux de recherche en reconnaissance de formes 2D, initiés par la thèse que j'ai réalisée au sein de l'équipe *GRIF* de TÉLÉCOM LILLE 1. C'est d'ailleurs dans le cadre de mes relations fructueuses depuis une dizaine d'années avec mon directeur de thèse, Faouzi Ghorbel, que cette thématique s'est développée par le biais de co-encadrements doctoraux dans son équipe de recherche à l'ENSI de Tunis (équipe *GRIFT* du Laboratoire Cristal).

J'ai donc tout naturellement structuré mon rapport en deux parties, l'une portant sur des aspects statistiques pour la segmentation et l'autre sur des aspects géométriques pour la reconnaissance. La suite de cette synthèse souligne mes principales contributions sur ces deux thématiques, depuis mon arrivée en 2001 dans l'équipe *GSM*, dirigée par Salah Bourennane, de l'INSTITUT FRESNEL.

## 1.1 Segmentation Statistique d'Images

La télédétection civile, de part la diversité et le renouvellement des applications qu'elle propose, mais aussi de part la difficulté des problèmes qu'elle pose, est le champ privilégié des applications sur lequel j'ai été amené à confronter les différents modèles markoviens proposés. Cette thématique est particulièrement stimulatrice car bien souvent intransigeante et sans concession pour les méthodes développées ! Nous nous sommes confrontés en particulier à la localisation de nappes d'hydrocarbure sur la surface de la mer, à la détection de changements entre deux prises de vue imageant la même scène terrestre à des instants différents, et à l'analyse des prochaines données THR (Très Haute Résolution), dans le cadre du programme franco/italien *ORFEO* (« *Optical and Radar Federated Earth Observation* ») dont l'objectif est le déploiement d'une flotte de satellites

1. — La signification des sigles et des abréviations utilisés dans ce mémoire est détaillée en page 99.

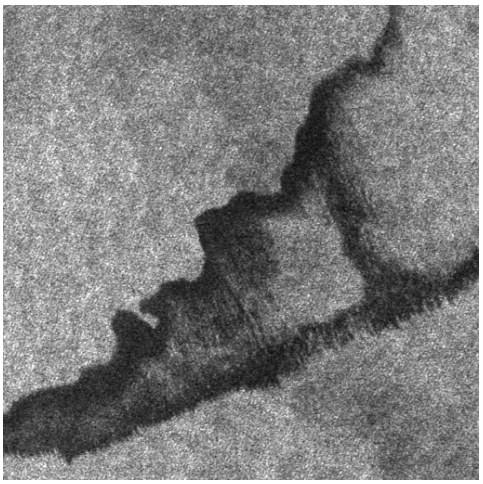


Figure 1.1 – Acquisition ASAR d'Envisat effectuée lors du naufrage du Prestige en 2002. *cf.* section 2.4, page 48.

optique et radar légers d'observation des ressources terrestres.

Dans le projet intitulé « *Multiresolution Oil Slick Segmentation* » (*MOSS* [122]), il s'agissait de détecter et de localiser des nappes d'hydrocarbure sur la surface de la mer provenant de déversements soit accidentels (naufrage), soit volontaires (dégazage). Dans ce contexte, les systèmes d'imagerie RSO procurent un avantage certain par rapport aux systèmes optiques passifs dans la mesure où ils permettent d'obtenir des images de jour comme de nuit, sous toutes conditions météorologiques. Cette caractéristique rend les données radar pleinement opérationnelles pour la protection des côtes maritimes et les opérations de pompage du pétrole où l'urgence d'une intervention prime pour limiter les conséquences humaines et écologiques d'une telle catastrophe. En contre-partie, l'interaction entre les vagues et l'onde électromagnétique émise par le capteur RSO est à l'origine d'effets de textures bruitées qui rendent l'interprétation automatique des images difficile ; difficulté encore accentuée par la présence de nombreux phénomènes océanographiques et météorologiques venant modifié l'aspect des images tels les courants marins, les coups de vent, la salinité des eaux... Les travaux de recherche sur le sujet abondent, comme l'attestent les récents états de l'art [18, 90]. La seule propriété de markovianité ne permet pas de tenir compte de tous ces phénomènes et les cartes de détection de nappe obtenues n'étaient généralement pas satisfaisantes d'un point de vue opérationnel avec la présence de nombreuses fausses alarmes. L'idée développée dans le cadre de ce projet, qui a permis des améliorations notables, consistait à présenter à l'entrée de l'algorithme, non plus l'image originale mais une représentation multiéchelle de cette image (décomposition en ondelettes), sans perte d'information puisqu'inversible [126, 57]. Pour cela, nous avons proposé une extension multidimensionnel du modèle original des CMCs utilisant une Analyse en Composantes Principales (ACP) et une modélisation statistique des lois d'attache aux données par (i) la famille des lois gaussiennes généralisées pour la composante basse fréquence de la décomposition et, (ii) le système de Pearson pour les coefficients d'ondelettes hautes fréquences.

D'une manière plus générale, nous nous sommes intéressés à développer une méthode de segmentation par CMC pour des données vectorielles, qui abondent dans le domaine des images satellitaires, qu'elles soient multi-spectrales, multi-capteurs, multi-dates, multi-échelles ou multi-

résolutions. La principale modification du modèle a consisté à intégrer des lois d'attaches aux données multidimensionnelles —la dimension de ces lois étant égale au nombre de bandes dans l'image vectorielle— pour prendre en compte simultanément et conjointement l'information disponible sur chacune des bandes. Or, dès lors que l'hypothèse de gaussianité n'est plus vérifiée, comme c'est le cas des modalités d'imagerie radar par exemple, l'estimation paramétrique de lois multidimensionnelles est problématique. Dans l'hypothèse d'un faible nombre de bandes, pour éviter le phénomène de Hughes et les problèmes de robustesse des paramètres estimés, nous avons adapté deux outils de l'analyse multi-variée : d'une part, l'ACP qui tient compte de la corrélation entre les composantes et, d'autre part, l'Analyse en Composantes Indépendantes (ACI) qui considère les liens statistiques à tous les ordres, pour des applications de détection de changements [61] et de segmentation d'images multi-spectrales *SPOT* [62]. Une autre solution, proposée dans les travaux de thèse de N. Brunel [21], consiste à utiliser une classe très générale de modèles paramétriques, appelée « copules » [20]. Bien connus dans les domaines de l'économie et de la finance, ces modèles ont retenu moins l'attention de la communauté du traitement du signal et de l'image. Ils présentent pourtant de nombreux avantages comme l'accès directe aux lois marginales permettant, à partir de connaissances physiques sur le capteur par exemple, d'intégrer une information *a priori* sur la forme des marginales. Ainsi, il devient possible de construire une copule gaussienne –qui modélise une structure de dépendance exclusivement avec les moments statistiques d'ordre 2– avec des marginales provenant de différentes familles de lois.

Parallèlement à ce travail, je me suis intéressé à deux extensions du modèle des CMCs qui permettent de relâcher certaines hypothèses classiquement admises, modifiant la topologie de la chaîne et aboutissant à de nouveaux modèles strictement plus généraux que le modèle classique : les « CMCs couples » et les « CMCs d'ordre quelconque ».

➽ Le modèle de *chaîne de Markov couple* (CMCc) considère directement la markovianité du couple des processus observé et caché, et non plus uniquement la markovianité du seul processus caché [146]. Dans ce modèle, le processus caché n'est plus nécessairement de Markov, mais la restauration bayésienne reste possible car, conditionnellement au processus observé, le processus caché est toujours de Markov [111]. Nous avons ainsi proposé une méthode originale d'estimation des paramètres du modèle, basée sur la méthode ECI (« Estimation Conditionnelle Iterative ») [142], pour la restauration non supervisée d'une CMC couple dans un contexte bayésien [66]. Ce modèle permet de prendre en considération des bruits avec des formes plus complexes que dans le cas classique, au détriment d'une complexité algorithmique accrue pour la phase d'estimation des paramètres. Cette extension est à l'origine d'une série de nouveaux modèles tels les modèles de CMC à bruit dépendant [19], les chaînes triplets [150] et les chaînes partiellement de Markov [107], développés dans l'équipe *CITI* de W. Pieczynski.

➽ Par extension du modèle original, nous avons également proposé une méthode d'estimation des paramètres pour le modèle de *chaîne de Markov d'ordre $R$* quelconque (CMCr) [53, 11], c'est-à-dire le modèle pour lequel la mémoire du processus de Markov ne se limite pas au seul état précédent mais aux $R$ états précédents. Pour $R = 2$, ce modèle a été évalué en traitement de la parole ou en reconnaissance de gestes, avec un succès modéré étant donné le faible gain sur les taux d'erreur de classification comparé au modèle classique (pour lequel $R = 1$). L'originalité de notre approche est d'avoir considéré que les lois d'attaches aux données dépendaient elles-mêmes des $S \leq R$ états précédents, favorisant la prise en compte dans le modèle de structures de bruit à plus longue portée. Ce modèle de CMCr a été testé expérimentalement sur des images RSO de

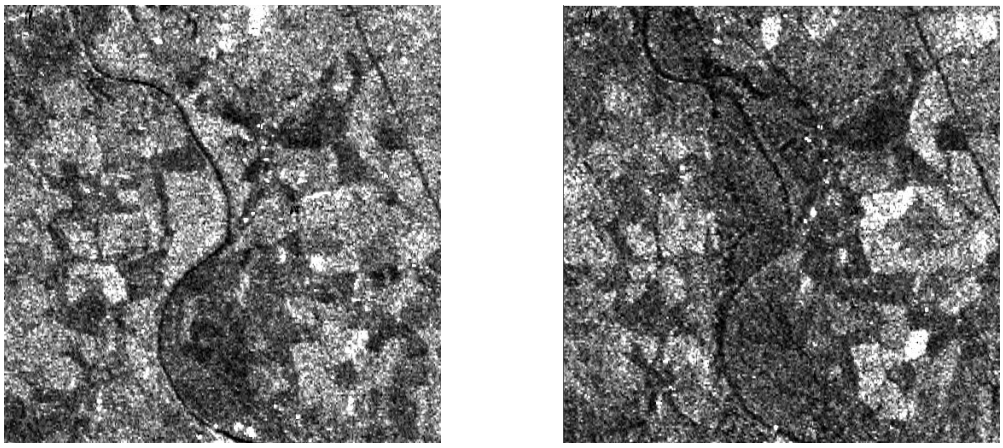


Figure 1.2 – Images ERS-2 PRI d'une région agricole anglaise, avant et pendant une inondation.*cf.* section 2.2, page 25.

nappe d'hydrocarbure, donnant des résultats comparables à ceux obtenus par une décomposition multi-échelle.

Enfin, dans le cadre des travaux de thèse de Cyril Carincotte [28], nous avons proposé une extension floue au modèle des CMCs. L'approche floue permet d'enrichir le modèle statistique (caractérisé par une mesure d'incertitude) en lui ajoutant une mesure d'imprécision. Ainsi, ce modèle prend en considération, conjointement et de manière dissociée, le bruit et le flou dans les observations. Il autorise donc la coexistence de pixels durs et flous dans une même image et permet, par exemple, de caractériser des dégradés de niveaux de gris comme ceux que l'on rencontre à la frontière de zones homogènes (effet de surface partielle bien connu en imagerie médicale mais que l'on rencontre aussi en imagerie satellite). Le modèle a été appliqué avec un certain succès pour détecter des changements entre deux images RSO de la même scène terrestre, le flou permettant de modéliser l'imprécision sur les changements [26].

Les modèles sont présentés dans le contexte de la segmentation statistique d'images. Ainsi, nous utiliserons le terme de 'classe' pour parler d'état et de 'pixel' pour parler des données. Mais leur très grande généralité ne les limite pas à ce domaine. Au contraire leur champ d'applications s'étend à tous les domaines classiques d'utilisation des modèles de Markov – génomique [165], économie, sciences de la santé, ... [139]– et en particulier à la reconnaissance de la parole, précurseur dans la communauté du traitement signal et à l'origine de nombreuses extensions du modèle original [155, 91]. Les algorithmes que nous avons développés ont été confrontés au thème de la détection de changements entre images radar, dans le cadre de divers projets comme le contrat d'études [58] soutenu par le CNES en 2004 dont j'ai assuré la coordination, ou comme le projet incitatif *DéCA* [124] financé par le GET en 2005. Je suis également à l'origine, avec Emmanuel Trouvé du laboratoire LISTIC et G. Mercier, de l'action thématique *CACHANT* [67], soutenue par le GDR ISIS (Thème B) et le CNES. L'objectif de cette action est de fédérer la communauté image intéressée par la problématique de la détection de changements et de l'analyse de données multi-dates, en mettant à disposition des chercheurs intéressés plusieurs jeux de données qui couvrent différentes situations : série d'images multi-capteurs *SPOT/Radarsat*, images présentant des angles d'incidence différents, images prises avec des longueurs d'onde différentes. Cette action est à l'origine de l'organisation des journées « GDR ISIS » à l'ENST des 4 avril et 24 novembre 2006.

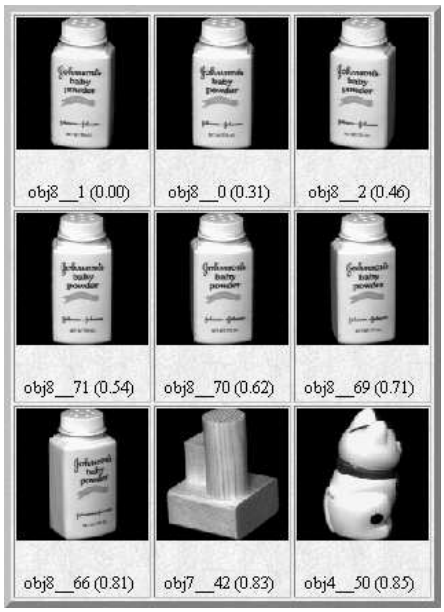


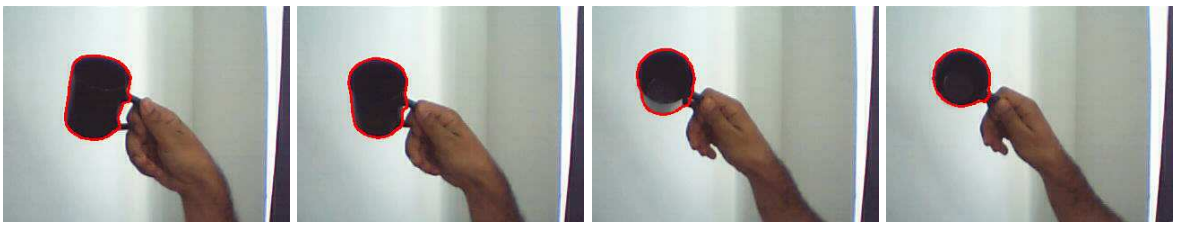

Figure 1.3 – Indexation et recherche par l'exemple. *cf.* section 3.1, page 60.

## 1.2 Détection et Reconnaissance de Formes 2D

Le second volet des travaux que je mène —et historiquement celui qui m'a introduit à la recherche en traitement d'images— concerne les méthodes et applications de la reconnaissance de formes 2D. L'étude conduite dans ma thèse [69] porte sur la description invariante par similitudes de formes planes à niveaux de gris à partir de la Transformée de Fourier-Mellin (*TFM*) et de son prolongement analytique (*TFMA*) [86]. Ces travaux ont donné lieu à deux publications en revue : le premier portant sur la comparaison de trois méthodes d'approximation numérique de la TFMA [55] ; le second traitant de la construction d'une famille complète de descripteurs invariants pour la reconnaissance de formes et, en particulier, pour son application à l'indexation de bases d'objets à niveaux de gris [63]. La poursuite de ces travaux a permis de proposer en 2004 un algorithme original de détection et d'estimation des axes de symétrie par rotation et réflexion d'un objet à niveaux de gris [56].

Par ailleurs, dans le cadre des relations scientifiques soutenues que j'entretiens depuis une dizaine d'années avec les membres du *GRIFT* de l'ENSI [2], j'ai participé au co-encadrement de la thèse de Rim Mezhoud-Khouaja [128], soutenue en novembre 2005. Son travail a consisté à développer une méthode d'indexation de bases d'images à niveaux de gris et de recherche par le contenu à partir d'une nouvelle famille complète de descripteurs invariants par similitudes issue des Moments Complexes (MC). Contrairement à la TFMA qui est issue de l'analyse harmonique sur les groupes, les MC proviennent de la théorie des moments algébriques et ont été introduits par Y.S. Abu-Mostafa et D. Psaltis en 1985 [1]. La famille d'invariants a été construite en appliquant un procédé de normalisation, similaire à celui développé dans le cadre de la TFMA, tirant partie du *théorème du retard* qui exprime le lien entre deux objets de même forme dans le domaine de représentation transformé. Le problème inverse de la reconstruction d'un objet à niveaux de gris

2. — avec notamment l'organisation, depuis sa première édition en 1999, des ateliers « Traitement et Analyse de l'Information - Méthodes et Applications » (*TAIMA*), qui ont lieu les années impaires à Hammamet,



(a) Trame 35 (b) Trame 68 (c) Trame 139 (d) Trame 191

Figure 1.4 – Poursuite d'une tasse. *cf.* section 3.3, page 74.

depuis sa représentation complète par les MC a été étudié par S. Dhabhi dans le cadre de son stage de master recherche PRISM de l'ENSI [88].

Actuellement, les travaux de recherche portant sur la description des formes se déclinent selon deux thèmes : *la détection et la reconnaissance de formes-contour* et *la poursuite d'objets d'intérêt dans les séquences vidéo*.

➽ *Détection et reconnaissance de formes-contour* — Lorsque les formes étudiées sont suffisamment simples ou lorsque l'application envisagée le permet, une description des objets par leur contour externe peut s'avérer satisfaisante pour réaliser des applications de reconnaissance de formes. Dans ce cadre, l'utilisation de certains résultats de l'analyse harmonique sur les groupes au cas de courbes fermées du plan permet de construire une description invariante complète par transformations euclidiennes ou affines [87].

Dans le cadre de son master recherche, M.-A. Chermi [36] a proposé d'utiliser ces descripteurs invariants pour intégrer un *a priori* de forme géométrique dans l'algorithme original des « *snakes* » introduit par M. Kass [99], pour la détection d'un objet dans une scène complexe à l'aide d'un modèle de la forme recherchée. Le nouveau modèle repose sur une force additionnelle, construite à partir de la représentation invariante du contour à chaque itération, qui force le *snake* à évoluer vers la forme d'un objet de référence (c'est l'*a priori* géométrique). Cette méthode améliore, au prix d'une légère augmentation de la complexité, la robustesse du *snake* au bruit et lui permet d'évoluer correctement dans de fortes concavités que des modèles reconnus comme le « *GVF* » ou le « *Balloon* » ne peuvent résoudre [54, 35]. Nous avons appliqué ce modèle au suivi d'un objet en mouvement selon le principe suivant : le *snake* obtenu à la trame courante est utilisé comme *a priori* de forme pour, d'une part, initialiser le *snake* à la trame suivante et, d'autre part, le contraindre vers une forme proche de celle trouvée dans l'image précédente. Ce modèle à l'avantage de s'adapter dynamiquement à l'évolution perçue de la forme de l'objet le long de la séquence [35].

➽ *Poursuite d'objets dans une vidéo couleur* — Que ce soit pour des applications de surveillance, d'authentification ou de vision industrielle, le développement de méthodes de poursuite d'objets d'intérêt dans les séquences vidéo reste une tâche difficile, notamment lorsque la caméra n'est pas dans un environnement confiné ou lorsqu'elle est elle-même en mouvement. Dans ce contexte difficile, nous avons développé une méthode de poursuite intégrant une modélisation originale de la distribution de couleurs de l'objet d'intérêt.

Ainsi, dans le cadre de sa thèse à vocation industrielle [3], William Ketchantang [104] a développé un système complet de reconnaissance d'iris pour des applications d'identification grand public comme, typiquement, le déverrouillage d'un téléphone portable à l'aide de l'appareil photo nu-

3. — soutenue par la région PACA et la société ST Microelectronics

mérique intégré. Pour rendre le système d'acquisition moins contraignant pour l'utilisateur, et finalement plus robuste à l'acquisition, nous comptons exploiter les caméras vidéo qui remplaceront très prochainement l'appareil photo. L'idée consiste à sélectionner les images les plus nettes à la cadence vidéo, en vue de leur utilisation pour la phase de reconnaissance qui s'en suit. Pour cela, a été mis en œuvre une méthode de suivi « temps réel » de la pupille basée sur la représentation de l'iris par un modèle probabiliste de mélange de lois dans l'espace des couleurs, s'inspirant des travaux de S. J. McKenna *et al* [116, 115]. Ainsi, ce n'est pas tant la forme de l'objet qui est suivie mais un mélange de couleurs qui est dynamiquement mis à jour au cours de la séquence. La taille et la position de la pupille sont prédites d'une image à l'autre par un filtre de Kalman pour améliorer la détection à chaque image et réduire les temps de calcul [100, 102]. La sélection des meilleures images, c'est-à-dire celles qui serviront à la reconnaissance, est alors réalisée à l'aide de critères de netteté sur l'iris, pour éviter les images présentant du flou de mouvement, une erreur de focalisation ou un problème d'occlusion partielle de l'iris par les paupières, les cils ou le cadre de l'image.

## 1.3 Plan du manuscript

La suite de ce mémoire reprend de façon détaillée les méthodes développées depuis l'obtention de ma thèse de 3$^{e}$ cycle. Les différentes applications pour lesquelles elles ont été développées et les résultats obtenus servent d'illustration aux algorithmes proposés. D'autre part, j'ai essayé de présenter de façon concise les différents aspects de mes recherches tout en faisant référence aux articles publiés. Ainsi, j'ai choisi de ne donner qu'un nombre limité de références bibliographiques dans ce document.

La **première partie** expose les diverses extensions de la modélisation probabiliste des images par les chaînes de Markov cachées que j'ai proposées. La **seconde partie** décrit mes travaux sur la reconnaissance de formes et la poursuite d'objets dans les séquences. La conclusion de chaque partie présente le bilan des diverses contributions de mon travail de recherche et les perspectives immédiates de ces travaux. La **conclusion générale** est l'occasion d'aborder une discussion sur les projets de recherche communs aux deux thématiques ainsi que les orientations à plus long terme que je souhaite donner à mes recherches.

CHAPITRE 2

# Segmentation Statistique d'Images

L'IDÉE FONDAMENTALE du cadre probabiliste de la segmentation est de considérer la classification d'une image comme l'estimation bayésienne d'un processus aléatoire caché [1] à partir d'un processus observé. Lorsque les observations sont supposées indépendantes entre-elles, nous aboutissons aux méthodes de segmentation qualifiées d'« aveugles » ou de « locales ». Ces méthodes ont connu un large succès avec, en particulier, le développement de méthodes automatique d'estimation des paramètres telle l'algorithme EM (Estimation-Maximisation) [51], rendant la segmentation non supervisée. Néanmoins cette hypothèse d'indépendance est rarement réaliste en segmentation d'images, du fait du lien intrinsèque qui lie les pixels voisins d'une scène observée. Dans le cadre proposé par D. Geman *et al* [84], la dépendance entre pixels est prise en compte à travers le concept fondamental de modèle de Markov, pour lequel le processus caché est supposé être un champs de Markov. Cette approche féconde a engendré de nombreux développements dont un panorama en traitement d'images peut-être consulté dans différents ouvrages de référence [172, 110, 147].

Dans ce contexte, nos travaux sont exclusivement consacrés à l'utilisation des chaînes de Markov pour la segmentation d'images, grâce au parcours de Hilbert-Peano [166] qui permet de transformer une image 2D en un signal 1D. Ce « dépliement » de l'image permet alors d'utiliser les algorithmes de chaînes qui sont particulièrement rapides lorsqu'on les compare à ceux développés pour les champs de Markov. En effet, le succès des Chaînes de Markov Cachées (CMC) dans de nombreux domaines (bio-informatique, finance, communication [76], ...) repose sur la possibilité, d'une part, de calculer (exactement ou approximativement) les probabilités *a posteriori* d'intérêt et, d'autre part, d'estimer les paramètres du modèles avant de réaliser la segmentation dans un contexte bayésien. Cette méthodologie en deux étapes (estimation / segmentation) est à l'origine d'un grand nombre d'extensions et de généralisations du modèle original qui visent soit à relâcher certaines hypothèses initiales, soit à introduire des propriétés particulières aux données cachées. Nous pouvons citer les CMC factorielles et les processus cachés semi-markoviens qui sont présentés, parmi d'autres, dans le livre récent de O. Cappé *et al* [23].

Ce chapitre fait la synthèse des travaux que j'ai menés depuis 2000 en segmentation statistique d'images par CMC. Aussi le **second paragraphe** est-il consacré à la présentation du modèle original (proposé initialement par B. Benmiloud dans le cadre de sa thèse [9]) et à certaines méthodes d'estimation des paramètres pour la segmentation bayésienne non supervisée d'images. Ce paragraphe est l'occasion de fixer les notations et décrit également les hypothèses —markovianité, indépendance conditionnelle, ...— et méthodes de référence —algorithmes de Baum, de Viterbi,

1. — que nous considérerons toujours à valeurs discrètes

. . . — par rapport auxquelles je situe mes différents apports au sujet dans la suite de ce chapitre.

Le **troisième paragraphe** est consacré à une extension vectorielle du modèle permettant de prendre en compte conjointement l'information fournie par différentes sources ; ces sources pouvant être des canaux spectraux, des images prises à des dates différentes ou à des échelles différentes, ou encore par des capteurs différents. Pour prendre en compte simultanément les informations des différentes bandes de l'image multi-composantes, le modèle conduit à modéliser les lois d'attache aux données par des densités multidimensionnelles qui, lorsque l'on sort d'un cadre gaussien, peuvent s'avérer difficile à modéliser. Ainsi, avons-nous proposé deux stratégies pour modéliser les liens statistiques entre les composantes : l'analyse multivariée et la théorie des copules.

Le **quatrième paragraphe** décrit deux extensions au modèle original qui lui sont strictement plus générales, remettant en cause l'hypothèse d'indépendance conditionnelle caractérisant le modèle original. Il s'agit

- d'une part, du modèle des CMCs d'ordre supérieur, qui prend en compte une mémoire d'ordre supérieur tant au niveau de la chaîne de Markov et qu'au niveau des lois d'attache aux données ;
- d'autre part, du modèle des CMCs couples, réalisé en collaboration avec W. Pieczynski, qui considère la markovianité du couple de processus modélisant les observations et les données cachées.

Nous présentons les algorithmes d'estimation associés à ces deux nouvelles topologies et illustrons les méthodes à l'aide d'exemples de segmentation non supervisée d'images simulées et radar à synthèse d'ouverture.

Le **cinquième paragraphe** est consacré aux travaux de thèse de C. Carincotte portant sur une extension flou du modèle des CMCs. Celui-ci permet d'enrichir la mesure d'incertitude du modèle statistique par une mesure d'imprécision. Aussi ce modèle est-il capable de prendre en considération, conjointement et de manière dissociée, le bruit et le flou dans les observations. Ce modèle s'est avéré très bien adapté à la détection de changements entre deux images radar de la même scène mais prises à deux dates différentes. Dans ce contexte, la classification floue permet de modéliser l'imprécision des changements observés, tout en conservant la caractère régularisateur des modèles markoviens.

Enfin, le **dernier paragraphe** fait la synthèse des travaux présentés et décrit succinctement les activités en cours et les projets de recherche en segmentation statistique d'images par modèles de Markov cachés.

**Remarque 2.1** *Pour ne pas surcharger la présentation de ce chapitre, j'ai fait le choix de limiter le nombre de résultats de segmentation et, par-contre, favoriser la description des modèles et des algorithmes d'estimation des paramètres sous-jacens. Tous les modèles présentés ont néanmoins été systématiquement évalués à partir de données simulées (signal ou image) et de données réelles (dans le cadre d'applications de télédétection) ; les études numériques complètes peuvent être consultées dans les articles et mémoires de thèses cités en référence.*

## 2.1 Segmentation non supervisée d'images par CMC

L'objectif de ce chapitre est de présenter, en détails, l'utilisation du modèle des chaînes de Markov cachées dans le contexte de la segmentation non supervisée des images. Il est aussi l'occasion de fixer les notations et de décrire également les hypothèses —markovianité, indépendance conditionnelle, ...— et méthodes de référence —procédure *avant-arrière* de Baum, algorithme de Viterbi, ...— par rapport auxquelles se situe dans la suite mes différentes contributions au sujet.

### 2.1.1 Modélisation probabiliste, parcours de Hilbert et décision bayésienne

Une image $\boldsymbol{y}$ de dimension $L \times C$ est considérée comme une réalisation d'un champ aléatoire $\boldsymbol{Y}$ de même dimension. Chaque pixel $y_{l,c}$ est modélisé par une variable aléatoire $Y_{l,c}$ à valeur dans $\mathbb{R}$ [2]. Segmenter une image, c'est estimer une réalisation $\boldsymbol{x}$ d'une champs aléatoire $\boldsymbol{X}$, dont chaque pixel $x_{l,c}$ est cette fois-ci représenté par une variable $X_{l,c}$ à valeur dans un espace fini de cardinalité $K$, que l'on peut identifier à $\Omega = \{1, \cdots, K\}$.

L'utilisation d'un modèle par chaîne de Markov impose de transformer l'image 2D en un signal 1D. Pour cela, il est possible d'utiliser les « Space Filling Curves » [162]. Une « courbe qui remplit l'espace » est une courbe qui parcourt un espace multi-dimensionnel (inclus dans $\mathbb{Z}^2$ dans notre cas) en passant une et une seule fois par tous les pixels de cet espace. Parmi celles-ci, la courbe de Hilbert se construit de manière récursive, ce qui lui confère une structure auto-similaire, *cf.* figure 2.1. Elle possède une propriété importante : le passage d'un point de la courbe à son voisin (sur la courbe) ne s'accompagne jamais d'un déplacement supérieur à un pixel dans l'image. Ainsi la courbe reste longtemps dans une zone restreinte de l'image. Pour quantifier la préservation du voisinage, J. Chanussot [33] propose le critère suivant : *étant donné le 8-voisinage d'un pixel, le critère proposé calcule le pourcentage de ses voisins qui se retrouvent encore parmi les 8 voisins du pixel après codage sur la courbe*. Comparé à d'autres courbes telles que le parcours de Peano, la courbe de Regazzoni, le « zig-zag scan » et l'ordre lexicographique, le parcours de Hilbert obtient le meilleur score selon ce critère. Ces parcours sont inversibles, ce qui permet, une fois les traitements effectués sur le signal 1D, de reconstruire l'image 2D. Ainsi, quand nous parlerons d'image, il faudra comprendre l'agencement spatial des pixels issu du dépliement de l'image en une chaîne grâce au parcours de Hilbert. Précisons enfin que ce parcours a été étendu par W. Skarbeck [166] pour prendre en compte des images dont les nombres de lignes $L$ et de colonnes $C$ sont paires et non plus des puissances de deux.

Ainsi, une image est-elle considérée comme la réalisation d'un processus stochastique monodimensionnel $\boldsymbol{Y} = (Y_1, \cdots, Y_n, \cdots, Y_N)$, ou $N$ représente le nombre total de pixels dans l'image ($N = L \times C$). Segmenter une image consistera donc à déterminer une réalisation du processus caché $\boldsymbol{X} = (X_1, \cdots, X_n, \cdots, X_N)$. Précisons que la modélisation par champs de Markov, qui utilise directement la représentation matricielle d'une image, est, d'un point de vue topologique, plus satisfaisante que le dépliement de l'image. Cependant, cette propriété de « préservation maximale du voisinage », ainsi que l'aspect pseudo-aléatoire des changements de directions du parcours, sont essentielles pour justifier *a priori* l'intérêt porté au modèle des chaînes de Markov. Il est néanmoins prévisible que certains effets de blocs ne pourront être évités du fait même de la structure fractale de la courbe de Hilbert. Un tel phénomène pourra être observé à la frontière entre les

2. — ou, plus généralement, dans $\mathbb{R}^d$ comme nous le verrons par la suite

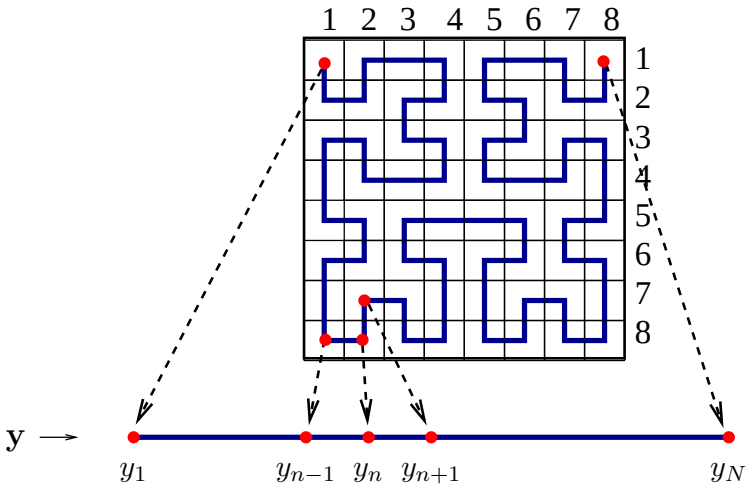



Figure 2.1 – Construction du parcours de Hilbert pour une image $8 \times 8$. À chaque pixel de coordonnées $(l, c)$ est associée une abscisse $n$ indiquant le numéro d'ordre dans le parcours de la courbe, depuis la position initiale $(1, 1) \rightarrow n = 1$ jusqu'à la position finale $(1, 8) \rightarrow n = 64$.

quadrants *haut-gauche* et *haut-droit* définie par les colonnes 4 et 5 de la figure 2.1.

L'estimation de $\boldsymbol{X}$ à partir de $\boldsymbol{Y}$ est réalisée en appliquant une règle de décision bayésienne. Celle-ci est construite à partir d'une fonction de coût $\mathcal{L}$ qui mesure la perte entre $\boldsymbol{x}$ et son estimé $\widehat{\boldsymbol{x}}$. L'estimateur bayésien de $\boldsymbol{X}$ correspondant au coût $L$ est celui qui minimise le coût moyen :

$$\begin{aligned} \widehat{\boldsymbol{x}}^{\mathrm{OPT}}(\boldsymbol{y}) &= \arg\min_{\widehat{\boldsymbol{x}} \in \Omega^N} \sum_{\boldsymbol{x} \in \Omega^N} \mathcal{L}\left(\boldsymbol{X} = \boldsymbol{x}, \widehat{\boldsymbol{x}}\right) p\left(\boldsymbol{X} = \boldsymbol{x} \,|\boldsymbol{Y} = \boldsymbol{y}\right) \\ &= \arg\min_{\widehat{\boldsymbol{x}} \in \Omega^N} E\left[\mathcal{L}\left(\boldsymbol{X} = \boldsymbol{x}, \widehat{\boldsymbol{x}}\right) |\boldsymbol{Y} = \boldsymbol{y}\right]. \end{aligned}$$

En traitement du signal et des images, deux fonctions de pertes sont le plus souvent utilisées, aboutissant aux estimateurs bien connus suivants :

**Maximum *A Posteriori* (MAP) –** La stratégie MAP consiste à pénaliser de la même façon toutes les configurations estimées qui sont différentes de $\boldsymbol{x}$. La fonction de coût correspondante est donnée par

$$\mathcal{L}_1\left(\boldsymbol{x}, \widehat{\boldsymbol{x}}\right) = 1_{\{\boldsymbol{x} \neq \widehat{\boldsymbol{x}}\}},$$

et l'estimateur a pour expression

$$\widehat{\boldsymbol{x}}^{\mathrm{MAP}}(\boldsymbol{y}) = \arg\max_{\boldsymbol{x} \in \Omega^N} p\left(\boldsymbol{X} = \boldsymbol{x} \,|\boldsymbol{Y} = \boldsymbol{y}\right). \tag{2.1}$$

**Maximum *a Posteriori* des Marges (MPM) –** La fonction de coût de l'estimateur du MPM est proportionnelle au nombre de pixels mal classés

$$\mathcal{L}_2\left(\boldsymbol{x}, \widehat{\boldsymbol{x}}\right) = \sum_{n=1}^{N} 1_{x_n \neq \widehat{x}_n},$$

et l'estimateur correspondant a pour expression

$$\forall n \in [1 \cdots N], \quad \widehat{x}_n^{\mathrm{MPM}}(\boldsymbol{y}) = \arg\max_{x_n \in \Omega} p\left(X_n = x_n \,|\boldsymbol{Y} = \boldsymbol{y}\right). \tag{2.2}$$

Comme nous le verrons par la suite, les chaînes de Markov permettent le calcul explicite des solutions bayésiennes de classification, contrairement au modèle des champs de Markov qui fait appel à des méthodes d'approximation de type Monte-Carlo (MPM) ou recuit simulé (MAP). Par la suite, nous utilisons exclusivement l'estimateur MPM qui minimise, en moyenne, le nombre d'observations mal classées [3], contrairement au MAP qui pénalise de la même façon une erreur sur un pixel et une erreur sur plusieurs pixels.

### 2.1.2 Chaîne de Markov cachée

En statistique bayésienne, les densités $p\left(\boldsymbol{x}\right)$ [4], $p\left(\boldsymbol{y}\,|\boldsymbol{x}\right)$ et $p\left(\boldsymbol{y}\right)$ sont respectivement appelées loi *a priori*, vraisemblance et loi d'observation. La loi *a posteriori* $p\left(\boldsymbol{x}\,|\boldsymbol{y}\right)$ se déduit des densités précédentes par la formule de Bayes :

$$\begin{aligned} p\left(\boldsymbol{x}\,|\boldsymbol{y}\right) &= \frac{p\left(\boldsymbol{y}\,|\boldsymbol{x}\right)\, p\left(\boldsymbol{x}\right)}{p\left(\boldsymbol{y}\right)}, \\ &\propto p\left(\boldsymbol{y}\,|\boldsymbol{x}\right)\, p\left(\boldsymbol{x}\right). \end{aligned}$$

Il suffit donc de connaître la loi *a priori* et la vraisemblance pour pouvoir calculer $\widehat{\boldsymbol{x}}^{\mathrm{MAP}}$ et $\widehat{\boldsymbol{x}}^{\mathrm{MPM}}$. Ces deux lois permettent d'introduire deux types de connaissances sur la scène observée. Par la loi *a priori*, il s'agit d'introduire la dépendance spatiale entre les observations. L'hypothèse de markovianité de $\boldsymbol{X}$ est le cas le plus simple qui permet de passer de données indépendantes à des données spatialement dépendantes, tout en restant calculable comme nous le verrons dans la suite. La vraisemblance, quand-à-elle, représente notre connaissance sur la formation de l'image observée et sera différente selon la modalité d'imagerie (optique, radar, ...). Elle est souvent associée à la notion de bruit dans les images, et plus généralement à la dispersion naturelle des niveaux de gris ou à la texture des classes thématiques.

Le modèle classiquement utilisé pour la modélisation des signaux et des images par les chaînes de Markov cachées repose sur les trois hypothèses suivantes :

3. — qui correspond au taux d'erreur de classification

4. — La notation $p\left(\boldsymbol{x}\right)$ sera utilisée pour simplifier l'écriture $p\left(\boldsymbol{X} = \boldsymbol{x}\right)$ dès lors qu'aucune confusion n'est possible. Nous noterons également $p\left(\boldsymbol{x}_{p:q}\right)$ avec $p \leq q$ la probabilité $p\left(x_p, \cdots, x_q\right)$.

**H1** Le processus caché $\boldsymbol{X}$ est une chaîne de Markov (d'ordre 1) :
$p\left(x_n \left| \boldsymbol{x}_{1:n-1}\right.\right) = p\left(x_n \left| x_{n-1}\right.\right)$.

**H2** Les variables aléatoire $Y_n$ sont indépendantes conditionnellement à $\boldsymbol{X}$ :
$p\left(\boldsymbol{y} \left| \boldsymbol{x}\right.\right) = \prod_{n=1}^{N} p\left(y_n \left| \boldsymbol{x}\right.\right)$.

**H3** La distribution de chaque $Y_n$ conditionnellement à $\boldsymbol{X}$ est égale à sa distribution conditionnellement à $X_n$ :
$p\left(y_n \left| \boldsymbol{x}\right.\right) = p\left(y_n \left| x_n\right.\right)$.

Si $\boldsymbol{X}$ est une CMC stationnaire, elle est entièrement définie par sa densité de transition $a(k,l) = p\left(X_2 = l \left| X_1 = k\right.\right)$ et sa loi *a priori* $\pi(k) = p\left(X_1 = k\right)$. En raison de l'hypothèse d'indépendance conditionnelle **H2**, ce modèle est appelé *chaîne de Markov cachée à bruit indépendant* (CMCbi)[5]. La densité du processus joint s'écrit

$$
\begin{aligned}
p\left(\boldsymbol{x},\boldsymbol{y}\right) &= p\left(\boldsymbol{x}\right) \prod_{n=1}^{N} p\left(y_n \left| x_n\right.\right), && (2.3)\\
&= p\left(x_1\right)\, p\left(y_1 \left| x_1\right.\right) \prod_{n=2}^{N} p\left(x_n \left| x_{n-1}\right.\right)\, p\left(y_n \left| x_n\right.\right), && (2.4)\\
&= \pi(x_1)\, f_{x_1}(y_1) \prod_{n=2}^{N} a(x_{n-1},x_n)\, f_{x_n}(y_n), && (2.5)
\end{aligned}
$$

où $f_{x_n}(y_n) = p\left(y_n \left| x_n\right.\right)$ est appelée *loi d'attache aux données*[6] et représente notre connaissance sur la dispersion des niveaux de gris associée à chaque classe. Ces $K$ densités sont généralement considérées gaussiennes mais d'autres formes de distributions (paramétrique ou non) restent possibles. De manière locale, nous pouvons écrire $p\left(y_1,y_2 \left| x_1,x_2\right.\right) = p\left(y_1 \left| x_1\right.\right)\, p\left(y_2 \left| x_2\right.\right)$ et donc

$$
p\left(x_1,x_2,y_1,y_2\right) = a(x_1,x_2)\, f_{x_1}(y_1)\, f_{x_2}(y_2). \tag{2.6}
$$

Le schéma de la figure 2.2 représente la structure de dépendance entre les processus $\boldsymbol{X}$ et $\boldsymbol{Y}$ obtenue grâce aux trois hypothèses énoncées ci-dessus. Les travaux présentés dans la suite de ce chapitre ont pour principal objectif commun de relâcher certaines des contraintes imposées par ces hypothèses, qu'il est parfois difficile de justifier en segmentation d'images. Nous aboutissons alors à des modèles plus généraux mais dont les calculs rapides restent toujours possibles, notamment grâce à la procédure de récurrence *avant-arrière* proposée par Baum [5] que nous rappelons maintenant.

Le principe à l'origine de la procédure de Baum[7] consiste à décomposer $\boldsymbol{Y}$ en deux sous-séquences :

5. — Cette hypothèse de bruit indépendant sera remise en question dans les modèles présentés dans la section 2.3, page 36.

6. — ou parfois *loi d'émission* selon que le modèle est utilisé en traitement du signal ou en traitement des images

7. — Il s'agit ici plus précisément des probabilités normalisées, introduites par Devijver [70] pour pallier à des difficultés numériques inhérentes à la précision des ordinateurs.

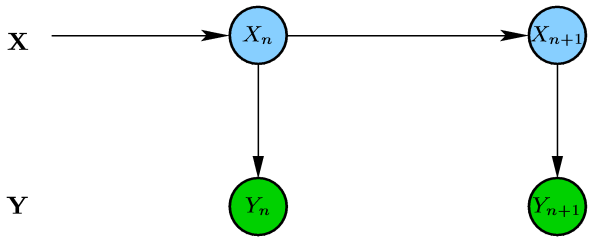



Figure 2.2 – Représentation de la structure de dépendance d'une CMC à bruit indépendant.

- une sous-séquence de l'instant initial 1 à l'instant $n$, où l'on attribue à $X_n$ la classe $k$

$$\alpha_n(k) = p\left(X_n = k \,|\boldsymbol{y}_{1:n}\right) = \frac{1}{S_n}\, p\left(X_n = k, y_n \,\big|\boldsymbol{y}_{1:n-1}\right), \tag{2.7}$$

  avec $S_n$, $n > 1$, le facteur de normalisation suivant :

$$S_n = p\left(y_n \,\big|\boldsymbol{y}_{1:n-1}\right) = \sum_{m=1}^{K} p\left(X_n = m, y_n \,\big|\boldsymbol{y}_{1:n-1}\right).$$

- une sous-séquence de l'instant $n+1$ à l'instant final $N$, où l'on attribue à $X_n$ la classe $k$

$$\beta_n(k) = \frac{p\left(\boldsymbol{y}_{n+1:N} \,|X_n = k\right)}{p\left(\boldsymbol{y}_{n+1:N} \,|\boldsymbol{y}_{1:n}\right)} = \frac{1}{S_{n+1}}\, \frac{p\left(\boldsymbol{y}_{n+1:N} \,|X_n = k\right)}{p\left(\boldsymbol{y}_{n+2:N} \,\big|\boldsymbol{y}_{1:n+1}\right)}. \tag{2.8}$$

Ces probabilités se calculent de façon récursive, à l'aide des formules suivantes

- Initialisation ($n = 1$) :

$$\alpha_1(k) = \frac{\pi(k)\; f_k(y_1)}{\sum_{l=1}^{K} \pi(l)\; f_l(y_1)}, \quad \forall k \in \Omega.$$

- Induction ($1 \leq n < N-1$) :

$$\alpha_{n+1}(k) = \frac{1}{S_{n+1}}\; f_k(y_{n+1}) \sum_{l=1}^{K} \alpha_n(l)\; a(l,k), \quad \forall k \in \Omega. \tag{2.9}$$

pour les probabilités *avant* et

- Initialisation ($n = N$) :

$$\beta_N(k) = \frac{1}{K}, \quad \forall k \in \Omega.$$

- Induction ($1 \leq n < N$) :

$$\beta_n(k) = \frac{1}{S_{n+1}} \sum_{l=1}^{K} a(k,l)\; f_l(y_{n+1})\; \beta_{n+1}(l), \quad \forall k \in \Omega. \tag{2.10}$$

pour les probabilités *arrière*.

Les probabilités marginales *a posteriori* $\xi_n(k) = p\left(X_n = k \,|\boldsymbol{y}\right)$ s'expriment en fonction des pro-

babilités *avant* et *arrière* selon :

$$\begin{aligned}
\xi_n(k) &= \frac{p\left(\boldsymbol{y}\,|X_n=k\right)\,p\left(X_n=k\right)}{p\left(\boldsymbol{y}\right)} \\
&= p\left(\boldsymbol{y}_{1:n}\,|X_n=k\right)\,p\left(\boldsymbol{y}_{n+1:N}\,|\boldsymbol{y}_{1:n},X_n=k\right)\,\frac{p\left(X_n=k\right)}{p\left(\boldsymbol{y}\right)} \\
&= p\left(\boldsymbol{y}_{1:n}\,|X_n=k\right)\,\frac{p\left(\boldsymbol{y}_{n+1:N}\,|X_n=k\right)\,p\left(X_n=k\right)}{p\left(\boldsymbol{y}\right)} \\
&= \frac{p\left(X_n=k\,|\boldsymbol{y}_{1:n}\right)\,p\left(\boldsymbol{y}_{1:n}\right)}{p\left(X_n=k\right)}\,\frac{p\left(\boldsymbol{y}_{n+1:N}\,|X_n=k\right)\,p\left(X_n=k\right)}{p\left(\boldsymbol{y}\right)} \\
&= p\left(X_n=k\,|\boldsymbol{y}_{1:n}\right)\,\frac{p\left(\boldsymbol{y}_{n+1:N}\,|X_n=k\right)}{p\left(\boldsymbol{y}_{n+1:N}\,|\boldsymbol{y}_{1:n}\right)} \\
&= \alpha_n(k)\,\beta_n(k). \qquad (2.11)
\end{aligned}$$

Ainsi, nous obtenons directement l'expression de l'estimateur MPM donné par l'équation (2.2). En ce qui concerne l'estimateur MAP, l'algorithme de Viterbi, qui peut être consulté par exemple dans le papier de L. R. Rabiner [155], propose une solution rapide au problème d'optimisation de l'équation (2.1).

### 2.1.3 Estimation non supervisée des paramètres par ECI

Les densités $p\left(x_n\,|\boldsymbol{y}\right)$ sont inconnues et nous devons les estimer avant de segmenter la série d'observations $\boldsymbol{y}$. Pour cela, en accord avec les équations (2.11), (2.9) et (2.10), il est nécessaire d'estimer la matrice de transition de la chaîne de Markov stationnaire $\boldsymbol{A}=\{a(k,l)\}_{k,l\in\Omega}$ et les paramètres $\Theta=\{\theta_k\}_{k\in\Omega}$ des $K$ lois paramétriques d'attache aux données $\{f(y;\theta_k)\}_{k\in\Omega}$. Pour simplifier l'exposé, et sans perte de généralités, nous considérerons que ces lois sont gaussiennes, auquel cas $\Theta$ est constitué des moyennes et variances de chaque loi. L'utilisation de familles paramétriques différentes (lois gamma, beta, ...) ou de distributions non paramétriques (histogramme, fonctions orthogonales, ...) ne pose pas de difficulté particulière, comme nous le préciserons dans la remarque 2.2, page 23.

Le modèle des CMCs est donc paramétré par $\Phi=\{\boldsymbol{A},\Theta\}$. Dans un contexte de segmentation non supervisée[8], l'estimation des paramètres doit se faire uniquement à partir des observations $\boldsymbol{y}$. Ce problème se situe dans le cadre général de l'estimation de modèles à partir de données manquantes. Depuis les travaux de Baum et Petri [5] à la fin des années 60, la méthode du maximum de vraisemblance est largement utilisée dans le contexte des modèles de Markov cachés. Il n'est reste pas moins que la recherche du maximum global de vraisemblance constitue une réelle difficulté à laquelle des solutions partielles ont été proposées, parmi lesquelles on compte l'algorithme EM (*Espectation/Maximization*), formalisé par A. P. Dempster *et al* [52] et ses nombreuses variantes dont notamment l'algorithme *Stochatic EM* (SEM) proposé par G. Celeux *et al* [30], *cf.* [117].

Le principe de l'algorithme itératif EM consiste à maximiser, par rapport au paramètre $\phi$ à

8. — Précisons ici que nous considérons toujours le nombre de classes connu. Il existe de nombreuses méthodes pour déterminer automatiquement $K$ (*cf. p. ex.* le troisième chapitre de la thèse de J. B. Durand [75]). Parmi eux, les critères informationnels BIC (*Bayesian Information Criterion*) et AIC (*Akaike Information Criterion*) sont largement utilisés dans le contexte des CMCs. Mais ils s'avèrent généralement insatisfaisants pour les images RSO, du fait de la nature du bruit qui les dégrade.

estimer, la fonction de log-vraisemblance définie par

$$Q\left(\phi, \phi^{[p]}\right) = E\left[\log p\left(\boldsymbol{X}, \boldsymbol{Y}, \phi \left| \boldsymbol{y}, \phi^{[p]}\right.\right)\right],$$

à partir d'une valeur initiale $\phi^{[0]}$. L'algorithme EM repose sur l'alternance de deux phases : une phase « Espérance » pour le calcul de la fonction $Q$, et une phase « Maximisation » : $\phi^{[p+1]} = \arg\max_{\phi} Q\left(\phi, \phi^{[p]}\right)$. La suite de paramètres $(\phi^{[p]})_{p\geq 1}$ ainsi construite augmente la vraisemblance des données incomplètes. Dans le cas de mélanges finis, EM fournit des expressions explicites pour la mise à jour des paramètres $\Phi = (\boldsymbol{A}, \Theta)$ ; le calcul de ces formules se base sur les probabilités *avant-arrière*. L'algorithme converge vers un maximum local de la vraisemblance. Il se montre donc très sensible au choix de la valeur initiale des paramètres $\Phi^{[0]}$, ce qui en fait l'une de ses principales faiblesses.

Pour palier à cette limitation, l'algorithme SEM propose de remplacer la phase « Espérance » par une phase « Espérance Stochastique ». Celle-ci consiste, à chaque itération $p$, à simuler des réalisations $\left(\boldsymbol{x}_r^{[p]}\right)_{1\leq r\leq R}$ de $\boldsymbol{X}$ conditionnellement à $\boldsymbol{Y}$, sur la base de la valeur courante du paramètre $\phi^{[p]}$, pour approcher l'espérance conditionnelle. L'estimation de la fonction $Q$ s'écrit alors

$$Q\left(\phi, \phi^{[p]}\right) = \frac{1}{R}\sum_{r=1}^{R} \log p\left(\boldsymbol{X}, \boldsymbol{Y}, \phi \left| \boldsymbol{x}_r^{[p]}, \boldsymbol{y}, \phi^{[p]}\right.\right).$$

Ainsi, l'algorithme consiste-t'il à chercher les paramètres qui maximisent la vraisemblance complétée, grâce aux réalisations *a posteriori*. La perturbation stochastique engendrée par la nouvelle phase de l'algorithme a pour principale conséquence de rendre l'algorithme moins sensible à l'initialisation, en permettant de sortir des « pièges » de maxima locaux de vraisemblance. Dans le cadre des CMCbi, l'algorithme SEM se traduit d'abord par la simulation, à chaque itération $p$, de $\boldsymbol{X}$ selon sa loi *a posteriori* $p\left(\boldsymbol{x} \left| \boldsymbol{y}, \Phi^{[p]}\right.\right)$, qui est celle d'une chaîne de Markov non stationnaire comme nous le verrons par la suite. Ensuite, les paramètres sont estimés grâce au calcul du maximum de vraisemblance au sein de chaque classe.

Dans nos travaux, nous avons considéré la plupart du temps une troisième méthode appelée ECI (*Estimation Conditionnelle Itérative*), introduite par W. Pieczynski dans [143]. ECI est également un algorithme itératif qui génère une suite d'estimations $\phi^{[p]}$ du paramètre $\phi$ et dont le principe consiste à calculer l'espérance d'un estimateur $\hat{\phi}(\boldsymbol{X}, \boldsymbol{Y})$ de $\phi$, défini sur les données complètes conditionnellement aux données observées

$$\phi^{[p+1]} = E\left[\hat{\phi}(\boldsymbol{X}, \boldsymbol{Y}) \left| \boldsymbol{y}, \phi^{[p]}\right.\right]. \tag{2.12}$$

Cette opération est réitérée jusqu'à la stabilisation des estimées, tout comme l'algorithme EM avec lequel il est équivalent pour des lois d'attaches aux données appartenant à la famille des exponentielles [50]. La méthode ECI présente l'avantage d'éviter le calcul et la maximisation des vraisemblances et est, en cela, fondamentalement différente de EM et de ses variantes. Notre choix pour cette méthode relève de la volonté d'évaluer le comportement des estimateurs qui en découlent sur différents modèles de Markov cachés, comme alternative aux méthodes basées sur la vraisemblance.

L'équation (2.12) conduit à deux situations distinctes, selon que l'on cherche à estimer les paramètres du modèle $\boldsymbol{A}$–formules de ré-estimation déterministe, à la EM– ou les paramètres des lois d'attache aux données $\Theta$–formules de ré-estimation stochastique, par simulation en utilisant

la loi des grands nombres, à la SEM– [8]. Nous avons besoin de connaître la loi de $\boldsymbol{X}$ conditionnellement à $\boldsymbol{Y}$, qui est celle d'une chaîne de Markov non homogène de densité de transition $\tilde{\boldsymbol{A}}_n$, où chaque élément $\tilde{a}_n(k,l) = p\left(X_{n+1} = l \,|X_n = k, \boldsymbol{y}\right)$ est calculé selon

$$\begin{aligned} p\left(X_{n+1} = l \,|X_n = k, \boldsymbol{y}\right) &= p\left(X_{n+1} = l \,|\boldsymbol{x}_{1:n-1}, X_n = k, \boldsymbol{y}\right) \\ &= \frac{p\left(\boldsymbol{x}_{1:n-1}, X_n = k, X_{n+1} = l \,|\boldsymbol{y}\right)}{p\left(\boldsymbol{x}_{1:n-1}, X_n = k \,|\boldsymbol{y}\right)} \\ &= \frac{\sum\limits_{x_{n+2},\ldots,x_N \in \Omega} p\left(\boldsymbol{x} \,|\boldsymbol{y}\right)}{\sum\limits_{x_{n+1},\ldots,x_N \in \Omega} p\left(\boldsymbol{x} \,|\boldsymbol{y}\right)} = \frac{\sum\limits_{x_{n+2},\ldots,x_N \in \Omega} p\left(\boldsymbol{x}, \boldsymbol{y}\right)}{\sum\limits_{x_{n+1},\ldots,x_N \in \Omega} p\left(\boldsymbol{x}, \boldsymbol{y}\right)}. \end{aligned}$$

Soit, en utilisant l'expression de la loi du couple (2.3),

$$\tilde{a}_n(k,l) = \frac{a(k,l) f_l(y_{n+1}) \sum\limits_{x_{n+2},\ldots,x_N \in \Omega} a(l,x_{n+2}) f_{x_{n+2}}(y_{n+2}) \ \ldots \ a(x_{N-1},x_N) f_{x_N}(y_N)}{\sum\limits_{x_{n+1},\ldots,x_N \in \Omega} a(k,x_{n+1}) f_{x_{n+1}}(y_{n+1}) \ \ldots \ a(x_{N-1},x_N) f_{x_N}(y_N)}.$$

Enfin, à partir de la définition récursive des probabilités *arrières* de l'eq. (2.10), nous obtenons

$$\tilde{a}_n(k,l) = \frac{a(k,l)\ f_l(y_{n+1})\ \beta_{n+1}(l)}{\sum\limits_{m=1}^{K} a(k,m)\ f_m(y_{n+1})\ \beta_{n+1}(m)}. \tag{2.13}$$

Par un calcul similaire, l'expression des probabilités conjointes *a posteriori* $\psi_n(k,l) = p\left(X_n = k, X_{n+1} = l \,|\boldsymbol{y}\right) = \tilde{a}_n(k,l)\,\xi_n(k)$ s'écrit :

$$\psi_n(k,l) = \frac{\alpha_n(k)\ \beta_{n+1}(l)\ a(k,l)\ f_l(y_{n+1})}{\sum\limits_{i=1}^{K}\sum\limits_{j=1}^{K} \alpha_n(i)\ \beta_{n+1}(j)\ a(i,j)\ f_j(y_{n+1})}. \tag{2.14}$$

- **Ré-estimation des paramètres de la chaîne de Markov–** Selon le principe d'ECI, les probabilités *a priori* et la matrice de transition de la chaîne de Markov stationnaire peuvent être ré-estimées de façon déterministe

$$E\left[\hat{\pi}(k)|\boldsymbol{y}, \Phi^{[p]}\right] \text{ donne :} \qquad \pi^{[p]}(k) = \frac{1}{N}\sum_{n=1}^{N} \xi_n^{[p]}(k), \quad \forall k \in \Omega.$$

$$E\left[\hat{a}(k,l)|\boldsymbol{y}, \Phi^{[p]}\right] \text{ donne :} \qquad a^{[p]}(k,l) = \frac{\sum\limits_{n=1}^{N-1} \psi_n^{[p]}(k,l)}{\sum\limits_{n=1}^{N-1} \xi_n^{[p]}(k)}, \quad \forall k,l \in \Omega. \tag{2.15}$$

- **Ré-estimation des paramètres des lois d'attache aux données–** En ce qui concerne les paramètres des lois d'attache aux données, l'espérance conditionnelle n'est pas calculable directement. On peut cependant l'approcher en calculant la moyenne empirique de plusieurs estimations selon

$$\theta^{[p]} = \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} \hat{\theta}(\boldsymbol{x}^{[p,\ell]}, \boldsymbol{y}),$$

où $\theta$ représente un paramètres a estimer (la moyenne de la densité par exemple) et $\hat{\theta}$ un estimateur de $\theta$ à partir des données complètes $(\boldsymbol{x}^{[p,\ell]}, \boldsymbol{y})$ (la moyenne empirique). La réalisation *a posteriori* $\boldsymbol{x}^{[p,\ell]}$ de $\boldsymbol{X}$ conditionnellement à $\boldsymbol{Y}$ est simulée, à l'itération $p$, grâce à la matrice de transition donnée par l'équation (2.13), en utilisant l'ensemble des paramètres $\Phi^{[p]}$. Dans le contexte du traitement d'images, l'expérience a montré qu'une seule réalisation s'avère généralement suffisante ($\mathcal{L} = 1$).
Ainsi, en connaissant la répartition des données dans chacune des classes à chaque itération $p$, nous pouvons estimer les moyennes $\mu_1^{[p]}(k)$ et variances $\mu_2^{[p]}(k)$ des lois gaussiennes à l'aide d'estimateurs empiriques classiques :

$$\begin{aligned}
\mu_1^{[p]}(k) &= \frac{1}{N_k} \sum_{n=1}^{N} \left( y_n \times \mathbf{1}_{\left[x_n^{[p,1]}=k\right]} \right), \quad \forall k \in \Omega, \\
\mu_2^{[p]}(k) &= \frac{1}{N_k} \sum_{n=1}^{N} \left( \left[ y_n - \mu_1^{[p]}(k) \right] \times \mathbf{1}_{\left[x_n^{[p,1]}=k\right]} \right)^2, \quad \forall k \in \Omega.
\end{aligned} \tag{2.16}$$

avec $N_k = \sum_{n=1}^{N} \mathbf{1}_{\left[x_n^{[p,1]}=k\right]}$ et la fonction $\mathbf{1}_{\left[x_n^{[p,1]}=k\right]} = 1$ si $x_n^{[p,1]} = k$ et 0 sinon.

**Remarque 2.2 (mélange généralisé)** *Dans la plupart des applications, toutes les lois d'attache aux données sont supposées appartenir au même modèle, en particulier à la famille des lois gaussiennes, pour simplifier la mise en œuvre informatique et faciliter les comparaisons. Cependant il existe des applications pour lesquelles les types de lois différent selon les classes. C'est par exemple le cas des images RSO où, par exemple, le signal rétrodiffusé par une banquise ne suit pas les mêmes lois statistiques que celui rétrodiffusé par la houle. Ainsi il est possible d'introduire,* a priori, *des informations physiques sur la formation des signaux reçus dans le modèle en proposant des familles adaptées aux données [49]. Lorsque, dans une même image, le comportement physique des classes est différent (les modèles de chatoiement sont* p. ex. *différents pour une banquise et pour une mer libre de toute glace), nous pouvons avoir recours aux mélanges généralisés pour lesquels plusieurs types de lois sont en concurrence ; la famille la plus adéquate est choisie –selon un critère basé* p. ex. *sur la distance de Kolmogorov [89]– lors de l'étape d'estimation, ce choix étant remis en cause à chaque itération. La même idée a été proposée dans le cadre du système de Pearson [48, 60]*[9] *qui réunit, dans le même formalisme, des familles de lois de support et de formes variées : lois gaussiennes, gamma, beta, Student…*

En conclusion, l'algorithme de segmentation non supervisé basé sur le modèle des CMCbi se déroule en trois étapes, *cf.* figure 2.3 :

❶ D'abord, nous produisons une estimation initiale des paramètres du modèle $\Phi^{[0]}$ à partir des

9. — Une description détaillée du système et de son implémentation en langage C grâce à la librairie GSL (`http://www.gnu.org/software/gsl/`) est disponible auprès de l'auteur.

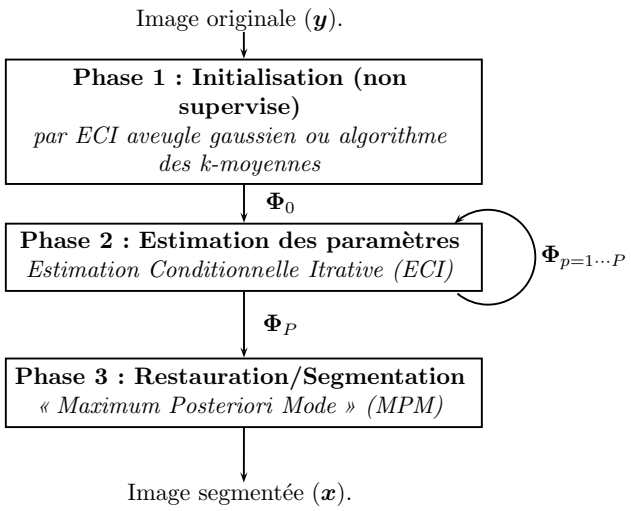



Figure 2.3 – Squelette général de l'algorithme de segmentation non supervisée.

données complétées $\left(\boldsymbol{y}, \boldsymbol{x}^{[0]}\right)$. La simulation $\boldsymbol{x}^{[0]}$ est obtenue à l'aide d'une méthode de segmentation non supervisée et rapide, telle l'algorithme des k-moyennes ou un classifieur non supervisé aveugle.

❷ Ensuite, nous estimons les paramètres du modèle selon le principe itératif d'ECI (ou un autre !). Les paramètres markoviens sont calculés à l'aide des expressions analytiques données dans l'équation (2.15), sur la base des probabilités *avant* (2.7) et *arrière* (2.8). Les paramètres des lois d'observations sont estimés à l'aide d'une ou plusieurs réalisations de $\boldsymbol{x}^{[p]}$ générées selon l'équation (2.13) et les expressions de l'équation (2.16) dans le cas d'un mélange gaussien. Les itérations sont stoppées lorsque les paramètres se stabilisent ($\Phi^{[P]}$).

❸ Enfin, sur la base des paramètres estimés, nous pouvons estimer les solutions bayésiennes de classification généralement retenues en traitement d'images : la solution MPM est obtenue en utilisant l'équation (2.11), alors que la solution MAP est obtenue par le biais de l'algorithme de Viterbi (*cf.* [155], non détaillé ici).

L'ensemble des modèles présentés dans cette partie du manuscript relève d'une méthodologie similaire, tant du point de vue de l'estimation que de la segmentation, ce qui justifie la présentation détaillée que nous venons de faire dans le cas des CMCs. Dans la suite, j'insisterai sur les particularités et l'originalité des nouveaux modèles.

**Illustration : Segmentation d'images RSO de mer glacée –** La figure 2.4 présente des résultats de segmentation obtenus sur une image de mer glacée (a), acquise en 1996 par le satellite RadarSat. L'objectif de la segmentation est d'aider à déterminer un itinéraire libre de glace pour la circulation des navires marchands. Il est en effet difficile de prévoir, dans une telle situation, l'apparition et le mouvement d'iceberg, et les spécialistes ont de plus en plus souvent recours aux méthodes de télédétection. Nous avons choisi de segmenter l'image en quatre classes sensées représenter différentes densités de glace : de la mer libre aux compacts icebergs. Le résultat de la segmentation est présenté dans le figure (d) avec l'initialisation des paramètres à l'aide d'une segmentation de l'image par l'algorithme des k-moyennes (b) et avec le système de Pearson pour représenter les lois d'observation (mélange généralisé). Les figures (e) et (f) montrent respectivement les lois et le mélange estimés après 70 itérations d'ECI. On peut remarquer que seule la loi associée à la queue

de l'histogramme et spécialisée sur l'iceberg présente une forme particulière, les autres lois étant de forme proche d'une gaussienne même si de légères dissymétries peuvent être observées.

Les images (b) et (c) présentent les résultats de segmentation de deux algorithmes non supervisés : l'algorithme déterministe bien connu des k-moyennes et un algorithme statistique aveugle pour lequel les paramètres ont été estimés par une adaptation de ECI à ce contexte. Cet exemple illustre clairement l'intérêt de la prise en compte d'un voisinage pour la classification puisque seule la segmentation issue du modèle des CMCs procure des zones homogènes, rendant la classification exploitable.

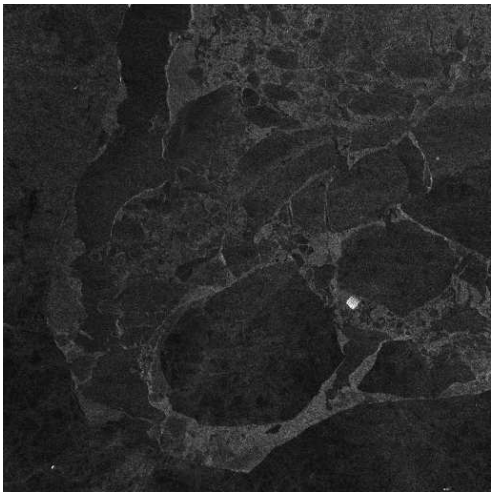
## 2.2 Segmentation multicomposantes par CMC

Dans le cadre des données de télédétection, les informations obtenues sur une scène terrestre ne se présente plus nécessairement sous la forme d'une seule image radiométrique, mais comme c'est de plus en plus le cas, sous forme de plusieurs bandes. Dans ce cadre, le problème consiste à concevoir des algorithmes de segmentation pour prendre en compte plusieurs sources d'information simultanément ; ces sources pouvant être des canaux spectraux (*images couleur* ou *multi-spectrales*), des images d'une même scène prises à des dates différentes (*multi-temporelles*) ou à des échelles différentes (*multi-échelles*), ou encore par des capteurs (*multi-capteurs*) ou à des polarisations différentes (*multi-polarisations*). Chaque composante de l'image met en exergue différentes caractéristiques de la scène observée, et l'objectif des travaux menés dans ce contexte est de prendre en compte l'ensemble de l'information disponible sur chacune des bandes de l'image dans la modélisation des CMCs.

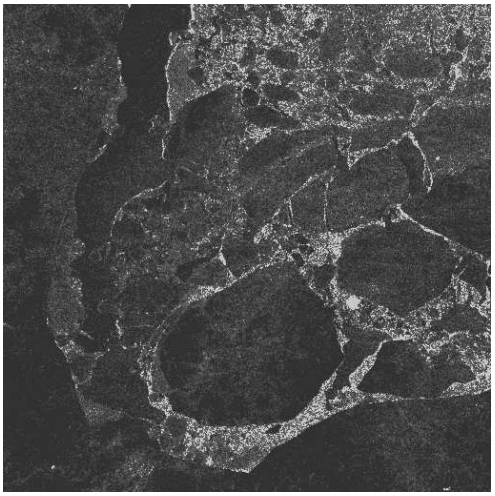
Ainsi se pose le problème de l'extension vectorielle du modèle des CMCs scalaires présenté précédemment. Nous sommes confrontés à un problème d'estimation de mélange multivarié, la dimension du mélange étant égale au nombres de composantes de l'image vectorielle. Ce problème, particulièrement difficile dans le cas général, se simplifie dans le cadre de la méthode d'estimation que nous avons considérée. En effet, comme l'indique les formules de ré-estimation des lois d'attache aux données de la page 23, la forme des lois n'a pas d'impact sur les formules d'estimation des autres paramètres du modèle car l'estimation est réalisée, à chaque itération $p$, à partir des données $\left(\boldsymbol{y}, \boldsymbol{x}^{[p]}\right)$ complétées par simulation. Ainsi, ce problème se réduit à l'estimation, à chaque itération de ECI, de la loi d'attache aux données multidimensionnelle associée à chaque classe du mélange. Nous pouvons donc construire des algorithmes de segmentation de données multivariées modulables en fonction de la manière dont ces lois sont modélisées et estimées [10].

Selon le type de données à traiter et les hypothèses concernant les fluctuations au sein de chaque classe, il est nécessaire de choisir un modèle paramétrique pertinent pour décrire la physique du phénomène observé. La plupart du temps, la loi normale est utilisée grâce à laquelle la notion de groupe est caractérisée par un vecteur moyen ($\boldsymbol{\mu}$) et une matrice de dispersion ($\boldsymbol{\rho}$). Cependant, la dispersion des données de certaines modalités d'acquisition s'écarte significativement de la loi normale et les modèles doivent en tenir compte pour garantir une bonne adéquation aux données. Il existe un grand nombre de modèles paramétriques multi-variés, dont notamment la famille des lois exponentielles (dont font partie les gaussiennes, gamma, gamma inverse, ...). No-

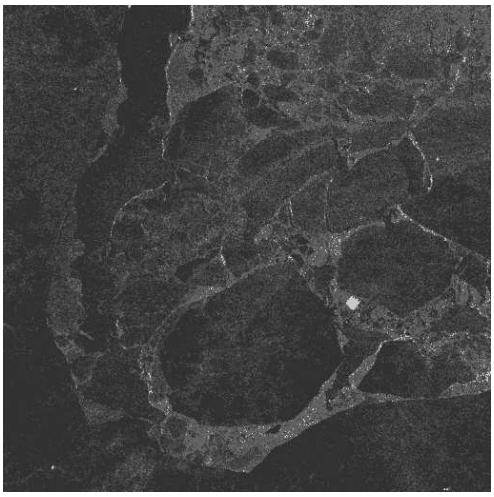
10. — Nous rappelons que le modèle paramétrique associé aux lois d'attache aux données, tout comme le nombre de classes dans l'image, est considéré comme une connaissance *a priori*. L'adéquation des lois d'émission dans les CMCs est examiné entre autres dans [112].



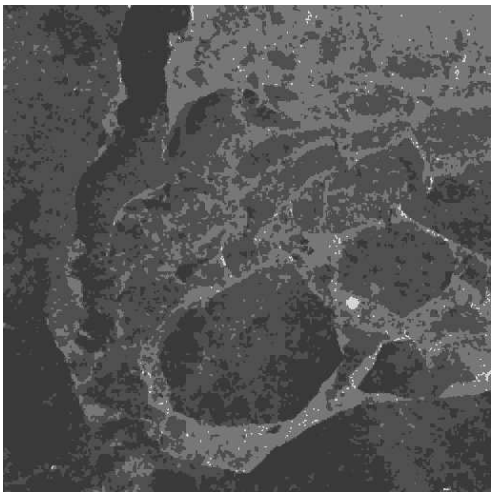
(a) Image originale



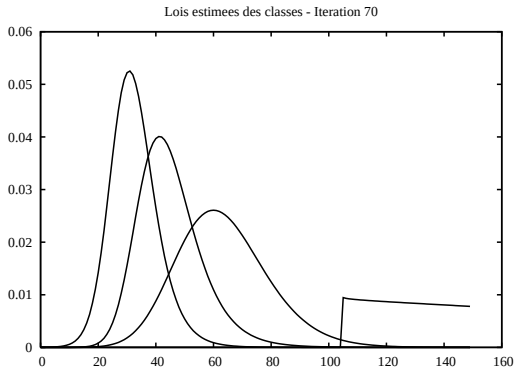
(b) Segmentation par k-moyennes



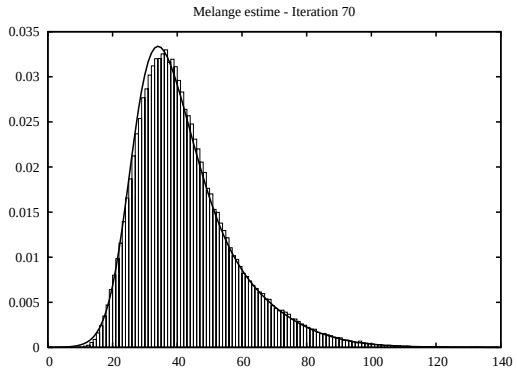
(c) Segmentation par ECI/MPM aveugle



(d) Segmentation par ECI/MPM CMCbi




(e) Lois estimées




(f) Mélange estimé superposé à l'histogramme

Figure 2.4 – Résultats de segmentation d'une image RadarSat de mer glacée, en mode « ScanWide » (*cf.* texte). ©IFRTP - $512 \times 512$ (24 octobre 1996 - Mer Dumont d'Urville, large de la Terre d'Adélie). Résultats extraits de [60].

tons que le livre de S. Kotz *et al* [106] recense une dizaine d'expressions différentes pour les lois gamma multivariées, dont certaines seulement sont à marges gamma. L'exemple du traitement simultané d'images de télédétection provenant de capteurs différents (optique/radar *p. ex.* ) montre que la nature et la forme des distributions de chaque classe dans les différentes bandes peuvent varier.

Nous souhaitons donc disposer de modèles généraux de lois multivariées qui permettent d'intégrer des connaissances *a priori* sur la loi de chaque composante. Pour cela, nous avons étudié deux stratégies très différentes dont nous présentons les avantages et inconvénients : (i) la première est basée sur l'analyse de données multivariées et repose sur la transformation des données en composantes principales ou indépendantes (*ACP/ACI*) [126, 57], et (ii) la seconde relève de modèles paramétriques très généraux appelés *copule* [20]. Nous commençons au préalable par situer le contexte et préciser les notations.

**Remarque 2.3 (phénomène de Hughes)** *La classification d'images multidimensionnelles souffre d'un problème majeur connu sous le nom de* phénomène de Hughes. *En effet, malgré l'information supplémentaire apportée par l'ajout d'une dimension (et donc l'amélioration de la séparabilité des classes), les performances de classification se dégradent (à nombre d'échantillons constant) à cause d'une diminution de la précision de l'estimation statistique des paramètres. De plus la relative complexité des modèles que nous manipulons (en terme de charge et de temps d'exécution) impose que les nombres de classes et de bandes soient « raisonnables ». Il ne s'agira donc pas de vouloir classer une image hyper-spectrale de 288 bandes avec ces modèles, sans prévoir au préalable une sélection des bandes pertinentes pour l'application ou une réduction de dimension adaptée aux données.*

### 2.2.1 Extension vectorielle des CMCs

En appliquant le parcours de Hilbert sur chacune des $M$ images, nous obtenons $M$ chaînes de $N$ données réelles, *cf.* figure 2.5(a). Pour chaque composante $m$, $1 \leqslant m \leqslant M$, la séquence de données est notée $\boldsymbol{y}^m = (y_1^m, \ldots, y_N^m)^t$. Pour chaque indice $n$, $1 \leqslant n \leqslant N$, le vecteur de $M$ composantes ainsi constitué est noté $\vec{y}_n = \left(y_n^1, \ldots, y_n^M\right)^t$. Finalement, l'ensemble des séquences est noté $\vec{\boldsymbol{y}} = \left(\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M\right)^t = (\vec{y}_1, \ldots, \vec{y}_N)$. L'approche statistique consiste à considérer que chaque observation $\vec{y}_n$, à valeurs dans $\mathbb{R}^M$, est la réalisation d'un vecteur aléatoire $\vec{Y}_n$, qu'il convient d'associer à l'une des $K$ classes. Les hypothèses de classification énoncées dans le paragraphe 2.1.2, page 17 restent valables, en remplaçant $y_n$ par $\vec{y}_n$. Il est important de noter qu'aucune hypothèse n'est faite quand à la structure de dépendance des variables $Y_n^m$ du vecteur aléatoire $\vec{Y}_n$ conditionnellement à $X_n$, ce qui rend possible la prise en compte des liens statistiques entre les bandes, *cf.* figure 2.5(b).

Pour chaque itération $p$ de l'algorithme ECI, nous sommes confrontés à l'estimation de $K$ lois $\{f_k(.;\theta)\}_{k\in\Omega}$ à $M$ dimensions. Pour chaque loi $k$, le principe est le suivant : à partir des données complétées par simulation $\left(\vec{\boldsymbol{y}}, \boldsymbol{x}^{[p]}\right)$, nous isolons les données de $\vec{\boldsymbol{y}}$ qui ont été affectées à la classe $k$ et nous les notons $\vec{\boldsymbol{z}}$ (de dimension $N_k$). Ainsi le principal problème de la segmentation statistique d'images vectorielle est le choix d'un modèle statistique adapté pour décrire les liens entre les pixels d'une même classe.

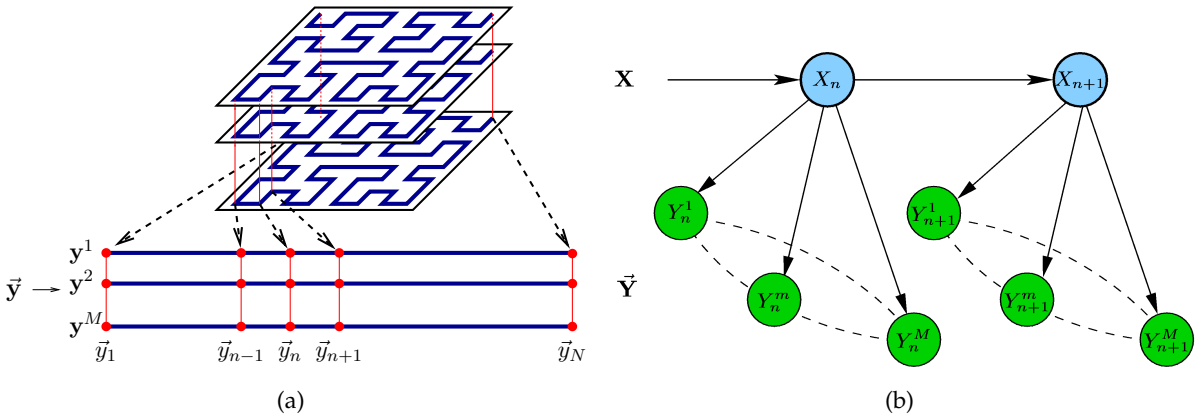



Figure 2.5 – À gauche : description des données vectorielles par le parcours de Hilbert (3 bandes). À droite : représentation de la structure de dépendance d'une CMC vectorielle (3 bandes).

### 2.2.2 Modèles spécifiés par analyse multivariée

Le point de vue de l'analyse multivariée consiste à décomposer l'estimation d'une loi $M$-dimensionnelle $f$ en l'estimation de $M$ lois monodimensionnelles $(f^m)_{1\leq m\leq M}$. Plusieurs stratégies sont possibles, en fonction des hypothèses faites sur les liens statistiques entre les données.

- Si les données sont supposées indépendantes [89], alors $f$ est simplement le produit direct de $M$ densités $f^m$ définies sur $\mathbb{R}$ :

$$f\left(\vec{z}_n\right) = \prod_{m=1}^{M} f^m\left(z_n^m\right). \tag{2.17}$$

  Cependant, cette hypothèse d'indépendance apparaît comme peu justifiée dans le contexte de la segmentation d'images multicomposantes.

- L'ACI permet de projeter les données $\vec{z}_n$ dans un nouvel espace où les composantes sont statistiquement indépendantes [24]. La stratégie sous-jacente utilise un résultat issu du théorème de la limite centrale : *la distribution de la somme de $N$ variables aléatoires indépendantes tend vers une gaussienne quand $N$ tend vers l'infini*, et plus précisément son corollaire : *la somme de deux variables aléatoires centrées est plus proche d'une gaussienne que n'importe laquelle des deux variables initiales*. Aussi l'ACI revient-elle à maximiser un critère de non gaussianité sur les composantes projetées de $\vec{z}$ [40]. Malheureusement la recherche de maximum de non-gaussianité engendre des types de distribution souvent multimodales. Il devient alors très difficile de considérer que ces lois multimodales soient caractéristiques de la distribution des observations restreintes à une seule classe !

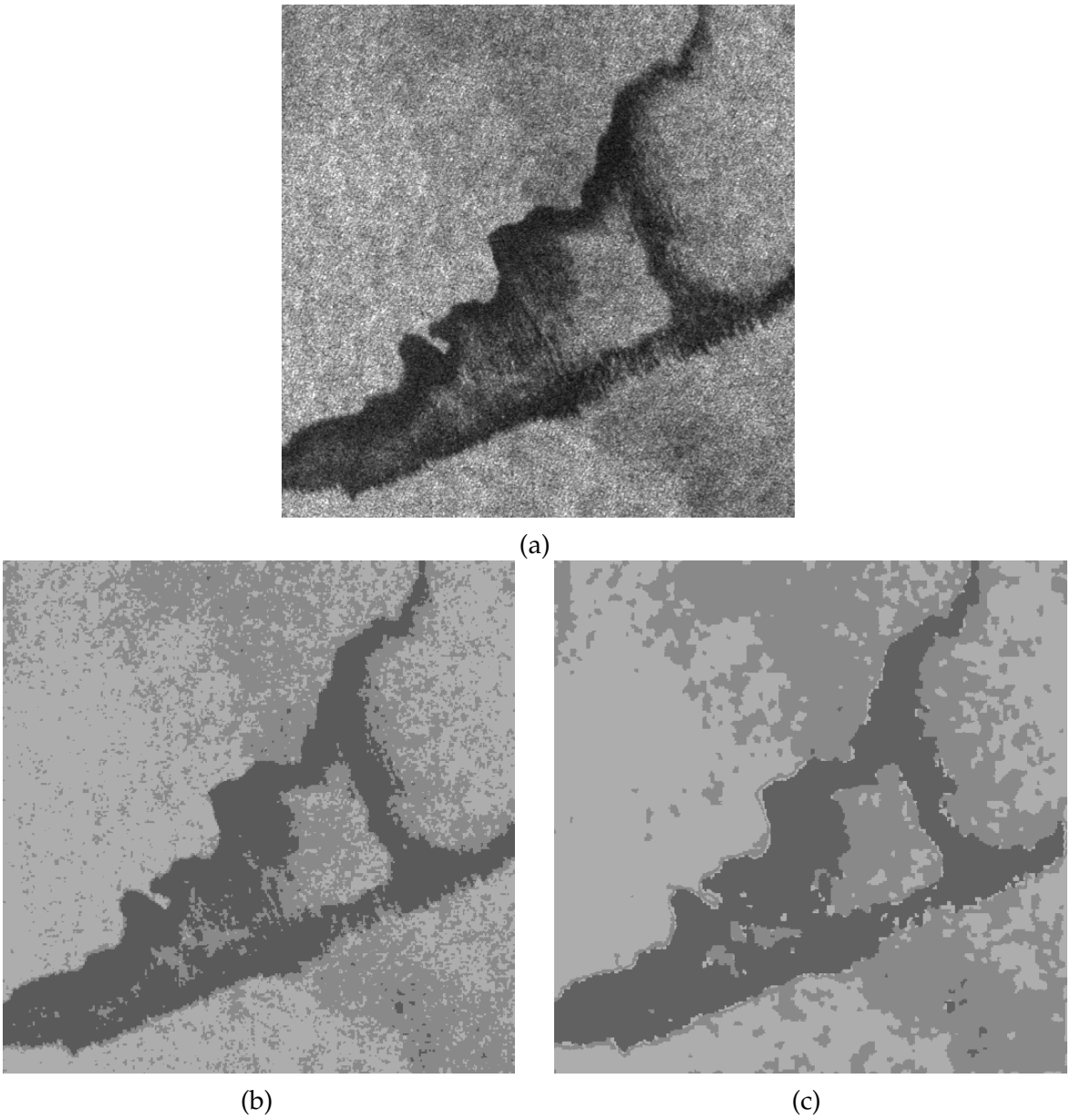
- Une solution sous-optimale, mais mieux adaptée dans la pratique, consiste à utiliser l'ACP. Le principe de l'algorithme est de décorréler les données originales pour revenir à la forme décrite par l'équation (2.17), *cf.* [149, 154]. S'il existe une matrice $\boldsymbol{B}$ telle que les données $\vec{t}_n = \boldsymbol{B}\vec{z}_n$ sont décorrélées, alors la matrice de covariance des données $\vec{t}_n$ est diagonale (matrice unité par exemple). Ainsi, nous pouvons écrire $\boldsymbol{B}\ \boldsymbol{\rho}_{\vec{\boldsymbol{z}}}\ \boldsymbol{B}^t = \boldsymbol{I}$, avec $\boldsymbol{\rho}_{\vec{\boldsymbol{z}}}$ la matrice de covariance des données $\vec{\boldsymbol{z}}$. En utilisant la décomposition de Cholesky (la matrice $\boldsymbol{\rho}_{\vec{\boldsymbol{z}}}$ est symétrique définie positive), une matrice $\boldsymbol{B}$ solution de cette équation s'écrit $\boldsymbol{B}^t = (\text{Cholesky}\,(\boldsymbol{\rho}_{\vec{\boldsymbol{z}}}))^{-1}$.
  On vérifie aisément que les variances des données projetées sont unitaires et que leurs covariances sont nulles. Soit $g^m(\cdot)$, $1 \leqslant m \leqslant M$, les lois monodimensionnelles estimées grâce aux données décorrélées $\vec{\boldsymbol{t}}_n$. Nous obtenons :

$$f\,(\vec{\boldsymbol{z}}_n) = |\det \boldsymbol{B}| \prod_{m=1}^{M} g^m\,(t_n^m). \tag{2.18}$$

Aucune hypothèse sur la forme des lois des composantes principales n'est nécessaire. La méthode peut donc très bien s'appliquer dans un cadre gaussien (il faut alors estimer les moyennes et variances des lois), ou dans le cadre d'un mélange généralisé où les lois peuvent appartenir à des familles différentes. C'est le cas des données multi-échelles, utilisées pour détecter et caractériser les nappes d'hydrocarbure, que nous présentons maintenant comme illustration de la méthode.

**Illustration : Détection de nappes d'hydrocarbure dans les images RSO –** Dans le cadre du projet MOSS [122], nous nous sommes intéressés à la détection et à la caractérisation des nappes d'hydrocarbure à partir d'images Radar à Synthèse d'Ouverture (RSO). La viscosité de l'hydrocarbure atténue sensiblement la rugosité de surface qui participe majoritairement à la rétro-diffusion [90]. En conséquence, un film visqueux est caractérisé par un déficit d'énergie rétro-diffusée et apparaît comme une zone sombre dans les images [81], ce qui donne visuellement l'impression d'une mer plus bruitée que la nappe dans les zones non polluées. En considérant le fait qu'un film visqueux a un impact caractéristique sur la répartition de l'énergie des vagues selon les différentes longueurs d'onde (des vagues de gravité-capillarité jusqu'à la houle), nous avons appliqué le modèle des CMCs vectorielles sur une représentation multi-échelle de l'image originale [126, 57]. La représentation multi-échelle fournit une représentation complète de l'image qu'elle décompose en une image passe-bas et des images passe-haut horizontales et verticales, à différentes échelles. En accord avec les lois à considérer pour les données RSO [93], le système de Pearson a été utilisé pour décrire les coefficients passe-bas, et la famille des gaussiennes généralisées a été choisie pour sa capacité à bien représenter les coefficients passe-hauts [73].

L'image de la figure 2.6(a) présente une vue satellitaire de la nappe d'hydrocarbure déversée juste après le naufrage du Prestige au large des côtes espagnoles, le 17 novembre 2002. Celle-ci se sépare en deux traînées à cause de l'état de la mer, du vent et de la viscosité du pétrole. Aussi, il devient difficile de caractériser la présence d'hydrocarbure entre ces deux traînées. De plus, les conditions atmosphériques induisent une pression particulière sur la surface de la mer et un front apparaît avec une radiométrie plus sombre sans pour autant être révélateur de la présence d'huile. L'image 2.6 (c) présente le résultat de segmentation en trois classes. Une première classe se spécialise sur la nappe d'hydrocarbure, une seconde sur une zone ambiguë dans laquelle la surface de la mer apparaît plus sombre, et la troisième caractérise la mer libre d'hydrocarbure. Le résultat de segmentation est bien plus homogène que lorsqu'on utilise le modèle des CMC



(a) (b) (c)

Figure 2.6 – (a) Acquisition ASAR d'Envisat effectuée lors du naufrage du Prestige en 2002, ©ESA. (b) Segmentation par CMC directement sur l'image. (c) Segmentation par CMC vectorielle le long d'une décomposition multi-échelle de l'image. Résultats extraits de [57].

directement sur l'image (b).

L'intérêt supplémentaire de l'approche présentée ici est d'effectuer une segmentation qui s'appuie sur l'estimation préalable d'un mélange de lois. Ainsi, une fois la segmentation effectuée, une analyse des types et des formes des lois obtenues permet de valider ou d'infirmer la présence effective d'une nappe d'hydrocarbure dans l'image. En effet, certains phénomènes météorologiques ou océanographiques –comme l'absence de vent, la présence de vents très violents ou de phénomènes océanographiques comme les « up-wellings »–, peuvent être confondus avec des nappes d'hydrocarbures.

### 2.2.3 Modèles spécifiés par copules

Les *copules* sont un outil statistique pour la modélisation et l'estimation de la dépendance entre variables aléatoires [133], et constituent une alternative à l'utilisation des méthodes de l'analyse multivariée. L'intérêt de ce modèle est, d'une part, d'introduire de nouvelles formes de dépendance entre les différentes composantes et, d'autre part, de permettre le contrôle les lois marginales unidimensionnelles des lois multivariées. La théorie sous-jacente permet de construire un nombre très important de nouveaux modèles [97] et notre objectif ici n'est évidemment pas de dresser un panorama du domaine mais plutôt de rappeler quelques principes et théorèmes fondamentaux à l'origine de cette théorie. Nous présentons ensuite l'utilisation (et l'estimation) de deux copules, intéressantes par leur simplicité, dans le cas de la segmentation d'images multi-composantes. Ce travail a été réalisé en collaboration avec W. Pieczynski et N. Brunel, qui ont été les premiers à introduire ces modèles en segmentation d'images [21], alors qu'ils ont été découverts à la fin des années 50 et sont utilisés depuis une vingtaine d'années dans le domaine de la finance [158].

**Définition** Une copule (bi-variée) est une fonction de répartition (f.d.r.) sur $[0,1]^2$ et dont les marges sont uniformes sur $[0,1]$.

Le résultat fondamental de la théorie des copules est le théorème d'existence de Sklar (1959) qui clarifie le lien entre les marges et la loi jointe :

> *Soit $F$ une f.d.r. sur $\mathbb{R}^2$ dont les distributions marginales ont pour f.d.r. $F_1$ et $F_2$. Alors il existe une copule $\mathcal{C}$ telle que*
>
> $$\forall z^1, z^2 \in \mathbb{R}, \quad F\left(z^1, z^2\right) = \mathcal{C}\left(F_1\left(z^1\right), F_2\left(z^2\right)\right).$$
>
> *La copule est unique si les f.d.r. des marges sont continues.*

Si la copule est dérivable, le théorème de Sklar permet d'écrire directement le lien entre les densités marginales $f_1$ et $f_2$ et la densité jointe $f$ de la loi multivariée

$$f\left(z^1, z^2\right) = f_1\left(z^1\right)\ f_2\left(z^2\right)\ c\left(F_1\left(z^1\right), F_2\left(z^2\right)\right), \tag{2.19}$$

où la densité de la copule $c$ est la dérivée de $\mathcal{C}$ selon les deux coordonnées. Nous pouvons alors écrire :

$$f\left(z^1 \middle| z^2\right) = f_1\left(z^1\right)\ c\left(F_1\left(z^1\right), F_2\left(z^2\right)\right).$$

Ces résultats restent valables pour une f.d.r. définie sur $\mathbb{R}^M$ et ses marges $F_1, \ldots, F_M$. Pour construire la loi conjointe d'un vecteur aléatoire $\vec{Z} = \left(Z^1, \ldots, Z^M\right)^t$, nous devons spécifier d'abord les lois marginales de chaque $Z^m$ (en utilisant des informations *a priori* sur leur forme), et ensuite une copule qui représente de manière exhaustive la dépendance entre les variables aléatoires.

**Exemples**

► ***Copule produit*** – La copule $\mathcal{C}_\perp\left(u^1,\ldots,u^M\right)=u^1\ \ldots\ u^M$ est appelée copule produit, et nous avons

$$c_\perp\left(u^1,\ldots,u^M\right)=1.$$

Nous en déduisons que toute distribution multidimensionnelle construite avec cette copule s'écrit

$$f\left(\vec{z}\right)=\prod_{m=1}^{M}f^m\left(z^m\right).$$

Nous retrouvons l'équation (2.17), page 28. La copule produit caractérise l'indépendance entre deux variables aléatoires.

Pour construire des copules, deux méthodes sont à signaler. La méthode de construction directe, qui aboutit aux copules archimédiennes telles les copules de Clayton, de Gumbel et de Franck. La seconde méthode consiste à déterminer, à l'aide du théorème de Sklar, la copule sous-jacente à partir de lois multivariées usuelles. C'est celles-ci que nous avons choisies pour modéliser les densités d'attache aux données multivariées car l'expression de leur d.d.p. est souvent plus maniable que celles des copules archimédiennes. Voici quelques exemples

► ***Copule gaussienne*** – À partir de l'éq. (2.19), nous pouvons écrire :

$$\frac{1}{2\pi\sqrt{1-\rho^2}}\,e^{-\frac{1}{2(1-\rho^2)}\left(x^2+y^2-2\rho xy\right)}=c_g\left(F_1^{-1}\left(x\right),F_2^{-1}\left(y\right);\rho\right)\ \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\,x^2}\ \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\,y^2},$$

soit, en notant $x=F_1^{-1}\left(u^1\right)$ et $y=F_2^{-1}\left(u^2\right)$,

$$c_g\left(u^1,u^2;\rho\right)=\frac{1}{2\pi\sqrt{1-\rho^2}}\,e^{-\frac{1}{2(1-\rho^2)}\left(x^2+y^2-2\rho xy\right)+\frac{1}{2}\left(x^2+y^2\right)}.$$

Dans le cas général à $M$ composantes, l'expression de la densité de la copule s'écrit

$$c_g\left(u^1,\ldots,u^M;\boldsymbol{\rho}\right)=|\boldsymbol{\rho}|^{-\frac{1}{2}}\ e^{-\frac{1}{2}\left(\boldsymbol{\zeta}^t(\boldsymbol{\rho}^{-1}-\boldsymbol{I})\boldsymbol{\zeta}\right)},$$

avec
– $\boldsymbol{\zeta}^t=\left(\Phi^{-1}\left(u^1\right),\ldots,\Phi^{-1}\left(u^M\right)\right)$ et $\Phi$ la f.d.r. d'une gaussienne normalisée.
– $\boldsymbol{\rho}$ désigne la matrice de corrélation de la copule gaussienne. Si $\boldsymbol{\rho}=\boldsymbol{I}$, alors $\mathcal{C}_g=\mathcal{C}_\perp$.

**Remarque 2.4** *La copule gaussienne est intéressante car les calculs sont particulièrement simples et la structure de dépendance entre les composantes est intuitive puisqu'elle est basée sur la notion usuelle de corrélation. Comparée à la loi gaussienne multivariée, une loi construite à partir d'un copule gaussienne n'a pas nécessairement ses marges gaussiennes ni des marges identiques : nous pouvons construire une structure de dépendance régie par une simple matrice de corrélation, et des lois avec des marges gamma, beta...* [11].

**Remarque 2.5 (ACP *vs* Copule gaussienne)** *On peut se poser la question de la différence entre l'utilisation de l'ACP et l'utilisation d'une copule gaussienne pour l'estimation de d.d.p. multidimensionnelles. Pour l'ACP, ce sont les lois des composantes principales de l'image qui sont estimées. Les marginales sont, quand à elles, difficiles à obtenir puisqu'elles nécessitent l'intégration de $M-1$ variables,* cf. *équation* (2.18)

11. — On peut aussi envisager des marges non paramétriques produisant un modèle semi-paramétrique.

*page 29. Ainsi, contrairement aux copules, l'ACP ne permet pas de sélectionner la forme des marginales, mais uniquement la forme des lois associées aux composantes principales de l'image.*

► ***Copule de Student*** – Des calculs similaires conduisent à l'expression de la copule de Student, construite à partir de l'expression d'une loi de Student multivariée avec $\nu$ degrés de liberté

$$c_s\left(u^1,\ldots,u^M;\boldsymbol{\rho},\nu\right)=|\boldsymbol{\rho}|^{-\frac{1}{2}}\ \frac{\Gamma\left(\frac{\nu+M}{2}\right)\ \Gamma\left(\frac{\nu}{2}\right)^{M-1}}{\Gamma\left(\frac{\nu+1}{2}\right)^{M}}\ \frac{\left(1+\frac{1}{\nu}\boldsymbol{\xi}^t\boldsymbol{\rho}^{-1}\boldsymbol{\xi}\right)^{-\frac{\nu+M}{2}}}{\prod_{m=1}^{M}\left(1+\frac{1}{\nu}\xi_m^2\right)^{-\frac{\nu+1}{2}}},$$

avec

- $\boldsymbol{\xi}^t=\left(T_\nu^{-1}\left(u^1\right),\ldots,T_\nu^{-1}\left(u^M\right)\right)$ et $T_\nu$ la f.d.r. d'une loi de student normalisée avec $\nu$ degrés de liberté.
- $\boldsymbol{\rho}$ la même matrice de corrélation que celle de la copule gaussienne. Par contre, si $\boldsymbol{\rho}=\boldsymbol{I}$, la copule de Student ne s'identifie pas à la copule produit (et ne signifie donc pas l'indépendance des variables aléatoires).

La figure 2.7 présente des exemples des trois copules avec, d'une part, des marges Gaussiennes (1re colonne) et, d'autre part, des marges Gamma (2e colonne) identiques. Les courbes iso-probabilités sont sensiblement différentes, mêmes dans le cas des copules $\mathcal{C}_g$ et $\mathcal{C}_s$.

**Estimation des paramètres** Au lieu de chercher une transformation adéquate pour rendre les composantes décorrélées (ACP) ou indépendantes (ACI), nous cherchons à estimer les paramètres des copules et des marges. Nous cherchons à estimer les paramètres $\left(\Theta=(\theta_m)_{1\leq m\leq M},\eta\right)$ d'une densité multivariée, dont les densités marginales appartiennent au modèle $\{f(.,\theta_m)\}$ et dont la copule appartient au modèle $\{c_\eta\}$.

Pour cela, nous pouvons utiliser la méthode d'estimation IFM (*Inference Functions for Margins*) [97] qui se décompose en deux étapes :

(i) estimation du maximum de vraisemblance $\hat{\theta}_m$ de chaque marge à partir de l'échantillon $(z_n^m)_{1\leq n\leq N_k}$ et création d'un nouveau jeu de données : $\forall n\in[1,N_k]$ and $m\in[1,M]$, $\quad u_n^m=F\left(z_n^m\middle|\hat{\theta}_m\right)$.

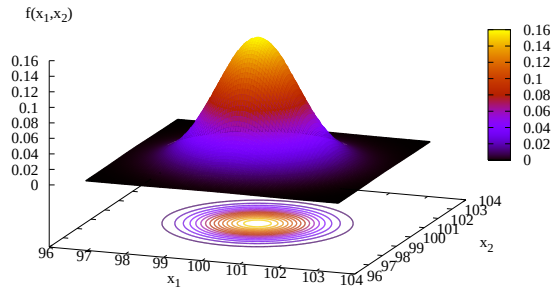
(ii) estimation du maximum de vraisemblance $\hat{\eta}$ à partir du nouvel échantillon $\vec{\boldsymbol{u}}=(\vec{u}_1,\ldots,\vec{u}_{N_k})$.

L'estimation des paramètres des marges ne pose pas de difficulté particulière, voyons maintenant à quoi aboutit la méthode IFM pour l'estimation des paramètres spécifiques des deux copules qui nous intéressent.

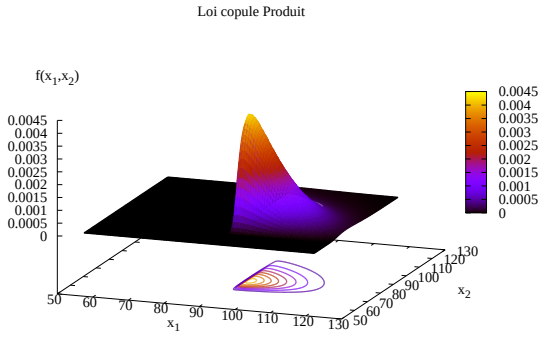
► ***Copule gaussienne*** – La matrice $\boldsymbol{\rho}$ peut être estimée de la manière suivante

$$\hat{\boldsymbol{\rho}}=\frac{1}{N}\sum_{n=1}^{N_k}\boldsymbol{\zeta}_n\,\boldsymbol{\zeta}_n^t, \tag{2.20}$$

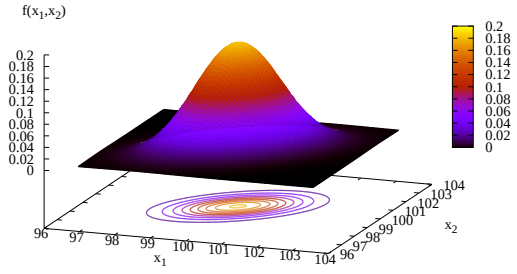
où $\boldsymbol{\zeta}_n^t=\left(\Phi^{-1}(u_n^1),\ldots,\Phi^{-1}(u_n^M)\right)$.



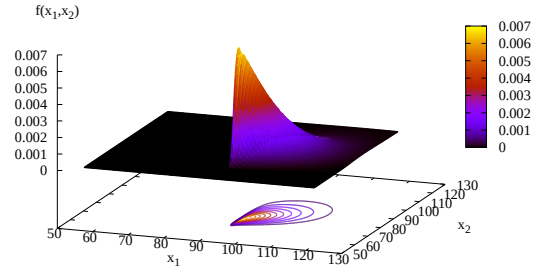

(a) $\mathcal{C}_\perp$



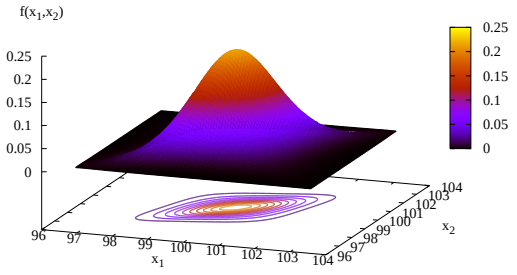

(b) $\mathcal{C}_\perp$



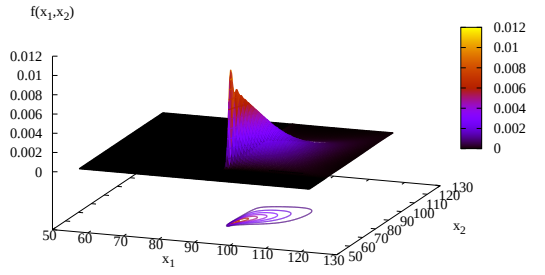

(c) $\mathcal{C}_g$ - $\rho = 0.5$




(d) $\mathcal{C}_g$ - $\rho = 0.5$




(e) $\mathcal{C}_s$ - $\rho = 0.5, \nu = 10$




(f) $\mathcal{C}_s$ - $\rho = 0.5, \nu = 10$

Figure 2.7 – Exemples de densités bi-variées pour les trois copules avec des marges gaussiennes (1re colonne) et des marges gamma (2e colonne). Résultats extraits de [62].

Figure 2.8 – Une image multicomposantes de quatre bandes. Taille : $256 \times 256$.

► ***Copule de Student*** – La matrice $\boldsymbol{\rho}$ peut être estimée avec l'algorithme itératif suivant [158] :

(i) Soit $\hat{\boldsymbol{\rho}}_0$ l'estimation de la matrice $\boldsymbol{\rho}$ obtenue avec l'équation (2.20).

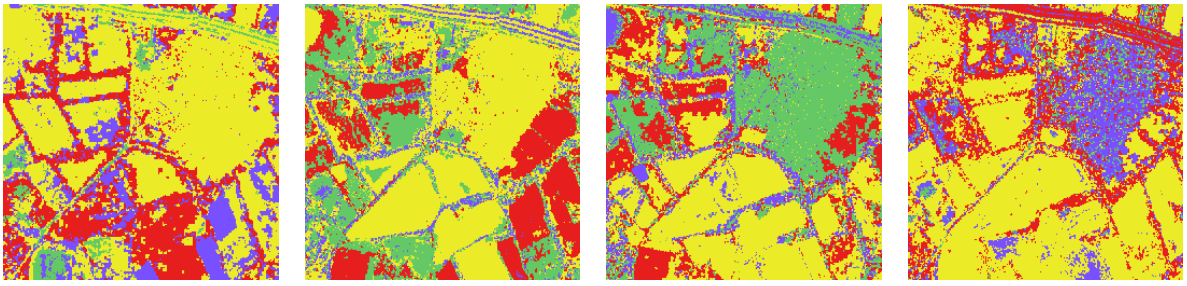
(ii) L'estimation $\hat{\boldsymbol{\rho}}_{q+1}$ est obtenue de la façon suivante :

$$\hat{\boldsymbol{\rho}}_{q+1} = \left( \frac{\nu + M}{\nu \, N_k} \right) \sum_{n=1}^{N_k} \frac{\boldsymbol{\xi}_n \, \boldsymbol{\xi}_n^t}{1 + \boldsymbol{\xi}_n^t \, \hat{\boldsymbol{\rho}}_q^{-1} \, \boldsymbol{\xi}_n}.$$

La méthode IFM fournit seulement une approximation du maximum de vraisemblance puisque la maximisation globale de la vraisemblance est remplacée par deux maximisations successives (et plus simples). Néanmoins la méthode donne de bons résultats et est simple à mettre en œuvre. Pour son utilisation dans les CMCbi, il s'agit seulement d'adapter la démarche de la stratégie ECI par cette procédure en deux étapes.

**Illustration : segmentation d'images multi-spectrales –** À titre d'illustration, nous avons segmenté l'image 4-bandes de la figure 2.8 à l'aide des trois copules décrites précédemment et de la méthode par ACI. Cette image vectorielle est le résultat de la réduction de dimension, à l'aide d'une méthode de poursuite de projection [119], d'une image hyperspectrale CASI (*Compact Airborne Spectrographic Imager*) de 17 bandes de 450 à 950 nm, avec $2m$ de résolution au sol. Dans tout ces cas, les densités des composantes indépendantes (méthode ACI - critère de non-gaussianité de Hyvärinen [94]) et des marginales (méthode par copules) sont issues de la famille gamma. Les résultats de segmentation en quatre classes (forêts, champs, routes et terres non cultivées), issus du critère MPM sont présentés dans la figure 2.9. Ces résultats sont tous différents et leur différence est exclusivement liée à la façon dont les lois multivariées ont été estimées. Il est cependant bien difficile, sans vérité-terrain, de déterminer quel résultat est le plus approprié ; la réponse dépendra certainement de l'application envisagée.

Depuis leur introduction dans le domaine du traitement du signal et de l'image par le biais des chaînes de Markov cachées, ces modèles par copule suscitent un intérêt croissant. Ils ont notamment été appliqués dans les CMCs couples et triplets [19] que nous discuterons plus loin. La copule de Marchal-Olkin est quand-à-elle utilisée pour définir une expression de l'information mutuelle indépendamment de la modélisation des lois marginales [118]. C'est un point de vue extrêmement intéressant en segmentation d'images hétérogènes puisqu'il permet de traiter conjointement, dans un contexte rigoureusement statistique, des observations optique et radar qui sont pourtant de nature très différentes. Ces outils ont été appliqués pour la détection supervisé de changements entre deux images d'une même scène [127, 125], en utilisant des noyaux stochas-



(a) Copule produit (b) Copule gaussienne (c) Copule de Student ($\nu = 10$) (d) ACI

Figure 2.9 – Résultats de segmentation. Les mêmes couleurs ont été choisies pour les classes les plus similaires. Résultats extraits de [62].

tiques à base de copule dans les méthodes de classification par SVM (*Support Vector Machine*).

## 2.3 Deux nouvelles topologies pour les CMCs

Pour améliorer la prise en compte du voisinage dans les modèles de CMC, de nombreux travaux ont essayé de pallier aux limitations du modèle en relâchant certaines des hypothèses **H1** à **H3**, page 18. Nous pouvons citer les modèles semi-markoviens, expérimentés avec succès en traitement de la parole [109], pour lesquels la probabilité d'un changement d'état dépend du temps passé dans l'état courant, contrairement au modèle classique où cette probabilité est constante. D'autres travaux proposent de relâcher l'hypothèse d'indépendance conditionnelle des observations tels que le modèle « *factorial HMM* » [77], le modèle « *HMM2* » [98] et le modèle « *double Markov chain* » [12] qui utilisent plusieurs chaînes de Markov à différents niveaux.

Je me suis intéressé à deux extensions originales du modèle des CMCs qui permettent également de relâcher certaines hypothèses, modifiant la topologie de la chaîne et aboutissant à de nouveaux modèles strictement plus généraux : il s'agit des « CMCs d'ordre supérieur » [65, 66] et des « CMCs couples » [53, 11]. Dans le premier cas, nous avons étendu la mémoire du processus de Markov et, ce qui est plus original, celle associée aux lois d'attache aux données. Dans le second, développé en collaboration avec W. Pieczynski, nous avons relâché l'hypothèse de markovianité du processus caché $\boldsymbol{X}$ au profit de la markovianité du couple des processus caché et observé $(\boldsymbol{X}, \boldsymbol{Y})$.

Ces deux techniques sont différentes dans leur modélisation des interactions entre étiquettes et données mais proches du modèle original dans leurs traitements mathématique et informatique. Dans les deux cas, nous définissons des procédures d'estimation et de classification similaires aux CMCs, ce qui permet de conserver la rapidité des traitements même si les algorithmes ainsi construits deviennent plus complexes : ils requièrent un temps de calcul plus important et consomment plus de mémoire vive.

### 2.3.1 Modèle des CMCs d'ordre supérieur

Par extension du modèle classique, nous avons également proposé une méthode d'estimation des paramètres pour le modèle des *chaînes de Markov d'ordre* $R$ quelconque [53], c'est-à-dire le

modèle pour lequel la mémoire du processus de Markov ne se limite pas au seul état précédent mais aux $R$ états précédents. Le modèle pour lequel $R = 2$ a été testé dans de nombreux domaines scientifiques y compris le traitement de la parole, la reconnaissance de gestes ou la robotique [113, 47, 4]. Le succès a été modéré étant donné le faible gain sur les taux d'erreur de classification comparé au modèle classique (pour lequel $R = 1$) et l'augmentation des temps de traitements et de mémoire vive associée. Ce résultat découle du fait que n'importe quelle chaîne de Markov stationnaire d'ordre $R$ peut être transformée en une chaîne équivalente d'ordre 1 [96]. L'originalité de notre approche est de considérer que les lois d'attaches aux données dépendent elles-mêmes de plusieurs états précédents, intégrant ainsi une mémoire d'ordre supérieur pour modéliser les interactions entre étiquettes des classes et données.

**Modèle d'ordre** $R$**–** L'hypothèse de markovianité (**H1** page 18) s'écrit $p\left(x_{n+1}\left|\boldsymbol{x}_{1:n}\right.\right) = p\left(x_{n+1}\left|\boldsymbol{x}_{n-R+1:n}\right.\right)$. Une chaîne de Markov stationnaire d'ordre $R$ est entièrement caractérisée par sa loi dite *a priori* :

$$p\left(x_{n-R+1}, \cdots, x_n + 1\right)$$

indépendante de $n$. Par marginalisation et conditionnement, la distribution de $\boldsymbol{X}$ est déterminée par sa loi initiale et les matrices de transitions suivantes :

- **Matrice de transition d'ordre** $R$ **:** $\boldsymbol{A}^R$, valide pour $n \geq R$, et dont les éléments sont définis par
$$a^R_{j|i_1,\cdots,i_R} = p\left(X_{R+1} = j\left|X_1 = i_1, \ldots, X_R = i_R\right.\right),$$
où $i_1, \ldots, i_R$ et $j$ appartiennent à $\Omega$.
- **Matrices de transition intermédiaires :** $\boldsymbol{A}^n$, pour les pixels $1 \leq n < R$, obtenues en marginalisant $\boldsymbol{A}^R$. Pour $n = R - 1$, la matrice de transition intermédiaire est définie par
$$a^{R-1}_{j|i_1,\cdots,i_{R-1}} = p\left(X_R = j\left|X_1 = i_1, \ldots, X_{R-1} = i_{R-1}\right.\right) = \sum_{i_R=1}^{K} a^R_{j|i_1,\cdots,i_R}.$$
La matrice $\boldsymbol{A}^{R-2}$ est obtenue en marginalisant $\boldsymbol{A}^{R-1}$, et ainsi de suite.
- **Loi initiale :** $\boldsymbol{\pi}$ caractérisant la distribution de $X_1$ : $\pi_j = p\left(X_1 = j\right)$.

Dans ce cadre, la loi *a priori* de $\boldsymbol{X}$ s'écrit

$$p\left(\boldsymbol{x}\right) = \pi_{x_1} \prod_{n=1}^{R-1} a^n_{x_{n+1}|\boldsymbol{x}_{1:n}} \prod_{n=R}^{N-1} a^R_{x_{n+1}|\boldsymbol{x}_{n-R+1:n}},$$

et, sous les hypothèses **H2** et **H3** page 2, la loi jointe de l'eq. (2.3) devient

$$p\left(\boldsymbol{x}, \boldsymbol{y}\right) = \pi_{x_1} f_{x_1}(y_1) \prod_{n=1}^{R-1} a^n_{x_{n+1}|\boldsymbol{x}_{1:n}} f_{x_{n+1}}(y_{n+1}) \prod_{n=R}^{N-1} a^R_{x_{n+1}|\boldsymbol{x}_{n-R+1:n}} f_{x_{n+1}}(y_{n+1}),$$

avec $f_k(y_n) = p\left(y_n\left|x_n = k\right.\right)$ les lois d'attache aux données que nous supposerons gaussiennes sans perte de généralité.

L'hypothèse **H3** n'est pas strictement nécessaire et peut être relâchée de la manière suivante

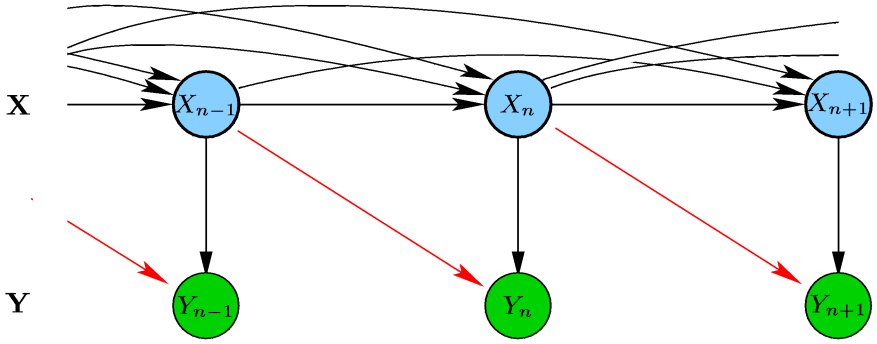



Figure 2.10 – Hypothèses d'indépendance d'un modèle $\text{CMC}_{3-2}$. Les lignes en pointillés représentent les nouvelles connections obtenues en relâchant l'hypothèse **H3**.

> **H3 bis** La distribution de chaque $Y_n$ conditionnellement à $\boldsymbol{X}$ est égale à sa distribution conditionnellement à $\{X_{n-S+1}, \ldots, X_n\}$, avec $S \leq R$.

La chaîne de Markov d'ordre $R$ utilisant l'hypothèse **H3 bis** est notée $\text{CMC}_{R-S}$. $\text{CMC}_{1-1}$ désigne donc le modèle classique. Le paramètre $S$ représente la longueur de la mémoire associée aux lois d'attache aux données et doit être inférieur à $R$ (*cf.* [11]). Pour $n \geq S$, leur expression devient

$$f_{\boldsymbol{x}_{n-S+1:n}}(y_n) = p\left(y_n \,|\boldsymbol{x}_{n-S+1:n}\right).$$

De ce fait, nous comptons $K^S$ lois d'attache aux données monodimensionnelles pour caractériser le mélange. Pour $n < S$, $f_{\boldsymbol{x}_{n-S+1:n}}(y_n)$ est remplacé par $\underbrace{f_{x_n, \ldots, x_n}}_{S \text{ termes}}(y_n)$.

Le schéma de la figure 2.10 illustre les interactions entre étiquettes et données dans le cas du modèle $\text{CMC}_{3-2}$. Les lignes continues connectant les variables $X_n$ symbolise l'ordre de la chaîne de Markov, alors que la ligne continue connectant $Y_n$ avec $X_n$ symbolise l'hypothèse **H3**. Les lignes en pointillés montrent l'apport de la nouvelle hypothèse **H3 bis** sur la structure de dépendance. L'hypothèse **H2** restant inchangée, nous sommes toujours dans le cadre d'un modèle à bruit indépendant. Il est intéressant de noter que l'ordre 2 permet, d'un certaine manière, de prendre en compte un voisinage diagonale dans les images étant donné les circonvolutions du parcours de Hilbert-Peano, *cf.* figure 2.1, page 16. Ceci le rapproche du système de cliques d'ordre 2 généralement utilisées dans le cadre des champs de Markov.

**Classification MPM–** L'estimateur du MPM adapté à ce nouveau contexte s'écrit

$$\forall n \in [R, \ldots, N], \quad \widehat{x}_n^{\text{MPM}}(\boldsymbol{y}) = \arg\max_{x_n \in \Omega} \xi_n(\boldsymbol{x}_{n-R+1:n}),$$

avec $\xi_n(\boldsymbol{x}_{n-R+1:n}) = p\left(\boldsymbol{x}_{n-R+1:n} \,|\boldsymbol{y}\right)$ les probabilités marginales *a posteriori* [12]. En adaptant les probabilités *avant* et *arrière*

$$\begin{aligned} \alpha_n(\boldsymbol{x}_{n-R+1:n}) &= p\left(\boldsymbol{x}_{n-R+1:n} \,|\boldsymbol{y}_{1:n}\right). \\ \beta_n(\boldsymbol{x}_{n-R+1:n}) &= \frac{p\left(\boldsymbol{y}_{n+1:N} \,|\boldsymbol{x}_{n-R+1:n}\right)}{p\left(\boldsymbol{y}_{n+1:N} \,|\boldsymbol{y}_{1:n}\right)}, \end{aligned}$$

12. — Dans la suite de ce paragraphe, toutes les formules seront données pour $n \geq R$ ; celles correspondant à $n < R$ s'en déduisant simplement.

nous vérifions [13] également que

$$\xi_n(\boldsymbol{x}_{n-R+1:n}) = \alpha_n(\boldsymbol{x}_{n-R+1:n})\, \beta_n(\boldsymbol{x}_{n-R+1:n}), \tag{2.21}$$

et que les probabilités conjointes *a posteriori* $\Psi_n(\boldsymbol{x}_{n-R+1:n+1}) = p\left(\boldsymbol{x}_{n-R+1:n}, x_{n+1} \,|\boldsymbol{y}\right)$ de l'eq. (2.14) s'écrivent

$$\Psi_n(\boldsymbol{x}_{n-R+1:n+1}) \propto \alpha_n(\boldsymbol{x}_{n-R+1:n})\, a^R_{x_{n+1}|\boldsymbol{x}_{n-R+1:n}}\, f_{\boldsymbol{x}_{n-S+2:n+1}}(y_{n+1})\, \beta_{n+1}(\boldsymbol{x}_{n-R+2:n+1}), \tag{2.22}$$

avec $\displaystyle\sum_{\boldsymbol{x}_{n-R+1:n+1}\in\Omega^{R+1}} \Psi_n(\boldsymbol{x}_{n-R+1:n+1}) = 1.$

**Estimation des paramètres avec ECI–** La procédure d'estimation des paramètres est essentiellement identique à celle présentée dans le cas classique, avec cependant un nombre bien plus important de paramètres à estimer.

- **Ré-estimation des paramètres de la chaîne de Markov d'ordre supérieur–** La matrice de transition d'ordre $R$ et la loi initiale sont obtenues en utilisant les probabilités *a posteriori* des équations (2.21) et (2.22)

$$\begin{aligned} a^{R,[p]}_{j|i_1,\cdots,i_R} &= \frac{\displaystyle\sum_{n=R}^{N-1} \psi_n^{[p]}(x_{n-R+1}=i_1,\ldots,x_n=i_R,x_{n+1}=j)}{\displaystyle\sum_{n=R}^{N-1} \xi_n^{[p]}(x_{n-R+1}=i_1,\ldots,x_n=i_R)}. \\ \pi_j^{[p]} &= \frac{1}{N-R+1}\sum_{n=R}^{N}\sum_{\boldsymbol{x}_{n-R+2:n}\in\Omega^{R-1}} \xi_n^{[p]}(x_{n-R+1}=j,\boldsymbol{x}_{n-R+2:n}). \end{aligned}$$

  En toute rigueur, les termes correspondant aux pixels d'indice $n < R$ doivent être ajoutés aux sommes. Cependant leur contribution numérique est très faible puisque $N \gg R$ pour une image. Nous avons donc préféré les omettre pour des raisons de clarté.

- **Ré-estimation des paramètres des lois d'attache aux données–** Selon ECI, les paramètres des lois d'attache aux données s'estiment en simulant $\boldsymbol{X}$ conditionnellement à $\boldsymbol{Y}$ dont la loi est celle d'une chaîne de Markov d'ordre $R$ non-homogène avec pour transitions

$$\begin{aligned} \tilde{a}^{R,[p]}_{x_{n+1}|\boldsymbol{x}_{n-R+1:n}} &= p^{[p]}\left(x_{n+1}|\,\boldsymbol{x}_{n-R+1:n},\boldsymbol{y}\right) \\ &= \frac{a^R_{x_{n+1}|\boldsymbol{x}_{n-R+1:n}}\, f_{\boldsymbol{x}_{n-S+2:n+1}}(y_{n+1})\, \beta_{n+1}(\boldsymbol{x}_{n-R+2:n+1})}{\displaystyle\sum_{x_{n+1}\in\Omega} a^R_{x_{n+1}|\boldsymbol{x}_{n-R+1:n}}\, f_{\boldsymbol{x}_{n-S+2:n+1}}(y_{n+1})\, \beta_{n+1}(\boldsymbol{x}_{n-R+2:n+1})}. \end{aligned}$$

13. — Toutes les démonstrations sont consultables dans [11].

Ainsi, en connaissant la répartition des données dans chacune des classes à chaque itération $p$, nous pouvons estimer les moyennes $\mu_1^{[p]}(i_{1:S})$ et variances $\mu_2^{[p]}(i_{1:S})$ [a], avec $i_1, \cdots, i_S \in \Omega$, des $K^S$ lois gaussiennes à l'aide d'estimateurs empiriques classiques :

$$\mu_1^{[p]}(i_{1:S}) = \frac{1}{N_{i_{1:S}}} \sum_{n=1}^{N} \left( y_n \times \mathbf{1}_{\left[\boldsymbol{x}_{n-S+1:n}^{[p,1]}=(i_{1:S})\right]} \right),$$

$$\mu_2^{[p]}(i_{1:S}) = \frac{1}{N_{i_{1:S}}} \sum_{n=1}^{N} \left( \left[ y_n - \mu_1^{[p]}(i_{1:S}) \right] \times \mathbf{1}_{\left[\boldsymbol{x}_{n-S+1:n}^{[p,1]}=(i_{1:S})\right]} \right)^2,$$

avec $N_{i_{1:S}} = \sum_{n=1}^{N} \mathbf{1}_{\left[\boldsymbol{x}_{n-S+1:n}^{[p,1]}=(i_{1:S})\right]}$ et la fonction $\mathbf{1}_{[\boldsymbol{x}_{n-S+1:n}=(i_{1:S})]} = 1$ si $\boldsymbol{x}_{n-S+1:n} = (i_{1:S})$.

*a.* — en notant l'ensemble d'indices $i_1, \cdots, i_S$ par $i_{1:S}$.

Il est important de noter la très grande similarité des formules de re-estimation avec le cas classique présenté dans le paragraphe 2.1.

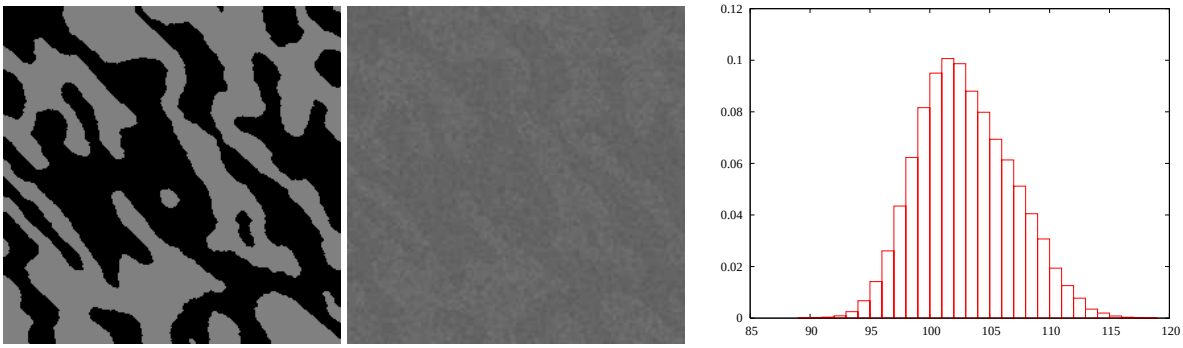
**Remarque 2.6** *L'augmentation de la mémoire de la chaîne de Markov entraîne une complexité accrue des algorithmes ainsi qu'une consommation de mémoire vive importante. Ceci peut être illustré par la taille des tableaux nécessaires pour enregistrer les probabilités* avant *et* arrière *ci-dessus. D'autre part, à nombre d'échantillons identique, l'augmentation du nombre de paramètres à estimer peut être à l'origine d'une diminution de la robustesse des estimées. Ainsi, $R$ et $S$ doivent prendre des valeurs maximales raisonnables (typiquement 3).*

**Illustration 1 : Segmentation d'une image simulée–** Pour évaluer ces modèles d'ordre supérieur en fonction des valeurs de $R$ et $S$ ainsi que les méthodes d'estimation associées, nous avons utilisé l'image bruitée de la figure 2.11. Les résultats de segmentation non supervisée sont présentés dans le tableau de la figure 2.12, pour $R$ et $S$ variant de 1 à 3, avec $S \leq R$. Les taux d'erreur diminuent sensiblement lorsque $R$ et $S$ croissent simultanément. Ces résultats confirment aussi le fait que l'augmentation seule de $R$ n'a pas d'impact sur les résultats. Nous pouvons aussi noter que la quantité de mémoire augmente exponentiellement avec $R$ et $S$ ; ceci pouvant créer des difficultés lors du traitement de grandes images. Enfin les temps de calculs rapportés dans le tableau mettent exergue que seule l'augmentation de $R$ a un impact significatif sur $\tau$ : lorsque $S$ augmente et pour $R$ fixé, la valeur de $\tau$ reste approximativement constante. Un exemple des paramètres estimés par ECI par le modèle $\text{CMC}_{2-2}$ est reporté dans [11].

**Illustration 1 : Segmentation d'une image radar de nappe d'hydrocarbure–** Ce second exemple illustre l'intérêt du modèle d'ordre supérieur pour des images réelles, en l'occurrence l'image ERS-SAR d'une nappe d'hydrocarbure acquise le 3 octobre 1992, au large des côtes égyptiennes, *cf.* Fig. 2.13.

La difficulté de la localisation des nappes de pétrole n'est pas tant de détecter les grandes nappes homogènes mais plutôt de déterminer précisément les frontières de cette nappe et également de détecter (sans fausses alarmes) des zones plus ambiguës telles les tâches sombres qui apparaissent sur le côté droit de l'image. Le problème de la détection de nappe d'hydrocarbures est loin d'être trivial [14], et de nombreuses études sont d'ailleurs menées à ce sujet [18]. Dans cette

14. — en particulier à cause du problème de distinction entre nappes de pétrole et nappes organiques, naturellement présentes en pleine mer.




(a) Taille : $256 \times 256$ (b) (c)

Figure 2.11 – Données simulées : (a) Image originale à deux classes, (b) Image bruitée et (c) son histogramme.

problématique, les capteurs radar présentent l'avantage de pouvoir fournir des données de jour comme de nuit et ce, quelles que soient les conditions météorologiques, ce qui leur confère un intérêt majeur dans un cadre opérationnel.

La figure 2.14 présente les résultats de segmentation pour $R$ et $S$ variant de 1 à 3. Dans tous les cas, la classification de l'image a été réalisée avec le critère bayésien du MPM. Le nombre d'itérations ECI a été fixé à 50 et les lois d'attache aux données ont été modélisées par des distributions gaussiennes. Il apparaît clairement que le modèle $\text{CMC}_{2-2}$ offre une segmentation beaucoup plus homogène, de la mer comme de la nappe, que le modèle classique pour lequel $R = S = 1$. Par contre, le modèle $\text{CMC}_{3-3}$ n'améliore pas la classification par rapport au modèle pour lequel $R = S = 2$, confirmant la nécessité de rechercher un compromis entre la complexité du modèle et l'application pour une image donnée.

Le modèle classique des CMCs ne peut caractériser que des interactions proches entre pixels voisins et jouit donc d'une capacité limitée pour décrire des phénomènes de plus grandes échelles. La prise en compte d'une mémoire plus étendue tant au niveau de la chaîne de Markov que des lois d'attaches aux données permet de mieux apprendre des structures de bruits plus complexes comme des motifs texturaux tels la houle qui module le bruit de chatoiement. Notons aussi qu'au sens du parcours de Peano, l'ordre $R = S = 2$ prend d'une certaine façon en compte un voisinage diagonale dans les images, qui le rapproche d'un système de cliques d'ordre 2 dans le cadre des champs de Markov.

### 2.3.2 Modèle des CMCs couples

Comme nous avons pu le souligner dans le paragraphe 2.1, le modèle classique des chaînes de Markov cachées est parfois appelé « CMC à bruit indépendant » du fait de l'hypothèse **H2** page 2, qui précise l'indépendance conditionnelle des observations par rapport à $\boldsymbol{X}$. Or, dans les situations réelles, cette indépendance est souvent sujette aux critiques et la possibilité de prendre en compte des interactions spatiales entres les pixels doit permettre d'améliorer l'adéquation du modèle aux images. C'est l'objectif de la « modélisation couple », proposée initialement par W. Pieczynski et ses co-auteurs dans le cadre de champs de Markov [151] et adaptée aux chaînes [146, 66] et aux arbres [144] de Markov. Le modèle de *Chaîne de Markov Couple* (CMCc) considère directe-

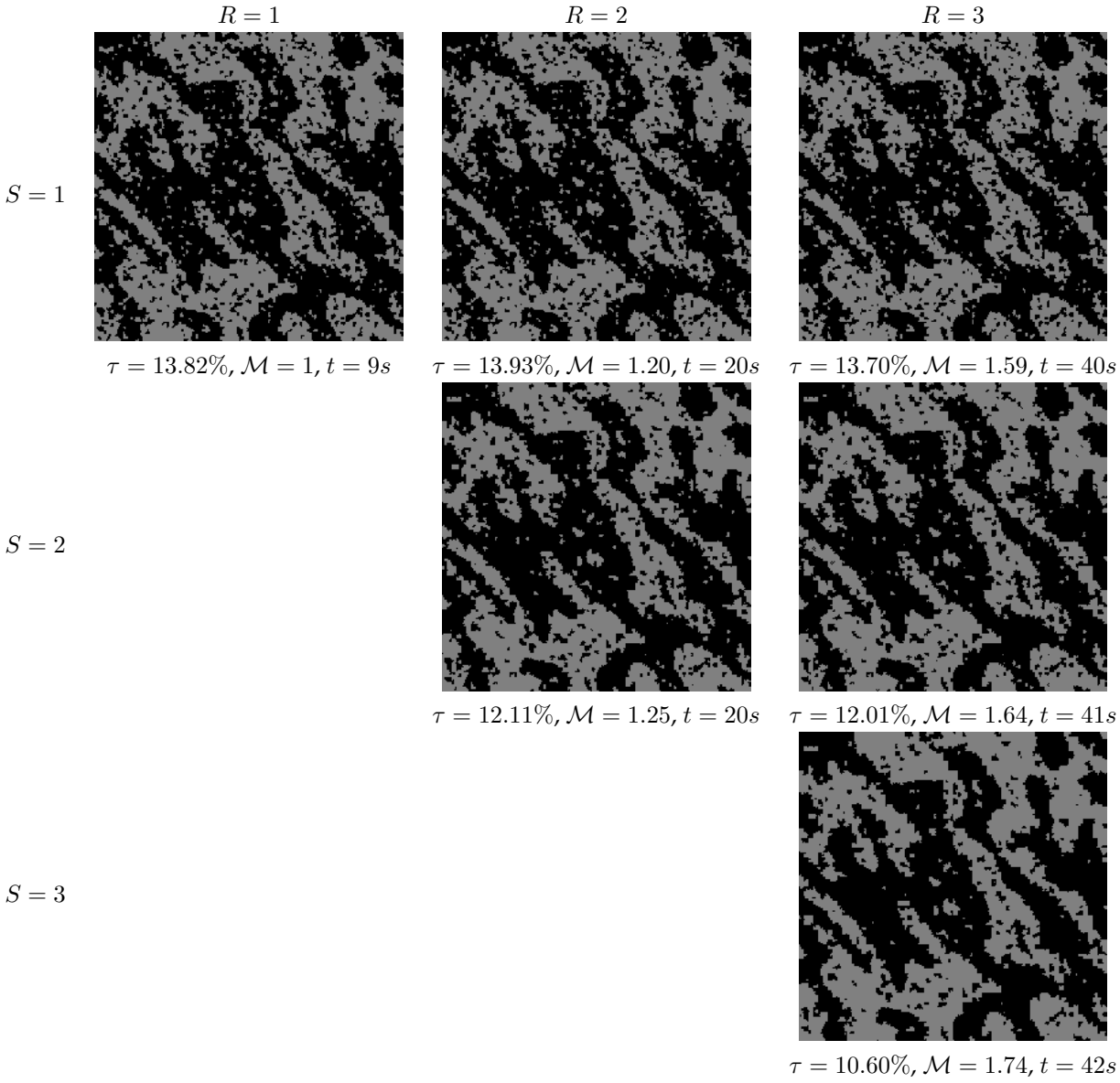



Figure 2.12 – Résultats de la segmentation non supervisée de l'image de la figure 2.11 pour différentes valeurs de $R$ et $S$. $\tau$ est le taux d'erreur. $\mathcal{M}$ est un indice sur la quantité de mémoire nécessaire à l'estimation de paramètres par rapport au modèle $\text{CMC}_{1-1}$. $t$ est le temps de calcul exprimé en secondes. Résultats extraits de [11].

ment la markovianité du couple des processus observé $\boldsymbol{Y}$ et caché $\boldsymbol{X}$, et non plus uniquement la markovianité du seul processus caché. Dans ce modèle, $\boldsymbol{X}$ n'est donc plus nécessairement de Markov, mais la restauration bayésienne reste possible car $\boldsymbol{X}\,|\boldsymbol{Y}$ l'est toujours [111].

**Modèle des CMCc–** En notant $Z_n = (X_n, Y_n)$, $\boldsymbol{Z} = (Z_1, \cdots, Z_n, \cdots, Z_N)$ est appelé « processus couple » associé à $\boldsymbol{X}$ et $\boldsymbol{Y}$. Le processus $\boldsymbol{Z}$ est une chaîne de Markov couple si sa distribution s'écrit de la manière suivante

$$p(\boldsymbol{z}) = p(z_1)\ p(z_2\,|z_1)\ \cdots\ p(z_N\,|z_{N-1})\,.$$

Nous considérons uniquement les CMCc stationnaires pour lesquelles $p(z_n, z_{n+1})$ ne dépend pas

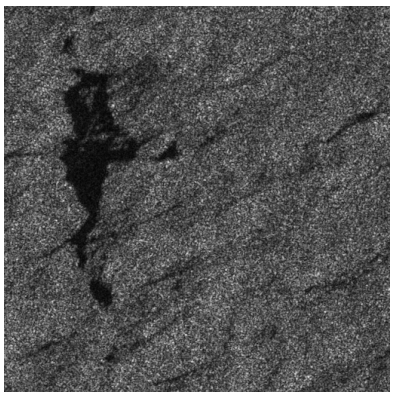
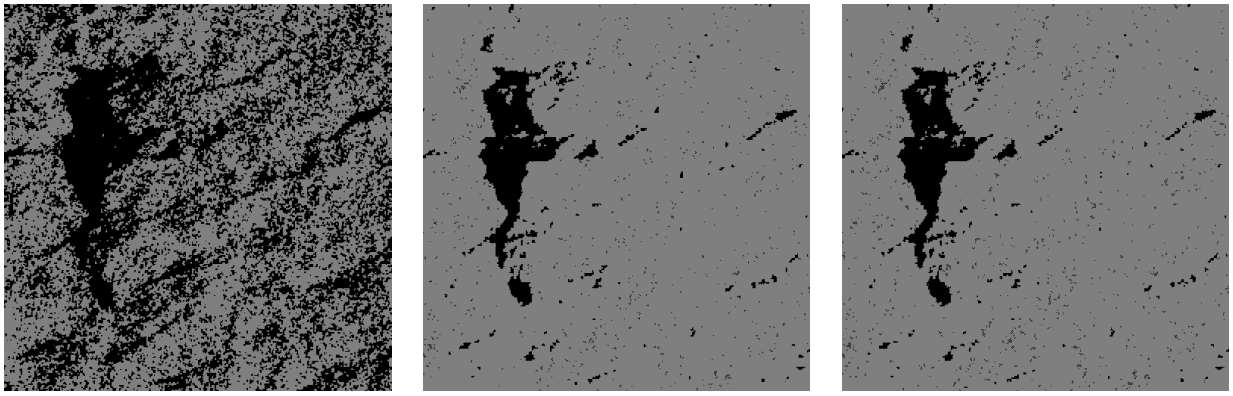


Figure 2.13 – Extrait (512 $\times$ 512) d'une image ERS-RSO d'une nappe de pétrole au large des côtes Égyptiennes, ©ESA.



(a) $\text{CMC}_{1-1}$ (b) $\text{CMC}_{2-2}$ (c) $\text{CMC}_{3-3}$

Figure 2.14 – Résultats de la segmentation non supervisée de l'image de nappe de pétrole de la figure 2.13 pour différentes valeurs de $R$ et $S$. Résultats extraits de [53].

de $n$. La distribution d'une CMCc est donc caractérisée par

$$p\left(z_1, z_2\right) = p\left(x_1, y_1, x_2, y_2\right) = p\left(x_1, x_2\right)\ p\left(y_1, y_2 \,|x_1, x_2\,\right), \quad \forall z_1, z_2 \in \Omega \times \mathbb{R}. \tag{2.23}$$

La distribution de $\boldsymbol{Z}$ est entièrement caractérisée par les $K^2$ probabilités $p\left(x_1, x_2\right)$ et les $K^2$ lois d'attache aux données bi-dimensionnelles $p\left(y_1, y_2 \,|x_1, x_2\,\right) = f_{x_1,x_2}(y_1, y_2)$. La stricte généralité du modèle des CMCc sur les CMCs est observable en comparant (2.23) avec son homologue (2.6) page 18 ; l'indépendance conditionnelle s'exprimant par $f_{x_1,x_2}(y_1, y_2) = f_{x_1}(y_1)\ f_{x_2}(y_2)$. À partir de l'eq. (2.23), la loi initiale d'une CMCc s'exprime comme un mélange de densités

$$p\left(z_1\right) = \sum_{x_2 \in \Omega} p\left(x_1, x_2\right)\ f_{x_1,x_2}(y_1), \tag{2.24}$$

et sa densité de transition s'écrit alors

$$p\left(z_2 \,|z_1\,\right) = \frac{p\left(z_1, z_2\right)}{p\left(z_1\right)}. \tag{2.25}$$

Le schéma de la figure 2.15(a) représente la structure de dépendance entre les processus $\boldsymbol{X}$ et $\boldsymbol{Y}$ d'un modèle CMCc. En comparant ce schéma à celui d'une CMC, *cf.* figure 2.2 page 19, on

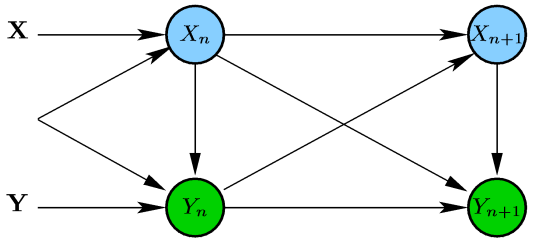



Figure 2.15 – Représentation de la structure de dépendance d'une chaîne de Markov couple.

peut mesurer la plus grande richesse d'interactions entre les observations et les classes du modèle CMCc.

**Classification MPM–** Pour segmenter l'image au sens du critère MPM, nous avons adapté l'algorithme de Baum et les probabilités *avant-arrière* au contexte des CMCc. En redéfinissant les probabilités *arrières* de l'eq. (2.8) de la manière suivante

$$\beta_n(k) = \frac{p\left(\boldsymbol{y}_{n+1:N}\,|X_n = k, Y_n\right)}{p\left(\boldsymbol{y}_{n+1:N}\,|\boldsymbol{y}_{1:n}\right)} = \frac{1}{S_{n+1}}\,\frac{p\left(\boldsymbol{y}_{n+1:N}\,|X_n = k, Y_n\right)}{p\left(\boldsymbol{y}_{n+2:N}\,|\boldsymbol{y}_{1:n+1}\right)},$$

nous vérifions toujours $\xi_n(k) = \alpha_n(k)\,\beta_n(k)$ [15], ce qui permet de calculer le critère Bayésien du MPM. Les formules de récurrence permettant le calcul rapide de ces probabilités s'écrivent

- Initialisation ($n = 1$) :

$$\alpha_1(k) = \frac{p\left(x_1 = k, y_1\right)}{\sum_{l=1}^{K} p\left(x_1 = l, y_1\right)}, \quad \forall k \in \Omega.$$

- Induction ($1 \leq n < N - 1$) :

$$\alpha_{n+1}(k) = \frac{1}{S_{n+1}} \sum_{l=1}^{K} \alpha_n(l)\, p\left(x_{n+1} = k, y_{n+1}\,|x_n = l, y_n\right), \quad \forall k \in \Omega.$$

pour les probabilités *avant* et

- Initialisation ($n = N$) :

$$\beta_N(k) = \frac{1}{K}, \quad \forall k \in \Omega.$$

- Induction ($1 \leq n < N$) :

$$\beta_n(k) = \frac{1}{S_{n+1}} \sum_{l=1}^{K} \beta_{n+1}(l)\, p\left(x_{n+1} = l, y_{n+1}\,|x_n = k, y_n\right), \quad \forall k \in \Omega.$$

pour les probabilités *arrière*.

L'expression des probabilités conjointes *a posteriori* $\psi_n(k,l) = p\left(X_n = k, X_{n+1} = l\,|\boldsymbol{y}\right)$ s'écrit :

$$\psi_n(k,l) \propto \alpha_n(k)\, p\left(x_{n+1} = l, y_{n+1}\,|x_n = k, y_n\right)\, \beta_{n+1}(l),$$

15. — La démonstration est similaire à celle présentée page 20 mais en écrivant $p\left(\boldsymbol{y}\,|X_n = k\right) = p\left(\boldsymbol{y}_{1:n}\,|X_n = k\right)\, p\left(\boldsymbol{y}_{n+1:N}\,|y_n, X_n = k\right)$ au lieu de $p\left(\boldsymbol{y}\,|X_n = k\right) = p\left(\boldsymbol{y}_{1:n}\,|X_n = k\right)\, p\left(\boldsymbol{y}_{n+1:N}\,|X_n = k\right)$.

avec $\sum_{k,l\in\Omega} \Psi_n(k,l) = 1$.

**Remarque 2.7 (algorithme de Viterbi pour les CMCc)** *Dans [66], nous avons proposé un équivalent de l'algorithme de Viterbi pour le critère MAP dans le contexte des CMCc. Cet algorithme est démontré de manière rigoureuse dans [21].*

**Estimation des paramètres avec ECI–** À partir de l'équation (2.23), les paramètres à estimer consistent en les $K^2$ probabilités $p\,(x_1, x_2)$ et les $K^2$ lois d'attache aux données $f_{x_1,x_2}(y_1, y_2)$.

- **Ré-estimation des paramètres associés à $\boldsymbol{X}$–** Selon le principe d'ECI, les probabilités conjointes *a priori* $p\,(x_n, x_{n+1})$ du processus stationnaire $\boldsymbol{X}$ (mais non markovien) peuvent être ré-estimées de la façon suivante :

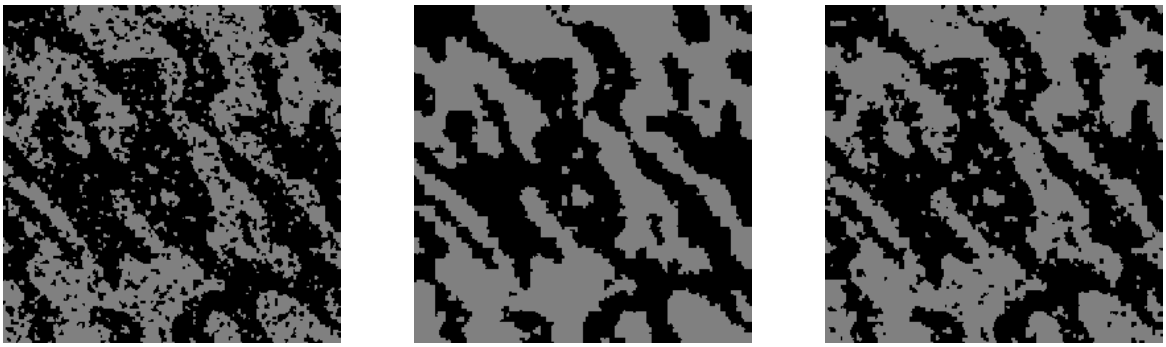
$$E\left[\hat{p}(k,l)|\boldsymbol{y}, \Phi^{[p]}\right] \text{ donne : } \quad p^{[p]}(k,l) = \sum_{n=1}^{N-1} \psi_n^{[p]}(k,l), \quad \forall k,l \in \Omega.$$

- **Ré-estimation des paramètres des lois d'attache aux données–** Les paramètres des lois d'attache aux données s'estiment en simulant $\boldsymbol{X}$ conditionnellement à $\boldsymbol{Y}$ dont la loi est celle d'une chaîne de Markov non-homogène avec pour transitions

$$\begin{aligned} \tilde{a}_n^{[p]}(k,l) &= p\,(X_{n+1} = l\,|X_n = k, \boldsymbol{y}\,) = \frac{\Psi_n(k,l)}{\xi_n(k)} \\ &\propto p\,(k,l)\ f_{k,l}(y_n, y_{n+1})\ \beta_{n+1}(l). \end{aligned}$$

Pour chaque itération $p$ de l'algorithme ECI, nous sommes confrontés à l'estimation de $K^2$ lois $f_{x_1,x_2}(y_1, y_2)$ bi-dimensionnelles, à partir des données complétées $(\boldsymbol{y}, \boldsymbol{x}^{[p]})$. Dans le cas d'un mélange gaussien, les matrices de corrélation et les moyennes des lois s'estiment très classiquement. Lorsque l'hypothèse de gaussianité n'est plus vérifiée, on est confronté, tout comme dans le paragraphe 2.2 sur les CMC vectorielles, au choix d'un modèle statistique adapté pour décrire les liens entre les observations d'une même classe. Les deux options de l'analyse multivariée et de la théorie des copules restent possibles :

> ❶ **Cas non gaussien / ACP –** Dans [66], nous avons proposé une solution de type ACP qui consiste à décorréler les données avant d'estimer les lois des composantes (parmi la famille des lois Beta, Gamma, ... ou dans le système de Pearson). Nous avons abouti a une expression similaire à l'eq. (2.18). Comme cette spécification des lois multidimensionnelles ne nous permet pas d'accéder directement aux marges, nous avons eu recours a des intégrations numériques pour le calcul de $p\,(z_1)$ et $p\,(z_2)$, *cf.* equations (2.24) et (2.25), nécessaires à l'échantillonnage de $\boldsymbol{X}$.



(a) $\tau = 13.80\%$, $t = 67s$ (b) $\tau = 7.60\%$, $t = 83s$ (c) $\tau = 10.04\%$, $t = 70s$

Figure 2.16 – Résultats de segmentation non supervisée de l'image simulée de la figure 2.11 par le modèle CMCc. (a) Copule produit ; (b) copule gaussienne ; (c) copule logistique de Gumbel.

**❷ Cas non gaussien / Copule –** Les copules permettent de résoudre ce problème [21]. La stationnarité du processus $\boldsymbol{Z}$ est équivalente à $p(z_1) = p(z_2)$, ce que nous pouvons encore écrire, grâce à l'équation (2.24),

$$\sum_{x_2 \in \Omega} p(x_1, x_2)\ f_{x_1,x_2}(y_1) = \sum_{x_1 \in \Omega} p(x_1, x_2)\ f_{x_1,x_2}(y_2),$$

où $f_{x_1,x_2}(y_1)$ et $f_{x_1,x_2}(y_2)$ représentent respectivement les marges gauche et droite de $f_{x_1,x_2}(y_1, y_2)$. Pour satisfaire cette contrainte, N. Brunel propose d'utiliser les copules pour représenter la densité jointe et d'imposer les conditions suivantes

$$f_{x_1,x_2}(y_1) = f_{x_1,x_2}(y_2) \quad \text{et} \quad p(x_1, x_2) = p(x_2, x_1).$$

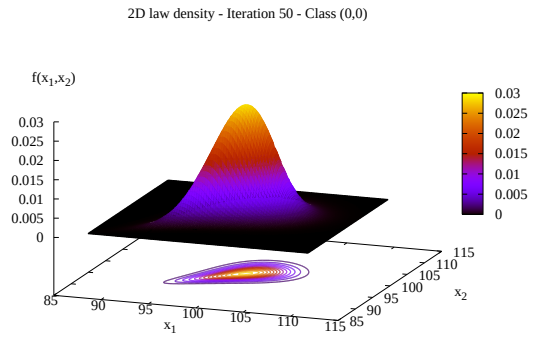
L'estimation des lois d'attache aux données du modèle CMCc requiert donc l'estimation de $K^2$ copules et de $K^2$ marges (issues de n'importe quel modèle, paramétrique ou non) [a]. Notons que le modèle CMCa, pour lequel $\boldsymbol{Z}$ *et* $\boldsymbol{X}$ sont de Markov, impose $p(x_1, x_2 | y_1) = p(x_1 | y_1)$ et $p(x_1, x_2 | y_2) = p(x_2 | y_2)$, ne requiert donc que l'estimation de $K$ marges (et toujours $K^2$ copules).

*a.* — Attention, il n'y a pas $2\,K^2$ marges, du fait de la symétrie imposée par la stationnarité.

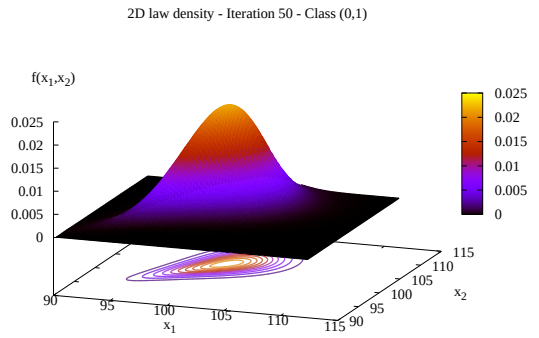
**Illustration 1 : Segmentation d'une image simulée–** Pour illustrer cet algorithme, nous avons segmenté la même image que celle utilisée pour le modèle CMC$_{R-S}$, *cf.* figure 2.11 page 41. Les résultats de segmentation non supervisée sont présentés dans la figure 2.16, pour les copules produit, gaussienne et logistique de Gumbel (marges gaussiennes, 50 itérations ECI, critère MPM).

Comme attendu étant donné la façon dont l'image bruitée à été construite, c'est la copule gaussienne qui procure le meilleur taux d'erreur de segmentation ($\tau = 7.60\%$). On peut aussi remarquer que celui-ci améliorer sensiblement les taux d'erreur du modèle d'ordre supérieur CMC$_{R-S}$, quelque soit les valeurs de $R$ et $S$ (rappel : $\tau_{\text{CMC}_{3-3}} = 10.60\%$). La copule logistique de Gumbel permet de mieux prendre en compte la corrélation du bruit que la copule produit [16] ($\tau = 10.04\%$), sans toutefois rivaliser avec la copule gaussienne ($\tau = 13.80\%$). À titre d'exemple, nous avons

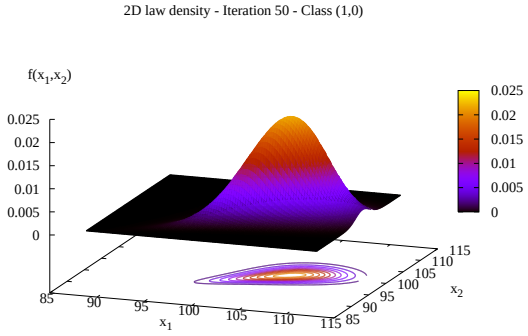
16. — qui, nous le rappelons, correspond à l'indépendance



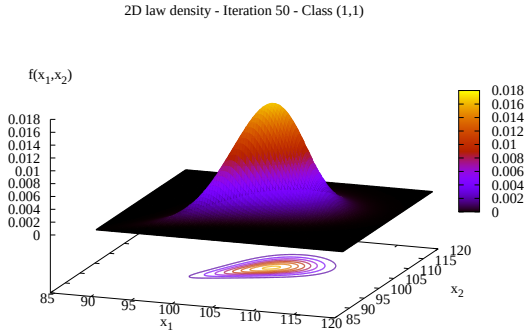

(a) Classes $x_1 = 1$ et $x_2 = 1$




(b) Classes $x_1 = 1$ et $x_2 = 2$




(c) Classes $x_1 = 2$ et $x_2 = 1$




(d) Classes $x_1 = 2$ et $x_2 = 2$

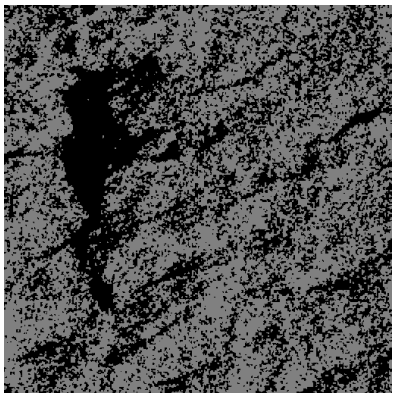
Figure 2.17 – Lois bidimensionnelles estimées à l'aide de la copule logistique de Gumbel avec marges gaussiennes, aboutissant à l'image segmentée 2.16(c).

rapporté dans la figure 2.17 les quatre lois 2D estimées par ECI pour le cas de la copule logistique de Gumbel.

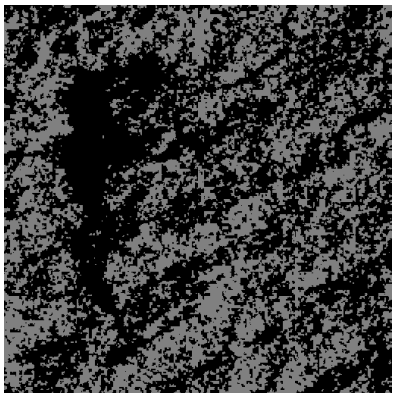
**Illustration 2 : Segmentation d'une image radar de nappe d'hydrocarbure–** Nous avons segmenté l'image radar de nappe de pétrole de la figure 2.13 pour quatre copules différentes et des marges gaussiennes, *cf.* figure 2.18. Les résultats montrent clairement que les copules gaussienne et Student procurent des résultats de segmentation bien plus homogènes que les copules produit et logistique de Gumbel. Sur cette image réelle, les résultats sont légèrement meilleurs que ceux obtenus avec le modèle $\text{CMC}_{2-2}$, *cf.* figure 2.14(b). Des travaux sur la segmentation d'images radar de rizière [65, 64] confirment le bon comportement de la méthode.

**Remarque 2.8 (modèle couple et CMCs)** *Soit $\boldsymbol{Z} = (\boldsymbol{X}, \boldsymbol{Y})$ une chaîne de Markov couple. Les trois propositions sont équivalentes :*

- $\boldsymbol{X}$ *est une chaîne de Markov,* i.e. $(\boldsymbol{X}, \boldsymbol{Y})$ *est une chaîne de Markov cachée,*
- $p(y_n | x_{n-1}, x_n) = p(y_n | x_n)$ *[146],*
- $p(y_n | \boldsymbol{x}) = p(y_n | x_n)$ *[148].*

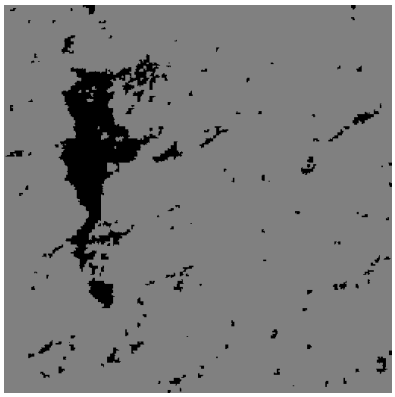


(a) Copule produit



(b) Copule log. de Gumbel

(c) Copule gaussienne



(d) Copule de Student

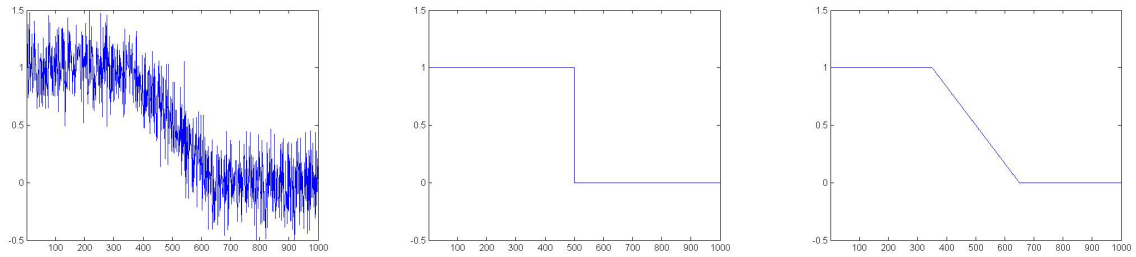
Figure 2.18 – Segmentation non supervisée de l'image radar de nappe de pétrole présentée page 43, pour différentes copules (marges gaussiennes).

**Remarque 2.9 (modèles triplets)** *Le modèle des* Chaînes de Markov Triplet *(CMT), introduit récemment par W. Pieczynski [145, 150], constitue une généralisation des chaînes de Markov couples, et augmente les capacités de modélisation des modèles CMC*[17]*. Ce modèle introduit une troisième chaîne aléatoire* $\boldsymbol{U} = \{U_1, \cdots, U_N\}$*, appelé* processus auxiliaire *ou* latent*, de telle manière que le processus triplet* $\boldsymbol{T} = (\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{Y})$ *est une chaîne de Markov appelée* CMT, i.e. $p(\boldsymbol{t}_n \,|\boldsymbol{t}_1, \cdots, \boldsymbol{t}_{n-1}) = p(\boldsymbol{t}_n \,|\boldsymbol{t}_{n-1})$*. Cette extension permet de gérer la non stationnarité des paramètres (en particulier les modèles semi-markoviens [107] permettant de gérer des bruits à longue dépendance) et peut être interprétée au travers de la théorie de l'évidence [145]. Une version floue, au sens du modèle que nous présentons, de ce modèle a également été proposée [163].*

## 2.4 Segmentation imprécise par CMC floue (thèse de C. Carincotte)

La restauration bayésienne dans le contexte des CMCs est une technique de classification très robuste, comme nous avons pu le souligner à plusieurs reprises dans ce document. Elle peut cependant s'avérer peu satisfaisante dans certaines situations. Il existe en effet des contextes pour lesquels une segmentation thématique de l'image ne permet pas de représenter la réalité intrinsèquement floue de la scène observée et/ou les phénomènes physiques flous liés à l'acquisition

17. — Ce modèle a aussi été adapté au cadre des champs de markov, *cf.* [7].



(a) Signal bruité (b) Segmentation « dure » (c) Segmentation « floue »

Figure 2.19 – Objectif de la segmentation statistique floue : trouver (c) plutôt que (b) à partir de l'observation (a). Extrait de [28].

des images. Le flou est bien souvent confondu avec le bruit des observations, ce qui déstabilise l'estimation des paramètres et conduit à des résultats de segmentation qui occultent la richesse de l'information présente dans l'image observée.

L'objectif de ce travail, résumé par la figure 2.19, concerne donc l'introduction d'une extension floue au modèle original des CMCs et constitue, pour l'essentiel, le travail de thèse de Cyril Carincotte [28, 27]. Ce nouveau modèle permet de caractériser, conjointement et de manière dissociée, le bruit et le flou associés aux observations en enrichissant la mesure d'incertitude du modèle stochastique des chaînes de Markov par une mesure d'imprécision. À la différence des modèles purement flous qui considèrent que tous les pixels sont nécessairement flous [131, 3], nous adoptons le point de vue développé dans le contexte de la segmentation aveugle et contextuelle [22], d'une part, et dans celui des champs de Markov [164], d'autre part.

Hormis le flou de mouvement et le flou de défocalisation que l'on connaît bien dans le cadre de l'imagerie vidéo, ce type de dégradé peut apparaître dans les images de télédétection suite à l'observation de phénomènes intrinsèquement flous :

- **La nature de la scène elle-même** Les images météorologiques de nuage constituent une bonne illustration d'images intrinsèquement floues. Dans ce cas, le cœur du nuage et le ciel constituent les deux classes de l'image, et le gradient de densité de vapeur d'eau entre les deux constitue la zone floue.
- **L'effet de surface/volume partiel** Cet effet, bien connu en imagerie médicale mais que l'on observe aussi dans les images de télédétection, est lié à la résolution spatiale : un pixel de l'image est le résultat de l'intégration sur une surface non nécessairement homogène (composée de plusieurs constituants mélangés dans des proportions différentes et inconnues) du signal réfléchi par cette même surface.

Dorénavant, nous qualifierons de classe « dure », l'ensemble des pixels correspondant à une classe thématique donnée (terre, mer, nuage…) et de classe « floue » l'ensemble des pixels correspondant à un dégradé de l'image entre deux classes dures (frontière terre/mer, dégradé nuage/ciel…). La classe floue correspond alors à un mélange de deux classes dures dans des proportions variées, et cette proportion peut être interprétée comme le degré d'appartenance à la classe 0. Pour des raisons de simplification de l'exposé, nous considérerons par la suite le cas de deux classes dures uniquement. Bien que cela ne soit pas une limitation théorique du modèle, le fait de considérer un nombre de classes dures supérieur à deux engendre des temps de calculs qui peuvent devenir rédhibitoire. L'application de détection de changements entre images RSO, présentée à la fin de ce

paragraphe, illustre un cas particulier d'un modèle de CMC floue à trois classes dures.

### 2.4.1 Chaîne de Markov floue

En terme de classification les pixels flous présentent une ambiguïté et nous cherchons non plus à obtenir une réponse dure (classe « 0 » ou « 1 ») mais une réponse graduée par une valeur dans l'intervalle $[0, 1]$. Dans le contexte des CMCs, ceci est réalisable en introduisant une mesure d'imprécision. Dans cette modélisation floue, chaque variable aléatoire $X_n$ est à valeur dans $\Omega_f = [0, 1]$ et nous adoptons la notation suivante :

$$X_n = u \text{ avec } \begin{cases} u = 0 \text{ si le pixel appartient à la classe dure } 0, \\ u \in ]0, 1[ \text{ si le pixel est flou,} \\ u = 1 \text{ si le pixel appartient à la classe dure } 1. \end{cases}$$

Ainsi, la chaîne $\boldsymbol{X}$ est un processus à valeur dans l'ensemble $\Omega_f^N = [0, 1]^N$. La nouvelle approche requiert donc la définition des probabilités de la chaîne sur $\Omega_f$. La loi de chaque $X_n$ est donnée par une densité qui inclue à la fois des composantes discrètes (classes « dures ») et une composante continue (classe « floue »). En notant $\delta_0$ et $\delta_1$ les mesures de Dirac en $0$ et $1$, en notant $\upsilon$ la mesure de Lebesgue sur $]0, 1[$, et en prenant $\nu = \delta_0 + \upsilon + \delta_1$ comme mesure sur $\Omega_f$, la distribution de $X_n$ est définie par une densité $h$ sur $[0, 1]$ selon la mesure $\nu$ :

$$p\left(X_n = u\right) = \begin{cases} \pi(0) = h\left(u = 0\right), \\ \pi(u) = h\left(u \in [a, b]\right) = \displaystyle\int_a^b h\left(v\right)\, d\nu\left(v\right), \text{ avec } 0 < a \leq b < 1, \\ \pi(1) = h\left(u = 1\right). \end{cases}$$

La densité $h$ vérifie $h(0) + h(1) + \int_0^1 h(u)\, d\nu(u) = 1$. La loi initiale de la chaîne, qui correspond à la distribution de $X_1$, est donc maintenant définie par la fonction $h$, et non plus par un vecteur comme dans le cas classique. De façon comparable, les transitions d'un état $u$ vers un état $v$ sont définies par une fonction $a(u, v)$, de support le carré unité.

La suite de cette section présente les algorithmes nécessaires à la mise en œuvre de ce modèle pour la segmentation non supervisée d'images à l'aide de la méthode d'estimation ECI. Fondamentalement, la démarche générale présentée dans la section 2.1 demeure inchangée. Ainsi, pour ne pas surcharger la présentation, nous nous limitons à décrire les concepts spécifiques et les aspects numériques propres au modèle.

### 2.4.2 Estimation des paramètres du modèle des CMCs floue

L'estimation des paramètres du modèle repose sur l'adaptation des probabilités *avant-arrière*, définies en page 19 pour le modèle original, au nouveau contexte :

$$\begin{aligned} \alpha_n(u) &= p\left(X_n = u \,|\boldsymbol{y}_{1:n}\right), \quad \forall u \in \Omega_f \\ \beta_n(u) &= \frac{p\left(\boldsymbol{y}_{n+1:N} \,|X_n = u\right)}{p\left(\boldsymbol{y}_{n+1:N} \,|\boldsymbol{y}_{1:n}\right)}, \quad \forall u \in \Omega_f. \end{aligned}$$

La procédure récursive de calcul s'écrit alors

- Probabilités *avant* ($1 \leq n < N-1$) :

$$\alpha_{n+1}(u) \propto f_u(y_{n+1}) \int_{v\in\Omega_f} \alpha_n(v)\, a(v,u)\, d\nu(v), \quad \forall u \in \Omega_f.$$

- Probabilités *arrière* ($1 \leq n < N$) :

$$\beta_n(u) \propto \int_{v\in\Omega_f} a(u,v)\, f_v(y_{n+1})\, \beta_{n+1}(v)\, d\nu(v), \quad \forall u \in \Omega_f.$$

**Remarque 2.10 (discrétisation)** *Les intégrales ne peuvent pas être calculées de manière analytique. D'un point de vue algorithmique, nous avons recours à une intégration numérique par discrétisation de l'intervalle* $[0,1]$. *On obtient ainsi* $F$ *classes floues « discrètes », dont la valeur floue correspond à la valeur médiane du sous-intervalle considéré. Par exemple,* $F = 2$ *implique* $u \in \{0.25, 0.75\}$ *et* $F = 3$ *implique* $u \in \{0.165, 0.5, 0.825\}$. *Plus* $F$ *est grand, plus l'estimation des paramètres est précise, au détriment d'un temps de calcul qui s'en trouve augmenté en conséquence. Ainsi, la pente régulière de la figure 2.19(c) est transformée en un escalier dont le nombre de marches (et donc la précision) est fixé par* $F$.

Les probabilités marginales et conjointes *a posteriori* définies par les équations (2.11) et (2.14) s'écrivent dans ce nouveau contexte

$$\begin{aligned} \xi_n(u) &= \int_{v\in\Omega_f} \psi_n(u,v)\, d\nu(v) = \alpha_n(u)\, \beta_n(u), \quad \forall u \in \Omega_f, \\ \psi_n(u,v) &\propto \alpha_n(u)\, \beta_{n+1}(v)\, a(u,v)\, f_v(y_{n+1}), \quad \forall u,v \in \Omega_f. \end{aligned}$$

La matrice de transition de la chaîne de Markov floue non-homogène *a posteriori*, permettant la simulation de $\boldsymbol{X}$ conditionnellement à $\boldsymbol{Y}$ et l'estimation des paramètres des lois d'attache aux données dans le cadre d'ECI, s'écrit

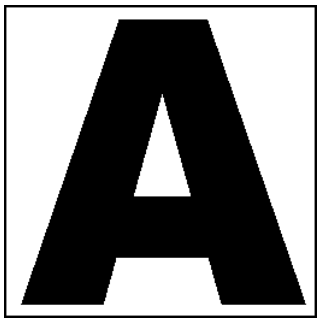
$$\tilde{a}_n(u,v) \propto \beta_{n+1}(v)\, a(u,v)\, f_v(y_{n+1}), \quad \forall u,v \in \Omega_f. \tag{2.26}$$

Dans le cadre discret fixé par la remarque 2.10, l'estimation des paramètres de la chaîne de Markov floue $\pi(u)$ et $a(u,v)$ ne pose pas de difficulté particulière, en adaptant les formules de l'équation (2.15) avec les expressions données ci-dessus. En ce qui concerne la ré-estimation des paramètres des lois d'attache aux données, il faut dissocier les deux classes dures de la classe floue. Pour simplifier les explications, nous supposons dans un premier temps que les répartitions des observations des classes 0 et 1 peuvent être modélisées par des lois gaussiennes de densité $\mathcal{N}(\mu_1^{(0)}, \mu_2^{(0)})$ et $\mathcal{N}(\mu_1^{(1)}, \mu_2^{(1)})$. L'estimation de ces deux lois d'attache aux données ne pose pas de difficulté, en utilisant des estimateurs à partir des données complétées par la simulation de $\boldsymbol{x}\,|\boldsymbol{y}$ selon l'équation (2.26). Pour caractériser la loi d'attache aux données de la classe floue, nous considérons deux variables aléatoires indépendantes $X^{(0)}$ et $X^{(1)}$ correspondant aux deux classes dures. En considérant qu'un pixel flou contient une « proportion » $u$ de classe 1 (et par conséquent une proportion $1-u$ de classe 0), la variable aléatoire associée à la classe floue s'écrit

$$X^{(u)} = (1-u)\, X^{(0)} + u\, X^{(1)}. \tag{2.27}$$

Nous montrons alors que la loi de la classe floue est une gaussienne de paramètres

$$\mu_1^{(u)} = (1-u)\, \mu_1^{(0)} + u\, \mu_1^{(1)}, \quad \mu_2^{(u)} = (1-u)^2\, \mu_2^{(0)} + u^2\, \mu_2^{(1)} \tag{2.28}$$



(a) Taille : $256 \times 256$

(b)

| | **Image 1** | | **Image 2** | | **Image 3** | |
|---|---|---|---|---|---|---|
| **Classe** | $\mu_1$ | $\mu_2$ | $\mu_1$ | $\mu_2$ | $\mu_1$ | $\mu_2$ |
| $\varepsilon = 0$ | 50 | 30 | 50 | 30 | 50 | 60 |
| $\varepsilon = 1$ | 110 | 45 | 75 | 45 | 75 | 85 |

(c)

Figure 2.20 – (a) Image originale et (b) image floue correspondante. Le tableau (c) présente les paramètres des bruits gaussiens utilisés dans la segmentation des images floues de la figure 2.21.

Ainsi, l'intégration d'une mesure floue sur l'espace des états de la chaîne de Markov revient donc à introduire une mesure d'imprécision sur l'espace des moments associé à chaque état. Nous obtenons une répartition linéaire du moment d'ordre $1$ associé à la loi d'attache aux données floue. Pour le moment d'ordre $2$, on peut constater que la variance de la classe floue peut devenir inférieure à celles des deux classes dures. D'autre part, d'autres expressions de $X^{(u)}$ sont possibles donnant en conséquence une répartition du flou différente sur l'espace des moments, *p. ex.*

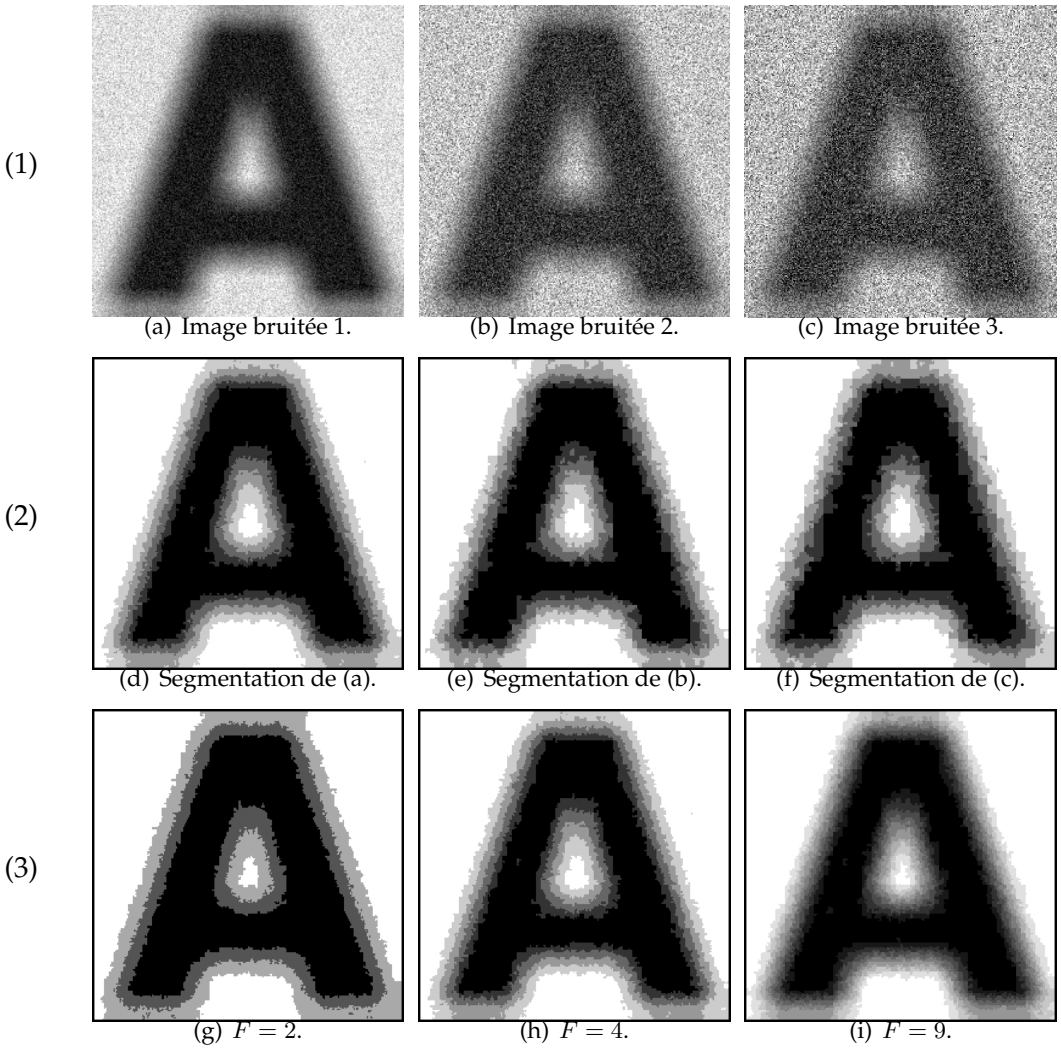
$$X^{(u)} = \frac{X^{(0)} + X^{(1)}}{2} + 4\,(X^{(1)} - X^{(0)})\,(u - 0.5)^3.$$

Ce type de répartition permet de privilégier la présence de pixels flous dans la segmentation. Une autre répartition permettra au contraire de favoriser l'existence de pixels durs. Des explications complémentaires sont consultables dans le manuscrit de C. Carincotte [28].

**Remarque 2.11 (mélange généralisé)** *Dans le cas non gaussien, c'est-à-dire lorsque les lois d'attache aux données des classes dures ne sont pas gaussiennes, les expressions des paramètres de la classe floue peuvent être déterminés à partir de ceux des deux classes dures, selon l'équation* (2.27)*. Ainsi, si nous faisons l'hypothèse que les deux classes dures appartiennent au système de Pearson, il suffit d'exprimer les quatre premiers moments de la classe floue en fonction de ceux des classes dures pour déterminer exactement la famille et les paramètres de la loi d'attache aux données floue. Il est intéressant de noter que si les classes dures sont distribuées selon des lois gamma, la classe floue sera quand-à-elle distribuée selon une loi beta. Une question se pose alors : est-il possible de trouver une répartition telle que la classe floue soit également de densité Gamma ? De manière plus générale, est-il possible de trouver une répartition qui donne une loi pour la classe floue qui se positionne, dans le diagramme de Pearson, sur le segment de droite liant les lois de deux classes dures quelconques ?*

L'extension vectorielle de ce modèle, sur la base des principes décrits dans la section 2.2, ne pose pas de difficulté particulière. Elle nous a permis de proposer un algorithme de détection et de caractérisation de nappes d'hydrocarbure à partir d'images multi-spectrales [25].

**Illustration : Segmentation d'une image floue simulée–** Pour cette expérience, nous avons utilisé l'image « $A$ » de la figure 2.20, que nous avons rendu floue par l'application d'un filtre moyenneur de taille $31$. Nous avons ensuite bruité l'image avec trois séries de bruits gaussiens dont les



(1) (a) Image bruitée 1. (b) Image bruitée 2. (c) Image bruitée 3.

(2) (d) Segmentation de (a). (e) Segmentation de (b). (f) Segmentation de (c).

(3) (g) $F = 2$. (h) $F = 4$. (i) $F = 9$.

Figure 2.21 – La ligne (1) présente les images bruitées obtenues à partir des paramètres du tableau 2.20(c). La ligne (2) montre les résultats de segmentation avec l'algorithme des CMCs floues (critère MPM) pour $F = 4$ sous-intervalles flous. La ligne (3) illustre le comportement de l'algorithme, sur l'image bruitée (a), pour un nombre croissant de sous-intervalles flous. Résultats extraits de [28].

paramètres pour les deux classes dures sont rapportés dans le tableau 2.20(c). Les paramètres des bruits pour les pixels flous sont calculés selon l'équation (2.28).

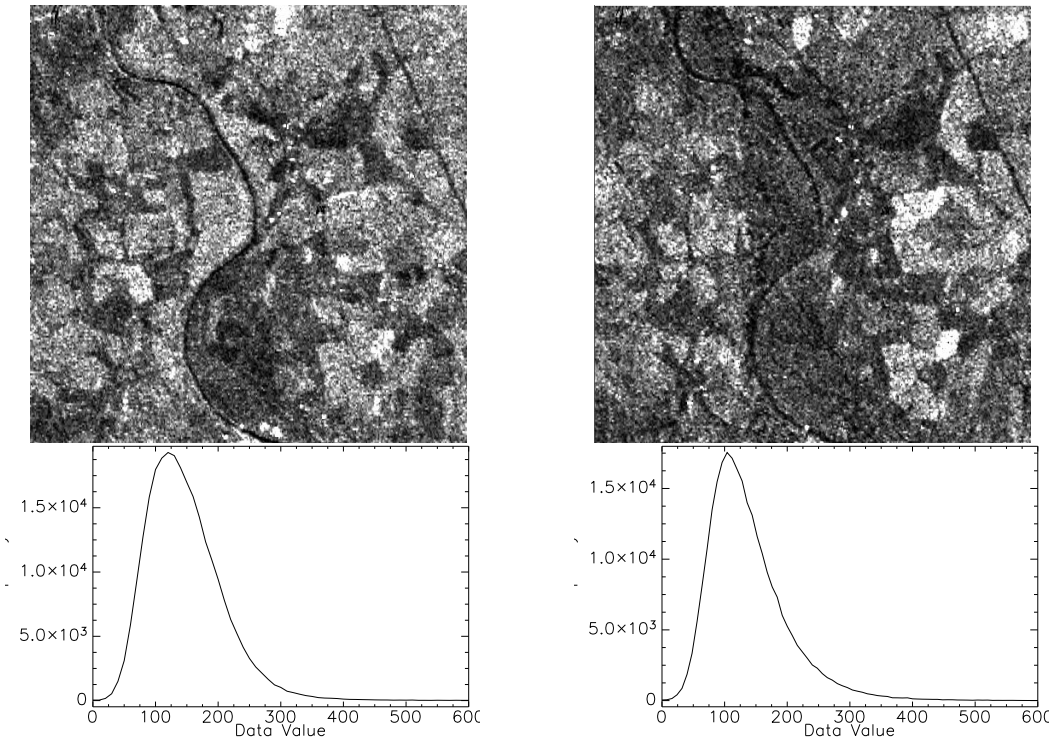
Les trois images floues et bruitées sont montrées dans la première ligne de la figure 2.21. La seconde ligne présente les résultats de segmentation de ces images (critère MPM, estimation par 70 itérations ECI et $F = 4$ sous-intervalles flous). Les résultats peuvent être comparés à l'image 2.20(b). Enfin, la troisième ligne présente les résultats de segmentation de l'image bruitée (a) pour $F = 2, 4$ et $9$ sous-intervalles flous. Plus le nombre de sous-intervalles flous est élevé, plus la restitution de la zone floue est visuellement satisfaisante. On peut également observer que, quel que soit le nombre de sous-intervalles flous considérés, la caractérisation des deux classes dures, c'est-à-dire le fond blanc de l'image et la lettre noire « $A$ », sont quasiment identiques ; il n'y a donc pas de confusion notoire entre les pixels bruités et les pixels bruités *et* flous. Une étude comparative de l'algorithme a été conduite avec la méthode de classification des k-moyennes floues [13], et un algorithme aveugle de classification statistique et flou [22], montrant la supériorité de l'approche considérée dans ce travail [28].

### 2.4.3 Détection floue de changements

Dans le cadre d'un travail publié dans [26], nous avons appliqué cet algorithme de segmentation pour détecter des changements survenus entre une paire d'images de télédétection de la même scène terrestre mais prises à deux instants différents (*avant*/*après* l'événement d'intérêt). La détection de changements correspond à la détection et à la localisation de zones ayant évoluées entre deux observations (ou plus) d'une même scène. Ces changements peuvent être de différents types, d'origine et de durée variées, ce qui permet de distinguer plusieurs familles d'applications :

- le *suivi de l'utilisation des sols*, qui correspond à la caractérisation du développement du tissu végétatif, ou à la détection des changements saisonniers de végétation ;
- la *gestion des ressources naturelles*, qui correspond entr'autres à la caractérisation du développement du tissu urbain, au suivi de la déforestation, … ;
- la *cartographie des dommages*, qui correspond à la localisation des dommages causés par des catastrophes naturelles lors d'une éruption volcanique, d'un raz de marée, d'un tremblement de terre, d'une inondation, …ou d'origine humaine.

Comme le montrent les récents états de l'art sur le sujet (*cf. p. ex.* [156]), un nombre relativement important de méthodes existent, mais la plupart d'entre-elles ont été développées pour des capteurs optique ou infra-rouge. Les études traitant de la détection de changements en imagerie radar sont bien moins nombreuses et plus récentes [6, 95, 132]. Il est cependant clair que les limites intrinsèques des capteurs optique ou infra-rouge, notamment leur dépendance aux conditions météorologiques et d'éclairement, sont particulièrement contraignantes et peu réalistes dans un cadre opérationnel. C'est par exemple le cas de la cartographie des dommages lors de catastrophes naturelles majeures (éruption, inondation, …) pour l'organisation des secours sur le terrain, situation dans laquelle la rapidité de réaction est primordiale. Dans ce contexte, l'utilisation de l'imagerie radar est donc une solution techniquement très intéressante, même si elle rend le problème de détection automatique de changements plus complexe. Les images radar sont, en effet, difficiles à comparer du fait des nombreuses dégradations et transformations des images liées à la géométrie d'acquisition (angles d'incidence des capteurs et sens de l'orbite) et au chatoiement résultant du phénomène de rétro-diffusion. L'introduction d'une mesure floue pour cette application permet



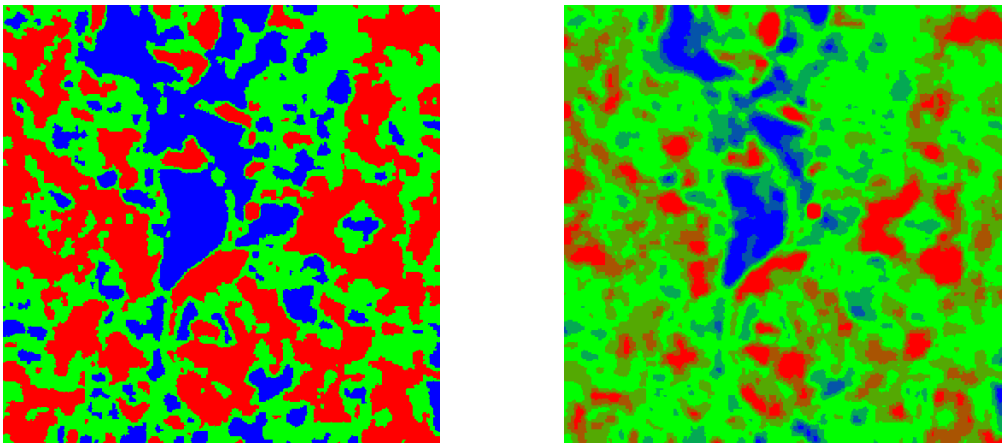

(a) 9 septembre 2000. (b) 21 octobre 2000.

Figure 2.22 – Images ERS-2 PRI (taille : $512 \times 512$) d'une région agricole près de Gloucester, Angleterre, avant (a) et pendant (b) une inondation. © ESA, distribution Eurimage.

de proposer non plus une réponse binaire (présence/absence de changements) mais une réponse graduée qui traduit l'imprécision sur le changement.

Pour détecter les changements entre l'*image avant* ($I_{t_1}$) et l'*image après* ($I_{t_2}$), nous avons appliqué l'algorithme de CMC floue décrit ci-dessus sur une *image de critère* calculée de la manière suivante : $I(x, y) = \ln(I_{t_2}(x, y)) - \ln(I_{t_1}(x, y))$. Ce critère a été choisi parmi d'autres pour son adéquation avec notre modèle et à la nature nature multiplicative du chatoiement. Pour réduire la sensibilité de ces détecteurs à la présence de bruit dans les images (chatoiement inhérent aux images radar), ce type d'opérateur n'est généralement pas appliqués pixel à pixel, mais sur des moyennes calculées sur une fenêtre de régularisation centrée autour du pixel d'intérêt. La détection de changements est alors considérée comme un problème de classification de l'image de critère en trois classes dures : deux classes sont utilisées pour caractériser les changements négatifs (classe « $-1$ ») et positifs (classe « $+1$ ») et la troisième classe pour l'absence de changement (classe « 0 »). Nous y avons associé deux mesures floues correspondant aux intervalles entre les classes dures : $]-1, 0[$ et $]0, +1[$.

**Illustration : Détection d'une inondation–** L'algorithme des CMCs floues a été appliqué pour caractériser l'évolution d'une inondation sur une zone agricole à partir des deux images ERS-2 de la figure 2.22. Ces images, géo-référencées et recalées (résolution $12.5m$) ont été acquises les 9 septembre et 21 octobre 2000 autour de la ville de Gloucester (Angleterre). Les secteurs inondés, qui apparaissent en noir sur l'image à cause de la réflectivité spéculaire de l'eau, sont identifiables le long de la rivière.



(a) CMC - $w = 15$ (b) CMC floue - $w = 15$

Figure 2.23 – Cartes de changements obtenues avec les modèles de CMC (a) et CMC floue (b) avec une modélisation non linéaire des deux classes floues et avec $F = 2$ pour les deux mesures floues. Résultats extraits de [26].

Du point de vue thématique, la difficulté de ces images provient du fait qu'au cours de l'inondation, plusieurs étendues d'eau ont été exposées au vent contrairement à d'autres. La friction à l'interface entre l'eau et le vent fait déplacer l'eau dans la direction du vent, et produit des changements de différents signes dans l'image de log-ratio. La figure 2.23 présente les résultats obtenus avec les modèles des CMCs (a) et CMCs floues (b) dans l'hypothèse de lois d'attache aux données gaussiennes. Les zones bleues, vertes et rouges présentent respectivement les classes dures de changements négatifs, de non-changement et de changements positifs. Les dégradés bleus et rouges correspondent aux $F = 2$ niveaux discrets de flou.

Dans ce cas de figure complexe, le modèle des CMCs produit un nombre très important de fausses alarmes par rapport aux CMCs floues. Les principales zones correspondant à l'inondation demeurent cependant facilement identifiables dans la figure (b). La prise en compte de l'imprécision associée aux changements semble donc avoir permis de faire la distinction entre les changements réellement survenus et les changements dus à la présence de vent. Néanmoins, sans « vérité-terrain » ou l'avis d'un expert, il est difficile de conclure sur l'adéquation des résultats, bien que les zones de changements obtenues par le modèle markovien flou semblent plus localisées et mieux en relation avec l'événement.

Ce travail sur la détection de changements entre images radar fait suite à une étude financée par le CNES en 2003 [59] dont les objectifs étaient de dresser un état de l'art des méthodes existantes et de tester un certain nombre d'algorithmes de segmentation sur des données réelles. L'activité sur la détection de changements s'est poursuivie dans le cadre de l'action *CACHANT* soutenu par le GDR-ISIS [68].

## 2.5 Conclusion & Perspectives

Mon activité de recherche en segmentation statistique d'images a consisté à enrichir le modèle classique des chaînes de Markov cachées pour augmenter ses capacités de modélisation. Ces extensions portent sur des modifications

Tableau 2.1 – Synthèse des modèles de chaînes de Markov développés (symbole ✓) et en cours (symbole ✗).

| | **Mélange généralisé** (non gaussien) | **Flou** (imprécision) | **Vectoriel** (multicomposantes) |
|---|---|---|---|
| **CMC** | ✓ | ✓ | ✓ |
| **CMCc** | ✓ | ✗ | |
| **CMC**$_{R-S}$ | ✓ | | |

- de la modélisation des données observées avec le modèle de *chaîne de Markov vectorielle* (paragraphe 2.2) et des données cachées avec le modèle de *chaîne de Markov floue* (paragraphe 2.4) ;
- de la topologie de la chaîne –et donc des hypothèses de dépendance statistique sous-jacentes–, aboutissant aux modèles appelés *chaîne de Markov d'ordre supérieur* (paragraphe 2.3.1) et *chaîne de Markov couple* (paragraphe 2.3.2).

La majeure partie de ces travaux a été réalisée dans le cadre de collaborations avec Wojciech Pieczynski et Grégoire Mercier, dans ce cadre de « projets inicitatifs » du GET [121, 123, 125] et avec Jordi Inglada, dans le cadre d'un contrat d'études « Recherche & Technologies » du CNES [59].

Pour tous ces modèles, la démarche originale reste essentiellement la même, en adaptant la définition des probabilités *avant-arrière* à chaque structure. Elle permet de conduire les calculs, pour l'estimation comme pour la segmentation, évitant ainsi les méthodes par simulations que l'on rencontre dans les modèles par champs de Markov. Ainsi, pour chacun d'eux une méthode originale d'estimation des paramètres, basée sur le principe de ECI, mais pour lesquels les méthodes EM et SEM restent possibles. La très grande généralité des modèles présentés ne les limite pas au traitement d'images, mais au contraire leur champ d'applications s'étend à tous les domaines classiques d'utilisation des modèles de Markov tels que traitement de la parole, génomique, économie, sciences de la santé.

Le tableau 2.1 synthétise les algorithmes déjà réalisés et met en exergue les combinaisons qu'il serait intéressant d'étudier dans un proche avenir. Nous pouvons notamment citer le modèle des *chaînes de Markov couple floues* qui doit permettre de segmenter des images floues et à bruit non indépendant. L'intérêt pour ce modèle se comprend pour les pixels à l'interface de deux classes. D'une part, le modèle couple permet de différentier les pixels à la frontière entre les classes (*p. ex.* $p\left(y_n, y_{n+1} \left| x_n = 0, x_{n+1} = 1\right.\right)$) et les pixels à l'intérieur d'une classe homogène (*p. ex.* $p\left(y_n, y_{n+1} \left| x_n = 0, x_{n+1} = 0\right.\right)$). D'autre part, le modèle flou est d'un grand intérêt aux frontières puisqu'il est susceptible de modéliser le dégradé qui peut apparaître à la jonction des deux classes ($p\left(y_n \left| x_n \in [0,1]\right.\right)$). Ainsi, la combinaison des modèles couple et flou permettra de donner une description fine de la transition des zones homogènes dans les images. Le modèle de *chaîne de Markov couple d'ordre supérieur*, où $\boldsymbol{Z}$ est une chaîne de Markov d'ordre supérieur, permettra quant à lui d'accroître la capacité de modélisation des zones texturées.

CHAPITRE 3

# Détection et Reconnaissance de Formes 2D

CE CHAPITRE décrit les travaux que j'ai réalisés dans le domaine de la description et de la représentation géométriques des formes 2D avec deux thèmes de recherche privilégiés : d'une part, la description invariante de formes pour des applications de reconnaissance ou d'indexation de bases d'images (paragraphes 3.1 et 3.2) ; d'autre part, la détection et le suivi d'objets d'intérêt dans les séquences vidéo, pour des applications de surveillance et de reconnaissance biométrique (paragraphes 3.3 et 3.4).

Historiquement, ces travaux ont débuté avec ma thèse [69] et je commence, dans le **premier paragraphe**, par résumer les résultats obtenus sur la Transformée de Fourier-Mellin (TFM) et son prolongement analytique (TFMA). Ce dernier a été introduit pour pallier à certains problèmes numériques qui avaient largement limité son utilisation ; depuis la publication de ces travaux [63, 55], la TFMA a par exemple été utilisée avec succès dans le cadre d'un algorithme de tatouage d'images permettant de répondre à certaines attaques de type géométrique [160]. Ce paragraphe est également l'occasion de décrire l'algorithme de détection des axes de symétrie par rotation et réflexion d'un objet à niveaux de gris que j'ai proposé en 2004 suite à mes travaux de thèse [56].

Ensuite, dans le **second paragraphe**, nous décrivons les travaux de thèse réalisés par Mme Rim Mezhoud-Khouaja [128] sous mon co-encadrement, qui utilisent les Moments Complexes (MC) pour réaliser l'indexation de bases d'objets à niveaux de gris. La famille de descripteurs invariants par similitudes planes, que nous avons proposé et sur laquelle repose l'algorithme d'indexation et de recherche par le contenu, est obtenue par un procédé de normalisation similaire à celui mis en œuvre pour la TFMA. Un travail récent portant sur le problème inverse de la reconstruction d'une image à partir de l'ensemble de ses MC est également présenté [88].

Le **troisième paragraphe** porte sur l'introduction d'un *a priori* géométrique dans l'algorithme original des *snakes* introduit par M. Kass [99] [36]. Celui-ci permet de détecter, dans une scène complexe, un objet dont on connaît un modèle de la forme, sans en connaître la pose exacte dans l'image. Nous avons ainsi proposé d'intégrer une force supplémentaire, construite à partir de la description du contour du *snake* et du modèle par une famille complète et stable de descripteurs invariants par similitudes planes (issue des séries de Fourier). Cette force rend l'algorithme des *snakes* plus robuste à l'initialisation, au bruit et permet de détecter des formes présentant de fortes concavités ; formes pour lesquelles l'algorithme original et ses extensions bien connues telles le GVF (*Gradient Vector Flow*) sont en échec. Nous avons finalement confronté cette méthode au suivi d'un objet en mouvement dans une séquence vidéo ; la forme de l'objet détecté à l'instant $t-1$ servant d'*a priori* pour détecter l'objet à l'instant $t$ [35, 54]. Nous faisons ainsi l'hypothèse que la représentation de l'objet dans le plan image se déforme essentiellement d'une similitude au cours de la vidéo.

La poursuite d'objets en mouvement est également le thème de l'étude présentée dans le **dernier paragraphe** de ce chapitre, et qui a fait l'objet des travaux de thèse de W. Ketchantang [104]. Cette fois-ci, l'objet à suivre est caractérisé par sa distribution de couleurs. Ainsi, ce n'est pas tant la forme de l'objet qui est poursuivie, mais une configuration de couleurs présentes dans la scène, indépendamment de l'organisation spatiale des couleurs entre-elles. Nous avons proposé de remplacer le modèle de mélange gaussien classiquement utilisé par un mélange de lois caractérisées par des copules gaussiennes (*cf.* section 2.2.3) et des marges issues du système de Pearson [102]. L'adaptation dynamique de ce mélange de couleurs au cours du temps permet une poursuite robuste, même si l'objet d'intérêt subit des déformations non rigides (avec occlusion partielle). Ce modèle a été adapté pour suivre une iris dans une séquence vidéo pour une application d'identification biométrique par de reconnaissance d'iris .

## 3.1 Transformée de Fourier-Mellin et reconnaissance de formes (thèse)

Mon travail de thèse s'inscrit dans les domaines de la description de scènes et de l'indexation de bases d'images par le contenu. Je me suis tout particulièrement intéressé au cas de la représentation globale des objets à niveaux de gris par la transformée de Fourier-Mellin.

La TFM est, selon la théorie de l'analyse harmonique sur les groupes [71], la transformée de Fourier sur le groupe des similitudes planes $\mathcal{G} = \mathbb{R}_+^* \times \mathbb{S}^1$, de mesure invariante $d\theta\ dr/r$. La TFM d'une fonction $f$ définie en coordonnées polaires, lorsqu'elle existe, s'écrit

$$\mathcal{M}_f(k,v) = \frac{1}{2\pi}\int_0^\infty\int_0^{2\pi} f(r,\theta)\, r^{-iv}\, e^{-ik\theta}\, d\theta\frac{dr}{r},$$

où l'espace des paramètres $(k,v)$ définit le groupe dual de $\mathcal{G}$ : $\widehat{\mathcal{G}} = \mathbb{Z}\times\mathbb{R}$. La TFM est définie pour toute fonction $f$ telle que

$$f \in \mathbf{L}^1(\mathcal{G}) \Rightarrow \|f\|_{\mathbf{L}^1(\mathcal{G})} = \int_0^\infty\int_0^{2\pi} |f(r,\theta)|\, d\theta\frac{dr}{r} < \infty. \tag{3.1}$$

### 3.1.1 Trois approximations numériques de la TFMA

On montre que si $f$ désigne la représentation d'un objet à niveaux de gris décrit en coordonnées polaires par rapport à son centre d'inertie, alors l'intégrale (3.1) diverge en général. Ce problème peut être résolu en calculant la TFM de la fonction $f_\sigma(r,\theta) = r^\sigma f(r,\theta)$ avec $\sigma > 0$. La transformée obtenue est appelée *prolongement analytique de la TFM* (TFMA) [86]

$$\forall (k,v)\in\widehat{\mathcal{G}},\quad \mathcal{M}_{f_\sigma}(k,v) = \frac{1}{2\pi}\int_0^\infty\int_0^{2\pi} f(r,\theta)\, r^{\sigma-iv}\, e^{-ik\theta}\, d\theta\frac{dr}{r}, \tag{3.2}$$

Comme $\mathcal{G}$ est commutatif, la transformée de Fourier sur $\hat{\mathcal{G}}$ existe, définissant ainsi la TFMA inverse de $f$ :

$$\forall (r,\theta)\in\mathcal{G},\quad f(r,\theta) = \int_{\mathbb{R}}\sum_{k\in\mathbb{Z}} \mathcal{M}_{f_\sigma}(k,v)\, r^{-\sigma+iv}\, e^{ik\theta} dv, \tag{3.3}$$

où $\delta k \otimes dv$ est la mesure invariante sur $\widehat{\mathcal{G}}$.

Les transformées directe et inverse sont définies pour des paramètres continus. Comme il n'existe pas de version discrète de la TFMA, nous avons proposé trois algorithmes d'approximation, basés sur des ré-échantillonnages polaire, log-polaire et cartésien de l'image. Ces approximations diffèrent : (i) par la manière dont les images sont ré-échantillonnées sur de nouvelles grilles et, (ii) par la méthode numérique d'intégration adaptée à ces grilles.

- **Approximation directe par échantillonnage polaire** L'*approximation directe* de la TFMA (TFMA-D) consiste à ré-échantillonner l'image $f(p,q)$ selon une grille polaire et à estimer directement les intégrales de Fourier-Mellin en les remplaçant par des sommes.

- **Approximation rapide par échantillonnage log-polaire** L'*approximation rapide* de la TFMA (TFMA-F) consiste à ré-échantillonner l'image $f(p,q)$ selon une grille log-polaire et à estimer les transformées de Fourier résultantes. En effet, par changement de la variable d'intégration de $r$ à $t = \ln(r)$, les transformées de Fourier-Mellin peuvent se réécrire sous la forme de deux transformées de Fourier

$$\begin{cases} \mathcal{M}_{f_\sigma}(k,v) = \dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}} \int_0^{2\pi} e^{t\sigma}\, f(e^t,\theta)\, e^{-i(k\theta+tv)} d\theta dt \\ \\ f(e^t,\theta) = e^{-t\sigma} \displaystyle\int_{\mathbb{R}} \sum_{k\in\mathbb{Z}} \mathcal{M}_{f_\sigma}(k,v)\, e^{i(k\theta+tv)} dv, \end{cases}$$

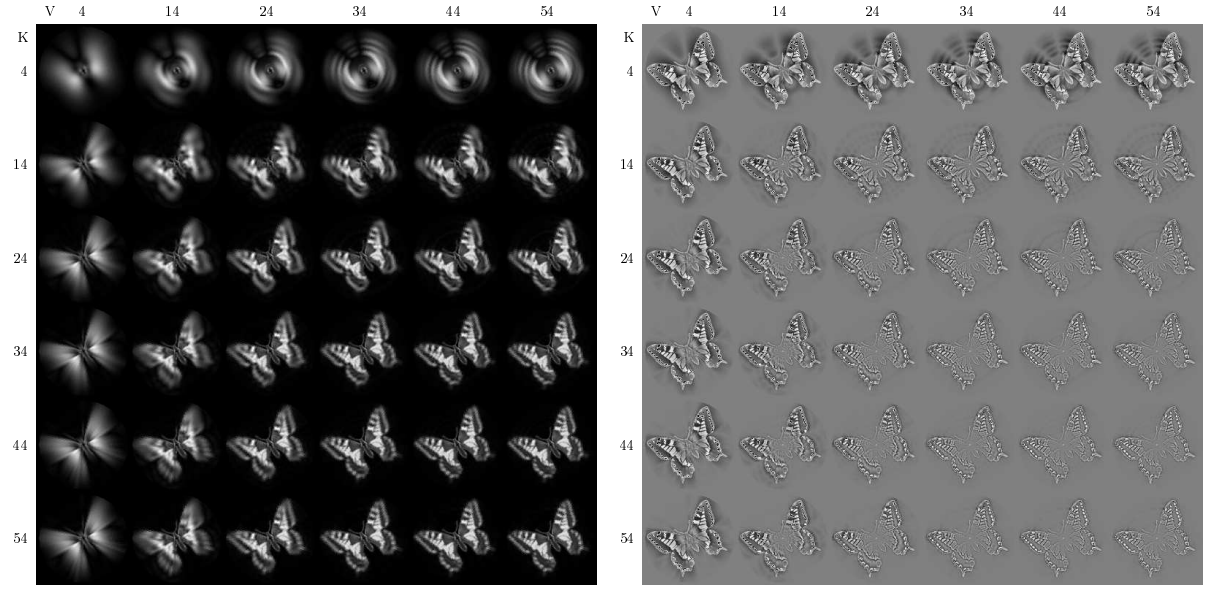
  de l'image déformée $e^{t\sigma}\, f(e^t,\theta)$.

- **Approximation cartésienne sans ré-échantillonnage préalable** La dernière méthode d'estimation est appelée *approximation cartésienne* de la TFMA (TFMA-C). Il est en effet possible d'effectuer un changement des variables d'intégration dans les équations (3.2) et (3.3) qui aboutit directement à des expressions en coordonnées cartésiennes :

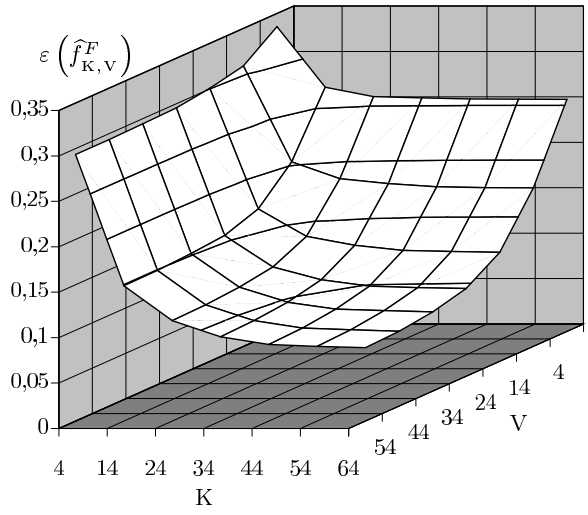
$$\begin{cases} \mathcal{M}_{f_\sigma}(k,v) = \dfrac{1}{2\pi} \displaystyle\iint_{\mathbb{R}^2} f(x,y)\, (x+iy)^{-k}\, (x^2+y^2)^{\frac{k-2+\sigma-iv}{2}}\, dxdy \\ \\ f(x,y) = \displaystyle\int_{\mathbb{R}} \sum_{k\in\mathbb{Z}} \mathcal{M}_{f_\sigma}(k,v)(x+iy)^k\, (x^2+y^2)^{\frac{-k-\sigma+iv}{2}}\, dv. \end{cases}$$

  Dans ce cas, aucun ré-échantillonnage n'est nécessaire et les TFMA sont estimées directement à partir de la grille rectangulaire.

Nous avons réalisé une comparison systématique entre les trois méthodes d'approximation portant sur la qualité de reconstruction (mesurée par l'erreur quadratique moyenne) en fonction du nombre d'harmoniques discrètes utilisées pour la reconstruction. Il est apparu que l'approximation TFMA-F fournissait le meilleur compromis des trois méthodes de part la qualité de reconstruction proposée et des temps de calculs sensiblement inférieurs aux deux autres méthodes. À titre d'illustration, nous avons reporté dans la figure 3.1 les résultats d'une expérience de reconstruction menée dans [55] sur une image de papillon.




(a) (b)




(c)

Figure 3.1 – Reconstruction de l'image d'un papillon en fonction du nombre d'harmoniques utilisées (algorithme TFMA-F). (a) Images reconstruites. (b) Différence avec l'image originale rehaussée de 128. (c) Erreur moyenne de reconstruction correspondante. Résultats extraits de [55].

### 3.1.2 Extraction de familles complètes d'invariants par similitudes

Par extension du théorème de Plancherel-Parseval au groupe des similitudes vectorielles planes, la TFMA est une isométrie de $\mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ dans $\mathbf{L}^2(\widehat{\mathcal{G}})$ [1]

$$\begin{aligned} \|f_\sigma\|_{\mathbf{L}^2(\mathcal{G})} &= \|\mathcal{M}_{f_\sigma}\|_{\mathbf{L}^2(\widehat{\mathcal{G}})}, \\ \int_0^\infty \int_0^{2\pi} r^{2\sigma} \, |f(r,\theta)|^2 \, d\theta \frac{dr}{r} &= \int_{\mathbb{R}} \sum_{k\in\mathbb{Z}} |\mathcal{M}_{f_\sigma}(k,v)|^2 \, dv. \end{aligned} \tag{3.4}$$

1. — Ces deux espaces seront appelés les espaces de représentation dans le domaine original et dans le domaine transformé.

On dit de deux objets $\mathcal{O}_1$ et $\mathcal{O}_2$ qu'ils ont même forme, si et seulement s'il existe une similitude $\mathcal{S}$ de paramètres $(\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{S}^1$ telle que $\mathcal{O}_2 = \mathcal{S} \circ \mathcal{O}_1$ soit $f^2(r, \theta) = f^1(\alpha r, \theta + \beta)$, où $f_\sigma^1$, $f_\sigma^2 \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ désignent deux représentations de $\mathcal{O}_1$ et $\mathcal{O}_2$. On montre aisément que l'action du groupe des similitudes sur l'espace de représentation de Fourier-Mellin se traduit par :

$$\forall (k, v) \in \widehat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma^2}(k, v) = \alpha^{-\sigma + iv} e^{ik\beta} \mathcal{M}_{f_\sigma^1}(k, v), \tag{3.5}$$

avec $\mathcal{M}_{f_\sigma^1}, \mathcal{M}_{f_\sigma^2} \in \mathbf{L}^2(\widehat{\mathcal{G}})$. L'équation (3.5) est appelée *théorème du retard de la TFMA*, par analogie au résultat classique pour la transformée de Fourier (et le groupe des translations). La TFMA convertit une similitude dans le domaine de représentation original en une multiplication complexe dans le domaine de Fourier-Mellin. Une forme est alors définie comme une classe d'équivalence sur $\mathbf{L}^2(\widehat{\mathcal{G}})$ et regroupe l'ensemble des objets identiques à une similitude près. Ainsi, le problème de *la reconnaissance de formes* consiste à déterminer si deux objets appartiennent ou non à la même classe.

Une famille complète et convergente de descripteurs invariants par similitude issue de la TFMA a été proposée [86]. Pour n'importe quel $\sigma > 0$ fixée, celle-ci peut s'écrire de la manière suivante :

$$\forall (k, v) \in \widehat{\mathcal{G}}, \quad \mathcal{I}_{f_\sigma}(k, v) = \mathcal{M}_{f_\sigma}(0, 0)^{\frac{-\sigma + iv}{\sigma}} \; e^{ik \arg(\mathcal{M}_{f_\sigma}(1,0))} \; \mathcal{M}_{f_\sigma}(k, v) \tag{3.6}$$

Chaque descripteur $\mathcal{I}_{f_\sigma}(k, v)$ est construit de manière à compenser l'action d'une similitude sur la TFMA (3.5). Cette compensation est réalisée grâce aux deux facteurs de normalisation :

$$\begin{cases} \Gamma(f_\sigma) = \mathcal{M}_{f_\sigma}(0, 0)^{\frac{1}{\sigma}}, \\ \\ \Theta(f_\sigma) = \arg(\mathcal{M}_{f_\sigma}(1, 0)). \end{cases}$$

qui agissent respectivement sur l'amplitude et la phase de la TFMA. Comme cette famille est complète et convergente ($\|\mathcal{I}_{f_\sigma}\|_{\mathbf{L}^2(\widehat{\mathcal{G}})} < \infty$), il est possible de définir une distance euclidienne sur l'espace des formes :

$$d_2(\mathcal{I}_{f_\sigma}, \mathcal{I}_{g_\sigma}) = \left( \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |\mathcal{I}_{f_\sigma}(k, v) - \mathcal{I}_{g_\sigma}(k, v)|^2 dv \right)^{\frac{1}{2}}.$$

Cette distance est nulle si et seulement si les objets ont même forme. Lorsque les objets sont différents, la distance est utilisée pour mesurer la ressemblance entre les objets. Ainsi, la recherche d'un objet inconnu $g$ parmi un ensemble d'objets de référence est réalisée : (i) en calculant la distance entre $g$ et tous les modèles ; (ii) en sélectionnant le ou les modèles qui présentent la plus faible distance avec $g$.

D'un point de vue numérique, la stabilité de la famille (3.6) dépend en grande partie de la qualité de l'estimation des deux facteurs de normalisation. Pour réduire cette sensibilité numérique, nous avons proposé d'utiliser des facteurs de normalisation construits par combinaison de plusieurs harmoniques de Fourier-Mellin, comme

$$\begin{cases} \Gamma(f_\sigma) = \left( \prod_{v_1=1}^{V} |\mathcal{M}_{f_\sigma}(0, v_1)| \right)^{\frac{1}{\sigma V}} \qquad V \in \mathbb{N}, \\ \\ \Theta(f_\sigma) = -\dfrac{1}{K} \displaystyle\sum_{k_1=1}^{K} \arg(\mathcal{M}_{f_\sigma}(k_1 + 1, 0)\, \mathcal{M}_{f_\sigma}(-k_1, 0)) \quad K \in \mathbb{N}^*. \end{cases} \tag{3.7}$$

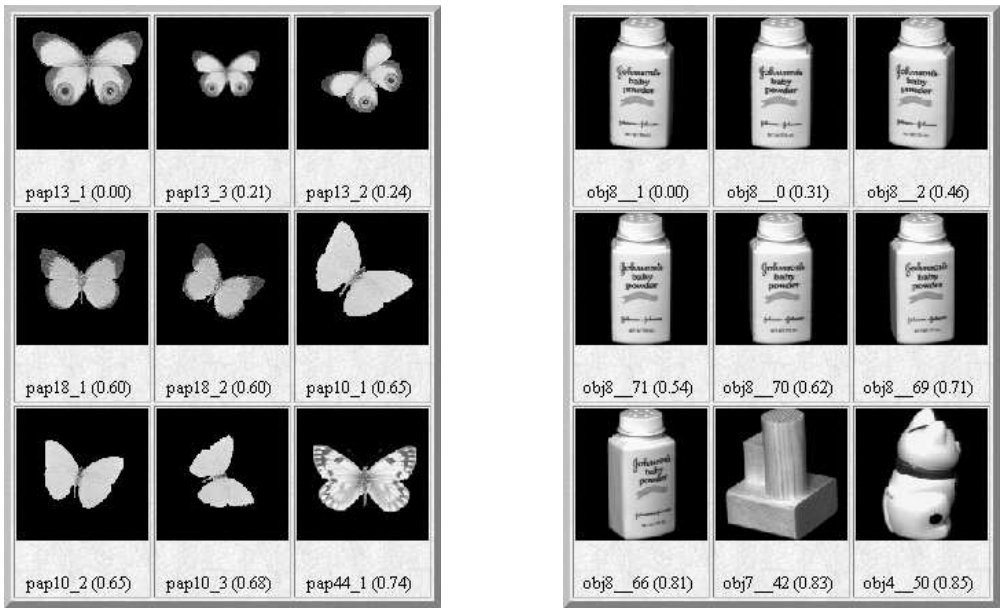



Figure 3.2 – Résultats de reconnaissance pour deux images requêtes (situées en haut et à gauche) à l'aide de la famille définie par les paramètres de normalisation de l'équation (3.7) et de l'approximation TFMA-F. Les images sont triées en fonction de la distance croissante à la requête, de gauche à droite et de bas en haut. Résultats extraits de [55].

À titre d'illustration, nous avons reporté dans la figure 3.2 les résultats d'une expérience de reconnaissance de formes portant sur une base d'une centaine d'images de papillons et la base d'objets de 1440 images de l'université de Columbia [134]. Ces résultats illustrent le bon comportement de cette famille d'invariants aux occlusions partielles et aux faibles déformations non assimilables aux similitudes planes.

### 3.1.3 Estimation des paramètres et détection des symétries

Nous avons également proposé un algorithme efficace dédié à l'estimation de la différence de taille et d'orientation entre deux objets de même forme à partir de leur représentation par la TFMA [56]. La méthode est basée sur le calcul d'une distance euclidienne, en utilisant le théorème de Plancherel-Parseval (3.4) et le théorème du retard (3.5).

La norme euclidienne $\mathbf{L}^2(\mathcal{G})$ peut être utilisée pour comparer deux objets $f$ et $g$ en calculant $\|f_\sigma - g_\sigma\|_{\mathbf{L}^2(\mathcal{G})}$. Elle vaut zéro si et seulement si les objets sont strictement identiques. Pour prendre en compte d'éventuelles différences de taille et d'orientation, nous avons formulé le problème de la manière suivante. Pour tout $(\rho, \psi)$ de $\mathbb{R}_+^* \times \mathbb{S}^1$, nous définissons la fonctionnelle $\mathbf{E}$

$$\mathbf{E}_{f_\sigma, g_\sigma}(\rho, \psi) = \|f_\sigma(r, \theta) - g_\sigma(\rho r, \theta + \psi)\|_{\mathbf{L}^2(\mathcal{G})} .$$

En utilisant l'égalité de Parseval et le théorème du retard pour $g_\sigma$, $\mathbf{E}$ peut être ré-écrite avec la

norme appropriée dans le domaine de représentation de Fourier-Mellin :

$$\begin{aligned}\forall(\rho,\psi)\in\mathbb{R}_+^*\times\mathbb{S}^1,\\ \mathbf{E}_{f_\sigma,g_\sigma}(\rho,\psi)&=\left\|\mathcal{M}_{f_\sigma(r,\theta)}-\mathcal{M}_{g_\sigma(\rho r,\theta+\psi)}\right\|_{\mathbf{L}^2(\widehat{\mathcal{G}})},\\ &=\left(\int_{\mathbb{R}}\sum_{k\in\mathbb{Z}}\left|\mathcal{M}_{f_\sigma}(k,v)-\rho^{-\sigma+iv}\,e^{ik\psi}\,\mathcal{M}_{g_\sigma}(k,v)\right|^2\,dv\right)^{\frac{1}{2}}.\end{aligned}\tag{3.8}$$

Lorsque les paramètres $\rho$ et $\psi$ parcourent le groupe des similitudes, nous obtenons une mesure de correspondance pour chaque lieu de l'espace des paramètres. Si $\mathbf{E}=0$ pour un certain couple $(\alpha,\beta)$, alors les deux objets s'apparient totalement (en conséquence les objets sont considérés comme similaires) et les paramètres $(\alpha,\beta)$ correspondent à la similitude qu'il faut appliquer à l'un des objets pour obtenir le second.

À cause des approximations numériques et de l'anisotropie de la grille cartésienne aux similitudes, la valeur minimale n'est jamais exactement zéro pour des objets similaires et l'estimation des paramètres consiste à recherche le couple $(\hat{\rho}_{\mathbf{E}},\hat{\psi}_{\mathbf{E}})$ qui minimise la fonctionnelle (3.8) :

$$\left(\hat{\rho}_{\mathbf{E}},\hat{\psi}_{\mathbf{E}}\right)=\arg\underbrace{\left(\min_{\rho\in\mathbb{R}_+^*,\psi\in\mathbb{S}^1}\widehat{\mathbf{E}}_{f_\sigma,g_\sigma}(\rho,\psi)\right)}_{\widehat{\mathbf{E}}^{\min}_{f_\sigma,g_\sigma}}.$$

Ce minimum peut être obtenu aisément en utilisant un algorithme d'optimisation classique tel l'algorithme de Powell [152]. L'algorithme de détection ainsi construit est illustré par l'exemple de la figure 3.3. Lorsque le groupe de transformations considéré se limite aux rotations, on montre que la fonctionnelle est exactement la distance de Hausdorff entre formes exprimée dans le domaine de représentation de Fourier-Mellin.

**Détection des symétries par rotation et réflexions–** Cette dernière caractéristique nous a permis de construire une méthode de détection des axes de symétrie par rotation d'une image. Pour cela, nous avons proposé de minimiser la fonctionnelle à un seul paramètre suivante

$$\forall\psi\in\mathbb{S}^1,\quad\mathbf{E}_{f_\sigma}(\psi)=\left(2\sum_{k\in\mathbb{Z}}\left(1-\cos\left(k\psi\right)\right)\,\int_{\mathbb{R}}\left|\mathcal{M}_{f_\sigma}(k,v)\right|^2\,dv\right)^{\frac{1}{2}}.$$

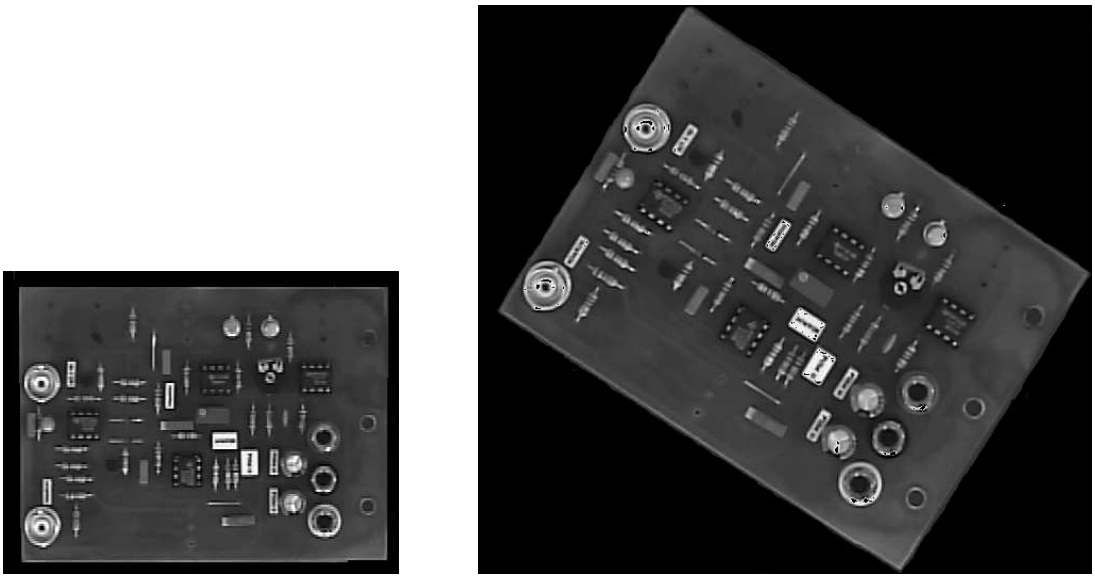
La fonction $\mathbf{E}_{f_\sigma,f_\sigma}$ est paire et périodique de période $\frac{2\pi}{\mathrm{P}}$ si l'image possède $P$ axes de symétrie par rotation. L'orientation des axes est donnée par les angles $\psi_i$ tels que $\mathbf{E}_{f_\sigma}(\psi_i)=0$. Notons que $\mathbf{E}_{f_\sigma}(0)=0$ est toujours vraie puisque qu'une image est toujours symétrique par rapport à un angle de $0°$.

Pour la détection des axes de symétrie par réflexion, nous avons proposé de minimiser la fonctionnelle suivante

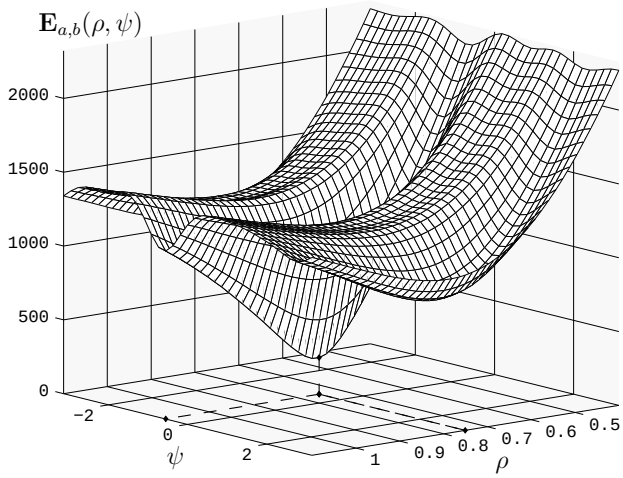
$$\forall\psi\in[0;\pi],\quad\mathbf{F}_{f_\sigma}(\phi)=\left(\int_{\mathbb{R}}\sum_{k\in\mathbb{Z}}\left|\mathcal{M}_{f_\sigma}(k,v)-e^{-2ik\phi}\,\mathcal{M}_{f_\sigma}(-k,v)\right|^2\,dv\right)^{\frac{1}{2}}.$$

$\mathbf{F}$ vaut zéro pour chaque angle $\phi$ qui définit un axe de symétrie par réflexion[2]. Il est facile de vérifier que $\mathbf{F}_{f_\sigma}$ est également paire et périodique. Ces deux algorithmes sont illustrés au travers de la figure 3.4 et permettent de détecter tous les axes de symétrie par rotation et réflexion d'une image réelle représentant une étoile de mer.

2. — Les angles $\phi+\pi$ caractérisent les mêmes axes que les angles $\phi$.



(a) (b)

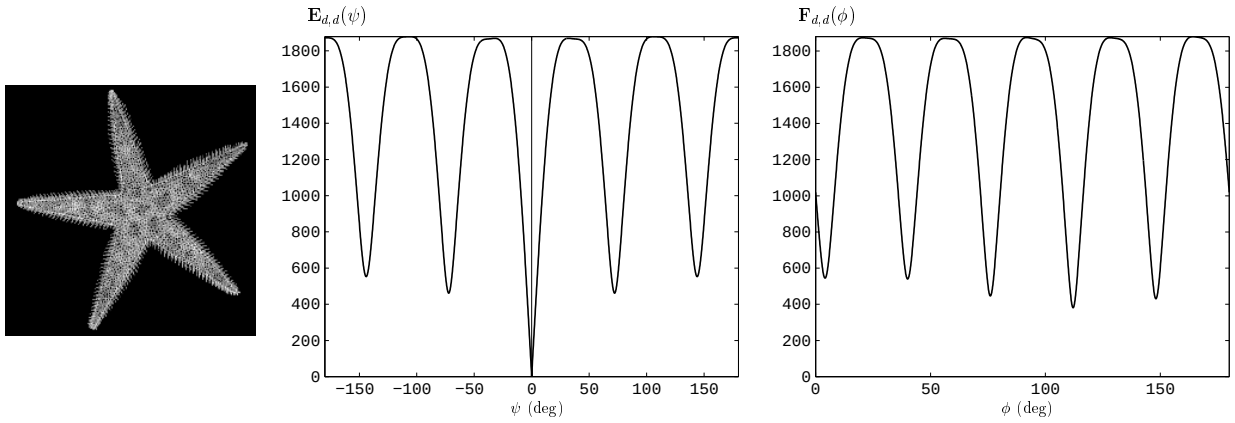



(c)

Figure 3.3 – (a) et (b) Deux images d'une même carte électronique. (c) La courbe 3D représente $\mathbf{E}$. Celle-ci montre un minimum pour les paramètres donnés par la différence de taille et d'orientation entre (a) et (b). Résultats extraits de [56].

## 3.2 Moments complexes et indexation de bases d'images (thèse de R. Mezhoud)

Les Moments Géométriques (MG) et les Moments Complexes (MC) sont des outils parmi les plus utilisés pour la construction de descripteurs de formes dans des applications de reconnaissance et de détection. Le travaux initiaux de H. Ming-Kuel [129] pour les MG et de P.J. Davis [45] pour les MC ont été étendus à une grande variété de nouveaux moments, particulièrement ceux construits sur la base de polynômes orthogonaux, comme les moments de Zernike, de Tchebichev et de Krawtchouk [175]. La propriété d'orthogonalité assure que l'information codée dans les mo-




(a) Image (b) $\mathbf{E}_{f_\sigma}(\psi)$ (c) $\mathbf{F}_{f_\sigma}(\phi)$

Figure 3.4 – Une image d'étoile de mer (a) et dessin des fonctionnelles pour l'estimation des axes de symétrie par rotation (b) et réflexion (c) de l'image. Résultats extraits de [88].

ments n'est pas redondante, ce qui est généralement considéré comme un avantage (compacité de la représentation) par rapport aux moments non orthogonaux dont font partie les MG et les MC. Cependant, il faut souligner que cette redondance d'information peut devenir salutaire dans le cas de données bruitées.

Ainsi, dans le cadre sa thèse [128], Rim Mezhoud-Khouaja s'est intéressée plus particulièrement aux MC et à leur évaluation pour l'indexation de bases d'objets à niveaux de gris. Elle a proposé une nouvelle famille complète de descripteurs invariants par similitudes, obtenue grâce à un procédé de normalisation similaire à celui présenté dans le cadre de la TFMA. La complétude des descripteurs invariants garantissant l'inversibilité de la représentation par les MC, nous avons proposé ultérieurement à cette thèse un algorithme de reconstruction d'une image à partir d'un ensemble fini de ses MC.

### 3.2.1 Moments géométriques et invariants de Hu

Le MG $m_f(p,q)$, d'ordre $(p+q)$, d'une fonction $f$ continue par morceaux est défini par

$$m_f(p,q) = \iint_{\mathbb{R}^2} x^p \, y^q \, f(x,y) \, dx \, dy,$$

où $p$ et $q$ désignent deux entiers naturels.

L'ensemble des moments $m_f(p,q)$ est unique si l'intégrale de $f(x,y)$ sur $\mathbb{R}^2$ existe. Réciproquement, l'image $f(x,y)$ est uniquement déterminée par l'ensemble de ses moments $m_f(p,q)$. Le moment d'ordre 0 représente l'énergie totale de l'image et les moments d'ordre 1 permettent de calculer le centre de masse d'une image selon

$$(\bar{x}, \bar{y}) = \left( \frac{m_f(1,0)}{m_f(0,0)}, \frac{m_f(0,1)}{m_f(0,0)} \right).$$

On définit ainsi les MG centrés de l'image $\mu_f(p,q)$, invariants par translation

$$\forall (p,q) \in \mathbb{N}^2, \quad \mu_f(p,q) = \iint_{\mathbb{R}^2} (x-\bar{x})^p \; (y-\bar{y})^q \; f(x,y) \, dx \, dy.$$

Si on considère maintenant deux images $g$ et $f$ représentant des objets identiques à un facteur d'échelle $\alpha$ près, c'est-à-dire $g(x,y) = f(\frac{x}{\alpha}, \frac{y}{\alpha})$, alors la relation entre les moments centrés d'ordre $(p+q)$ des deux images devient $\mu_g(p,q) = \alpha^\gamma \, \mu_f(p,q)$ avec $\gamma = p+q+2$. Ce résultat est à l'origine des MG réduits $\eta_f(p,q) = \mu_f(0,0)^{-\frac{\gamma}{2}} \, \mu_f(p,q)$ qui sont invariants par translation et homothétie du plan.

H. Ming-Kuel [129] a proposé un ensemble de sept descripteurs invariants, obtenus par des combinaisons non linéaires des moments réduits jusqu'à l'ordre trois [3] :

$$
\begin{aligned}
\phi_f^1 =& \eta_{2,0} + \eta_{0,2} \\
\phi_f^2 =& \left(\eta_{2,0} - \eta_{0,2}\right)^2 + 4\eta_{1,1}^2 \\
\phi_f^3 =& \left(\eta_{3,0} - 3\eta_{1,2}\right)^2 + \left(3\eta_{2,1} - \eta_{0,3}\right)^2 \\
\phi_f^4 =& \left(\eta_{3,0} + \eta_{1,2}\right)^2 + \left(\eta_{2,1} + \eta_{0,3}\right)^2 \\
\phi_f^5 =& \left(\eta_{3,0} - 3\eta_{1,2}\right)\left(\eta_{3,0} + \eta_{1,2}\right)\left[\left(\eta_{3,0} + \eta_{1,2}\right)^2 - 3\left(\eta_{2,1} + \eta_{0,3}\right)^2\right] \\
& + \left(3\eta_{2,1} - \eta_{0,3}\right)\left(\eta_{2,1} + \eta_{0,3}\right)\left[3\left(\eta_{3,0} + \eta_{1,2}\right)^2 - \left(\eta_{2,1} + \eta_{0,3}\right)^2\right] \\
\phi_f^6 =& \left(\eta_{2,0} - \eta_{0,2}\right)\left[\left(\eta_{3,0} + \eta_{1,2}\right)^2 - \left(\eta_{2,1} + \eta_{0,3}\right)^2\right] \\
& + 4\eta_{1,1}\left(\eta_{3,0} + \eta_{1,2}\right)\left(\eta_{0,3} + \eta_{2,1}\right) \\
\phi_f^7 =& \left(3\eta_{2,1} - \eta_{0,3}\right)\left(\eta_{3,0} + \eta_{1,2}\right)\left[\left(\eta_{3,0} + \eta_{1,2}\right)^2 - 3\left(\eta_{2,1} + \eta_{0,3}\right)^2\right] \\
& - \left(\eta_{3,0} - 3\eta_{2,1}\right)\left(\eta_{1,2} + \eta_{0,3}\right)\left[3\left(\eta_{3,0} + \eta_{1,2}\right)^2 - \left(\eta_{2,1} + \eta_{0,3}\right)^2\right]
\end{aligned}
$$

La famille ainsi construite n'est évidemment pas complète puisque ces coefficients ne suffisent pas à reconstruire l'image originale. Il n'existe pas de moyen systématique pour définir des descripteurs invariants par rotation, du fait de l'inadéquation des MG pour ce type de transformations. Les moments complexes, de part leur construction, permettent de résoudre ce problème.

### 3.2.2 Invariants complets des moments complexes

Pour une fonction intégrable $f(x,y)$, les moments complexes sont définis par

$$c_f(p,q) = \iint_{\mathbb{R}^2} (x+iy)^p \, (x-iy)^q \, f(x,y) \, dxdy,$$

où $p, q \in \mathbb{N}$. En coordonnées polaires, ils se ré-écrivent

$$c_f(p,q) = \int_0^\infty \int_0^{2\pi} r^{p+q+1} e^{i(p-q)\theta} \, f(r,\theta) \, drd\theta. \tag{3.9}$$

Il est aisé de vérifier que la relation entre les MC de deux images $f$ et $g$ de même forme mais de taille et d'orientation différentes, c'est-à-dire $g(r,\theta) = f(\alpha r, \theta + \beta)$, est donnée par

$$\forall p, q \in \mathbb{N}, \quad c_g(p,q) = \alpha^{-(p+q+2)} \, e^{-i(p-q)\beta} \, c_f(p,q).$$

Cette relation est appelée « Théorème du retard des MC » puisqu'elle joue le même rôle que le théorème du retard original issu de la transformée de Fourier.

3. — La notation $\eta_f(p,q)$ est remplacée par $\eta_{pq}$ pour des commodités d'écriture.

À partir de l'équation (3.9), J. Flusser et T. Suk [79] ont proposé une famille complète d'invariants par rotation ($\Phi_f$) que l'on peut écrire de la manière suivante

$$\forall p, q \in \mathbb{N}, \quad \Phi_f(p,q) = c_f(p_0 - 1, p_0)^{p-q}\, c_f(p,q),$$

où $p_0 > 0$ est un indice arbitraire et le moment $c_f(p_0 - 1, p_0) \neq 0$ joue le rôle d'un facteur de normalisation qui rend les MC invariants aux rotations. De manière indépendante et antérieurement à ces travaux, nous avons proposé une famille complète d'invariants simultanément aux rotations et aux homothéties ($I_f$) [63], dont l'expression est

$$\forall p, q \in \mathbb{N}, \quad I_f(p,q) = \Gamma_f^{-(p+q+2)}\, e^{-i(p-q)\,\Theta_f}\, c_f(p,q), \tag{3.10}$$

avec $\Theta_f = \arg\left(c_f(1,0)\right)$ et $\Gamma_f = \sqrt{c_f(0,0)}$. Les coefficients $\Gamma_f$ et $\Theta_f$ normalisent l'amplitude et la phase des MC de manière indépendante. Il est facile de vérifier que d'autres facteurs de normalisation peuvent être proposés dès lors que, pour $g(r,\theta) = f(\alpha r, \theta + \beta)$, les conditions suivantes sont remplies

$$\Theta_f - \Theta_g = \beta\ (2\pi) \quad \text{et} \quad \frac{\Gamma_f}{\Gamma_g} = \alpha. \tag{3.11}$$

Par exemple, si nous introduisons

$$\Theta_f = -\arg\left(c_f(p_0 - 1, p_0)\right) \quad \text{et} \quad \Gamma_f = \sqrt[2(p_1+1)]{|c_f(p_1,p_1)|},$$

dans l'équation (3.10), avec $p_0 > 0$ et $p_1 \geq 0$, nous obtenons une nouvelle famille d'invariants. Ainsi, le travail de J. Flusser apparaît comme un cas particulier pour lequel l'invariance unique aux rotations est obtenue, puisque la normalisation en amplitude ne rend pas la famille invariante aux facteurs d'échelle.

Il est facile de vérifier que les ensembles de la forme donnée par l'équation (3.10) sont complets puisque nous pouvons écrire

$$\forall p, q \in \mathbb{N}, \quad c_f(p,q) = \Gamma_f^{p+q+2}\, e^{i(p-q)\,\Theta_f}\, I_f(p,q).$$

Ainsi, la connaissance de tous les invariants $I_f(p,q)$ et des deux facteurs de normalisation $\Gamma_f$ et $\Theta_f$ permet de reconstruire tous les MC $c_f(p,q)$ et, par voie de conséquence, l'image originale puisque les MC sont complets et inversibles. Toutes les familles de la forme (3.10) avec les conditions imposées par l'équation (3.11) sont théoriquement équivalentes. Le choix des facteurs de normalisation doit être dicté par des considérations numériques telles que la robustesse au bruit. Généralement, des facteurs utilisant des moments d'ordre peu élevé sont préférés à cause de leur relative insensibilité aux approximations numériques.

Pour quantifier la ressemblance entre deux objets $f$ et $g$, nous pouvons classiquement utiliser la norme de la différence des descripteurs de chacune des deux images selon

$$d(f,g) = \left( \sum_{p,q \in \mathbb{N}} |I_f(p,q) - I_g(p,q)|^2 \right)^{\frac{1}{2}}.$$

Cette quantité sera nulle si les objets ne se distinguent que d'une similitude.

### 3.2.3 Reconstruction à partir des moments complexes

Nous venons de montrer qu'il était possible de reconstruire les moments complexes d'une images à partir de ses invariants, étant donné la propriété de complétude. Nous nous sommes alors intéressés à la seconde étape qui consiste à reconstruire l'image originale à niveaux de gris à partir d'un ensemble fini de ses MC. Ce problème est réputé difficile et les solutions telles que celle proposée par M. R. Teague [169] s'avèrent insatisfaisantes en qualité et/ou en temps de calcul [153]. Nous avons proposé récemment [88] une autre approche qui exploite le lien entre la Transformée de Fourier Discrète (TFD) d'une image et ses MC.

La TFD $F(u,v)$ d'une image peut s'exprimer en fonction de ses MC selon (*cf.* démonstration dans [88]) :

$$F(u,v) = \sum_{p=0}^{\infty} \frac{(-i\pi)^p}{p!} \sum_{k=0}^{p} \binom{p}{k} \left(\frac{u+iv}{N}\right)^{p-k} \left(\frac{u-iv}{M}\right)^{k} c_f\left(k, p-k\right).$$

Pour être reconstruite, chaque harmonique a besoin de l'ensemble des moments géométriques, ce qui n'est pas réaliste d'un point de vue numérique et informatique. Cependant, nous avons montré de manière expérimentale qu'un nombre limité de moments contribue significativement au calcul de l'harmonique et que l'on peut se satisfaire d'un nombre restreint de MC pour obtenir la TFD. L'image originale peut alors être reconstruite en calculant la TFD (tronquée) inverse.

Nous avons réalisé un certain nombre d'expériences numériques pour mesurer la qualité des images reconstruites, en fonction du nombre de ses moments, en les comparant à l'originale par le biais du calcul d'une erreur de reconstruction moyenne. Il est intéressant de constater qu'au fur et à mesure que le nombre de moments grandit, l'image reconstruite s'approche de l'image originale, comme l'illustre le résultat de reconstruction présenté dans la figure Fig. 3.5. Il faut cependant noter que cette méthode ne nous a pas permis à l'heure actuelle de reconstruire une image à niveaux de gris avec une qualité suffisante pour envisager de l'utiliser en compression d'images par exemple.

### 3.2.4 Indexation de bases d'objets à niveaux de gris

L'essor des dispositifs d'acquisition (caméras numériques, scanner, IRM), des capacités de stockage (mémoire de masse), des systèmes de transmission et de diffusion (réseau Internet, réseaux spécialisés à hauts-débits) a favorisé la création et le développement de bases d'images volumineuses et de natures très diverses. Ceci pose le problème de l'archivage, de l'exploitation et de la consultation de ces bases d'images. Les techniques traditionnelles d'indexation manuelle de ces bases (essentiellement sous forme de mots-clés) sont confrontées à deux problèmes majeurs. Le premier concerne la quantité de travail nécessaire au processus d'annotation textuel. Le second est lié à la richesse du contenu sémantique des images et à la subjectivité de la description.

Pour tenter de s'affranchir de ces problèmes, l'accès aux bases par le contenu visuel des images est devenu, depuis une vingtaine d'années, un sujet de recherche particulièrement dynamique. Au lieu d'une annotation manuelle par des attributs textuels, les images sont indexées automatiquement par leur propre contenu, grâce à des caractéristiques traduisant la couleur, la forme et la texture des images. De part les enjeux scientifiques mais aussi commerciaux, la recherche dans ce domaine reste particulièrement active.

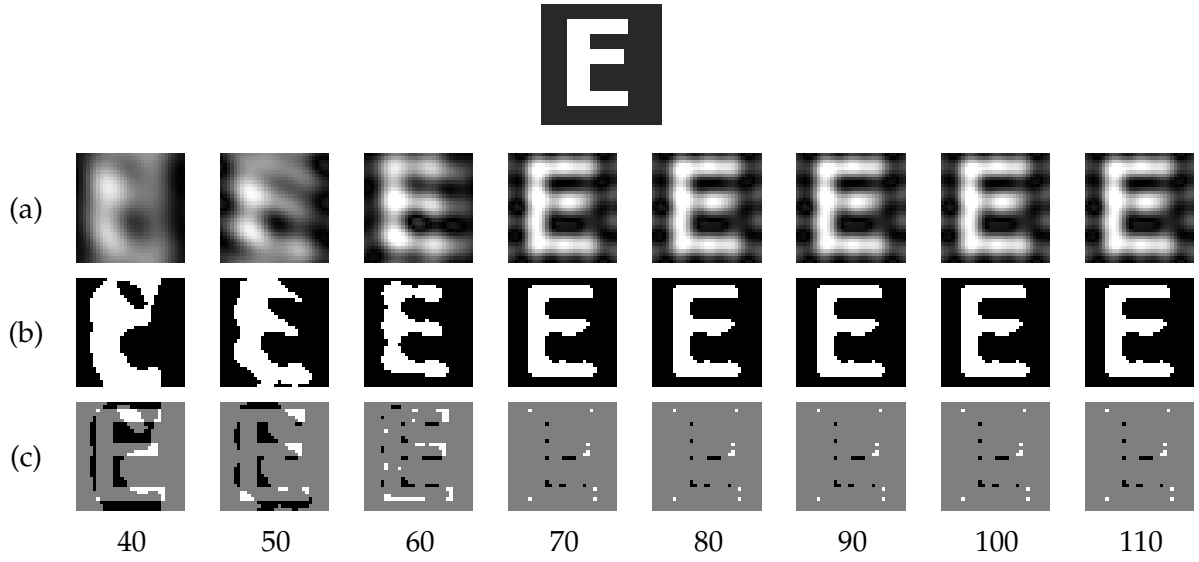



Figure 3.5 – La ligne (a) montre l'évolution de la reconstruction de la lettre 'E' avec un nombre croissant de MC. La ligne (b) montre la version binarisée de ces images reconstruites (seuil à $0.5$). La ligne (c) montre la différence pixel à pixel entre l'image originale et l'image reconstruite. Résultats extraits de [88].

**Principe :** L'architecture typique d'un système d'indexation est schématisée par la figure 3.6. Dans ce modèle, on distingue deux étapes : la phase *off-line* pour les pré-traitements, l'organisation et la gestion de la base d'images de référence (sous-système d'indexation de la base), et la phase de consultation *on-line* pour les traitements propres à chaque nouvelle requête de l'utilisateur (sous-système de recherche). Les deux sous-systèmes ont en commun les deux traitements suivants :

- ✓ L'extraction de descripteurs caractéristiques des images, que ce soit une image de la base durant la phase *off-line* ou une image requête proposée par l'utilisateur lors de la phase *on-line*.
- ✓ La construction des index à partir des descripteurs. Le rôle principal des index est d'accélérer la recherche des images similaires à la requête durant la phase de consultation.

Pendant la consultation de la base, l'utilisateur sélectionne une image par le biais d'une interface graphique. Les index de la requête sont confrontés aux index des images de références. Finalement, le système sélectionne et présente à l'utilisateur les « images les plus similaires » à la requête.

**Illustration :** Le système d'indexation que nous avons proposé à partir de la représentation des images par la famille complète d'invariants issue des moments complexes est illustré au travers de l'expérience rapportée dans la figure 3.7. La base utilisée est constituée de 1440 images représentant 20 objets 3D pris dans 72 positions différentes (par rotation de 5° autour d'un axe vertical) [4]. La principale difficulté de cette base réside dans le fait que le mouvement apparent des objets fausse sensiblement le modèle de transformation géométrique que nous considérons ici. Cette base permet de tester la stabilité et la robustesse de nos descripteurs invariants aux déformations de formes.

Nous avons reporté les résultats obtenus à l'aide des invariants complets issus des moments

4. — Base d'objets de l'université de Columbia.

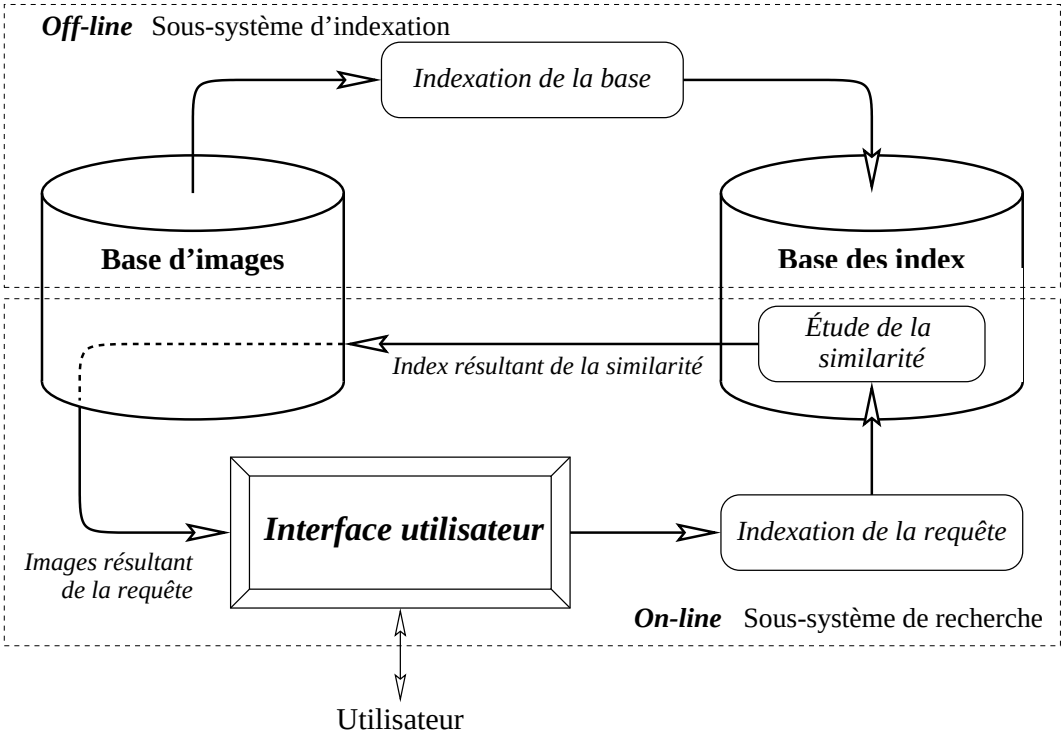



Figure 3.6 – Architecture générique d'un système d'indexation de bases d'images.

complexes et de la transformée de Fourier-Mellin Analytique, et les invariants de Hu. Chaque résultat présente, de gauche à droite et de haut en bas, les 16 premières images classées dans l'ordre de similarité avec l'image requête (que l'on retrouve dans le coin supérieur gauche). L'image située en première position est la requête elle-même.

**Amélioration de la recherche :** L'extraction de primitives pertinentes est un problème qui ne connaît pas de solution dans le cas général. D'une part, parce qu'un unique type de descripteurs ne permet pas de caractériser aussi bien une base d'empreintes digitales qu'une base de visages ou une base de tableaux de maîtres par exemple. Le type de descripteur à inclure dans la représentation des images dépend de la nature de la base et de l'application qui en est faite. D'autre part, il semble qu'il n'existe pas d'attributs qui puissent modéliser une base selon tous les points de vues, pour des raisons multiples qui peuvent être illustrées par la subjectivité de la requête de l'utilisateur.

Les moteurs de recherche tentent de combiner des primitives variées pour améliorer la discrimination et la classification. Par conséquent, le nombre de descripteurs de formes par image est de plus en plus élevé. Un des problèmes est alors de déterminer le poids approprié pour chaque type de descripteur. La plupart des moteurs font donc également intervenir l'utilisateur par le moyen d'une interface interactive qui l'associe à l'interprétation des résultats en permettant, par exemple, la sélection de plusieurs requêtes simultanées. Le système recherche alors les modèles les plus semblables à toutes les requêtes présentées. Pour améliorer les performances de recherche et pour s'adapter à la subjectivité de la requête de l'utilisateur, les efforts portent sur la conception de systèmes d'apprentissage dynamique qui modifient les poids des descripteurs au fur et à

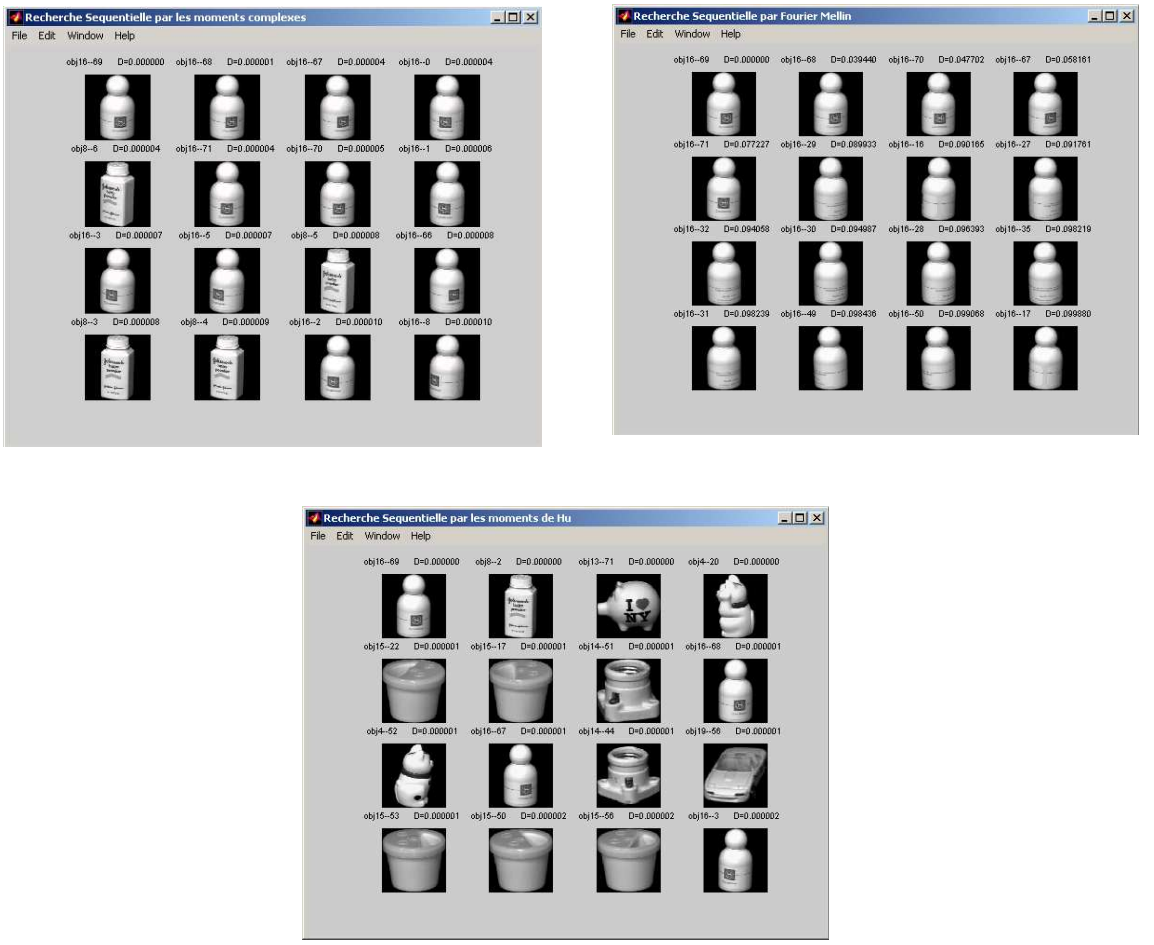
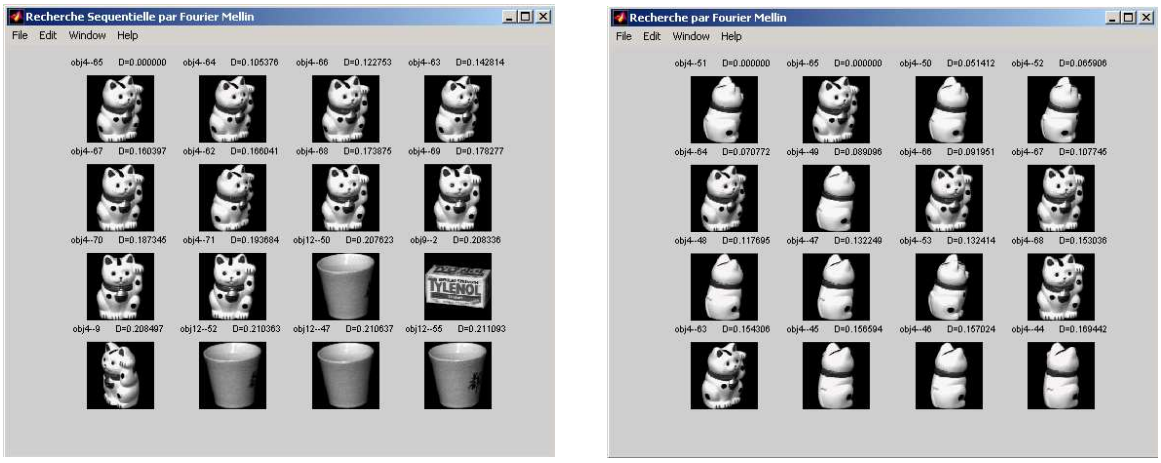



Figure 3.7 – Résultats d'une recherche séquentielle sur l'image requête Obj16_69 en utilisant (a) les invariants complets des MC, (b) les invariants complets de la TFMA et (c) les invariants de Hu. Résultats extraits de [128].

mesure des requêtes lancées par l'utilisateur [5].

Une autre solution pour améliorer les résultats de recherche consiste à ne plus considérer une unique image requête mais un ensemble d'images requêtes. Cela permet d'aider l'utilisateur à préciser ce qu'il recherche. Pour illustrer ce concept, nous présentons dans la figure 3.8, les résultats de recherche à partir d'une unique requête (a) et ceux obtenus avec deux images requêtes (b). Dans ce dernier cas, pour mesurer la similarité nous avons simplement trié les images selon la somme des erreurs obtenues individuellement sur chaque image requête. Le résultat de recherche est visuellement bien plus satisfaisant.

5. — Exemples : moteurs de recherche FOUREYES et PHOTOBOOK [130, 141] (`http://www.media.mit.edu/~tpminka/photobook/foureyes/`), MARS [161] (`http://jadzia.ifp.uiuc.edu:8000`) et PICHUNTER [41].




(a) Recherche avec une image requête (b) Recherche avec deux images requêtes

Figure 3.8 – Résultats d'une recherche pour les invariants complets. Résultats extraits de [128].

## 3.3 *A priori* géométrique pour les *snakes* et détection d'objet (master recherche de M.-A. Chermi)

Plutôt que d'utiliser l'information à niveaux de gris des objets comme dans les deux paragraphes précédents, nous avons aussi considéré le cas d'une description par le contour externe. Celui-ci peut s'avérer satisfaisant lorsque les formes étudiées sont suffisamment simples ou lorsque l'application envisagée le permet. Dans le cadre de la détection et la localisation d'objets, les méthodes de contours actifs appelées *snakes*, introduits en 1987 par M. Kass [99], ont connu un intérêt marqué à l'origine de nombreuses applications comme la stéréovision [32, 83, 105], l'imagerie médicale [38, 39] et la télédétection [157]. Un *snake* est une courbe fermée qui, à partir d'une position initiale, évolue vers un objet d'intérêt sous l'influence de forces internes –imposant des contraintes sur l'élasticité et la rigidité du contour– et de forces externes –par exemple le gradient de l'image.

L'engouement pour ce modèle se justifie principalement par sa facilité de mise en œuvre et par son coût algorithmique relativement faible, permettant ainsi d'envisager son application au suivi « temps réel » d'un objet en mouvement [32]. Cependant, ces modèles souffrent de certains problèmes maintenant bien connus tels :

- une forte sensibilité à l'initialisation du *snake*,
- un manque de robustesse vis-à-vis du bruit et d'un fond trop texturé, et
- une incapacité notable à gérer convenablement les objets présentant une forte concavité.

Malgré les améliorations comme le modèle des *balloons* [38] et celui plus récent du GVF (*Gradient Vector Flow*) [174] qui introduisent de nouvelles forces externes, certaines difficultés persistent qui ont amoindrit son attractivité ces dernières années ; ceci au profit d'un autre modèle de contour actif appelé « courbes géodésiques » (*level set* [29, 136]) dont le principal avantage par rapport aux *snakes* est de savoir gérer les changements de topologies et les objets multiples, en contre-partie d'une complexité algorithmique accrue.

Une autre solution qui a retenue récemment l'attention consiste à ajouter aux forces internes et externes des *snakes*, une nouvelle énergie qui contient une information *a priori* sur l'objet re-

cherché. Ces méthodes, qu'elles soient orientées contour ou région, peuvent être classées en deux catégories : celles utilisant un *a priori* statistique [167, 42, 37] et celles, plus récentes, intégrant un *a priori* géométrique [138, 80, 34] [6]. Notre contribution dans ce domaine a consisté à proposer une modèle de *snake* intégrant un *a priori* géométrique orienté contour, intégrant l'invariance à certaines transformations géométriques. Ainsi, dans le cadre de son master recherche [36], Mohamed Ali Chermi a proposé d'introduire une force additionnelle construite à partir de la description invariante du modèle et celle du contour à chaque itération de convergence du *snake*. Cette nouvelle force contraint le *snake* a évoluer vers une forme proche de celle du modèle [35, 34], indépendamment de leur pose respective. La représentation invariante que nous avons utilisée est inspirée des familles complètes et stables de descripteurs invariants introduites par F. Ghorbel à partir de la transformée de Fourier adéquate [87].

### 3.3.1 Le modèle original des *snakes*

**Description de la courbe** Un *snake* est représenté par une courbe paramétrée $\mathcal{C}$ :

$$\mathcal{C} = \{v(s,t) = x(s,t) + i\, y(s,t), s \in [0, L] \text{ et } t \in [0, T]\},$$

où $s$ est l'abscisse curviligne et $t$ représente le temps. Cette courbe peut être ouverte ou fermée. Sans perdre en généralité, la longueur de la courbe $L = \int v'(s)\ ds$ peut toujours être ramenée à $L = 1$ par normalisation ($s \in [0, 1]$), impliquant l'invariance de la représentation par rapport au facteur d'échelle. On notera $v(s,t)$ par $v(s)$ pour simplifier l'écriture dès lors que cela n'introduit pas de confusion.

Dans le modèle numérique, le temps $t$ représente l'itération du *snake* jusqu'à convergence, $T$ devenant le nombre d'itérations avant convergence. L'abscisse $s$ est discrétisée, aboutissant à paramétrisation euclidienne discrète du contour $v$ par un nombre fini $N$ de points : $v(n) = x(n) + i\, y(n)$, $n = 1, \ldots, N$. Pour garantir l'invariance de la description par rapport aux translations du plan, les coordonnées $v(n)$ sont exprimées par rapport au barycentre de $v$.

**Fonction d'énergie** Le *snake* se déplace sous l'action d'une fonctionnelle d'énergie composée de deux termes :

$$E^*_{snake} = \int_0^1 E_{int} + E_{ext}\, ds. \tag{3.12}$$

> **Énergie interne $E_{int}$ :** cette énergie ne dépend que de la courbe $v(s)$. Elle est constituée de la somme de deux termes correspondant aux dérivées premières et secondes de $v(s)$ par rapport à l'abscisse curviligne $s$
>
> $$E_{int} = \alpha \left\| \frac{\partial v(s)}{\partial s} \right\|^2 + \beta \left\| \frac{\partial^2 v(s)}{\partial s^2} \right\|^2.$$
>
> Son rôle est d'imposer des contraintes d'élasticité et de rigidité sur le *snake*. Le premier terme évite l'apparition des points isolés dans le contours. Le deuxième terme empêche l'apparition des angles aigus et des coins dans le contour. Les coefficients $\alpha$ et $\beta$ sont des poids généralement pris constants (indépendants de $s$ et de $t$).

6. — Des travaux récents tels [82, 108] intègrent les deux types d'*a priori* dans le même modèle.

**Énergie externe** $E_{ext}$ **:** cette énergie traduit la fidélité aux données et dépend donc de l'image. Elle permet de déplacer le *snake* vers les zones d'intérêts dans l'image qui peuvent être des lignes, des contours ou des terminaisons de droites. L'énergie externe la plus couramment utilisée est donnée par

$$E_{ext} = -|\nabla(G_\sigma * I)|,$$

où $\nabla$ désigne l'opérateur gradient, $G_\sigma$ est un filtre gaussien bi-dimensionnel d'écart type $\sigma$ selon les deux directions et $I$ est la matrice représentant les niveaux de gris de l'image.

La résolution du problème des *snakes* revient à minimiser la fonctionnelle d'énergie de l'équation (3.12). Plusieurs méthodes itératives ont été proposées, parmi lesquelles on compte la programmation dynamique [2], les éléments finis [39] et le calcul variationnel [99]. Cette dernière méthode, la plus utilisée, conduit à l'inversion, à chaque itération, d'une matrice carrée $N \times N$ pentadiagonale, avec une complexité $O(N)$. À chaque itération, on dispose des coordonnées de la nouvelle position du *snake*.

### 3.3.2 Introduction d'un *a priori* de forme dans les *snakes*

L'idée de la méthode proposée est d'ajouter une nouvelle force de forme ($E_{prior}$) construite à partir d'une description invariante des contours. Cette force permet de guider le *snake* vers la forme d'un objet de référence (ou modèle), indépendamment de leur orientation et taille respectives.

**Description invariante des contours** Nous considérons ici uniquement les transformations euclidiennes, même si la méthode n'est pas restreinte à ce type de transformations. Nous avons utilisé la famille de descripteurs invariants proposée par F. Ghorbel [87], dont l'expression pour un contour $v$ est la suivante :

$$\begin{aligned} I_{k_0}(v) &= |C_{k_0}(v)|^{k_0-k_1+p+q}, \\ I_{k_1}(v) &= |C_{k_1}(v)|^{k_0-k_1+p+q}, \\ I_k(v) &= \frac{C_k^{k_0-k_1}(v)\, C_{k_1}^{k-k_0}(v)\, C_{k_0}^{k_1-k}(v)}{|C_{k_1}(v)|^{k-k_0-p}\, |C_{k_0}(v)|^{k_1-k-q}}, \quad \text{pour } k \in [1,\ldots,N-1] \text{ et } k \neq k_0, k_1, \end{aligned} \tag{3.13}$$

avec $C_k(v)$ les coefficients de Fourier du contour $v$, $p$ et $q$ deux nombres réels strictement positifs et $k_0$ et $k_1$ deux entiers choisis tels que $k_0, k_1 \in [1,\ldots,N-1]$, $k_0 \neq k_1$, $C_{k_0}(v) \neq 0$ et $C_{k_1}(v) \neq 0$.

Cette famille de descripteurs est

**Invariante** aux rotations et au point initial de description du contour. L'invariance par rapport aux translations et au facteur d'échelle est assurée par la représentation barycentrique normalisée du contour.

**Complète (Crimmins [43])** c'est-à-dire que deux contours $v_1$ et $v_2$ ont la même forme si et seulement si $I_k(v_1) = I_k(v_2)$, $\forall k \in [1,\ldots,N-1]$. Ainsi, la reconstruction de la forme à partir de ses invariants et des phases de $C_{k_0}$ et $C_{k_1}$ est possible. Pour éviter toute ambiguïté lors de la reconstruction, nous choisissons $k_1$ tel que $k_0 - k_1 = 1$.

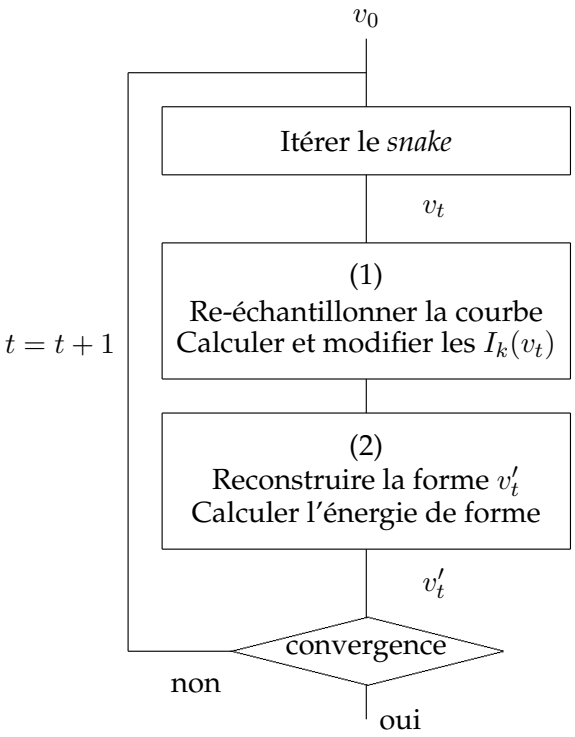



Figure 3.9 – Stratégie d'introduction des descripteurs invariants au modèle des *snakes*.

**Stable (Ghorbel [85])** une famille de descripteurs $\{I_k\}$ est stable si et seulement si ses descripteurs définissent une fonction continue de l'espace des paramétrisations dans l'espace des descripteurs. La stabilité se traduit par le fait qu'une faible variation de la forme induit une faible variation des invariants. Par rapport à la famille complète proposée par T. R. Crimmins, la stabilité est obtenue par l'ajout du dénominateur dans l'équation (3.13).

Des familles complètes et stables de descripteurs invariants aux transformations affines ont été proposées [31]. Leur utilisation comme *a priori* dans le modèle des *snakes* constitue une perspective immédiate à ce travail.

**Intégration de l'*a priori* invariant dans le modèle des *snakes*** Nous détaillons maintenant la stratégie adoptée pour ajouter cette information *a priori* dans le modèle des *snakes* (*cf.* figure 3.9). Soient $v_t$ le *snake* à l'itération $t$ et $v_{ref}$ le contour du modèle. L'introduction de l'information de forme à partir des invariants se fait en deux étapes :

> **1– Modification des invariants.** On calcule une combinaison des invariants du *snake* en cours et de la forme de référence selon
>
> $$I_k(v'_t) = c_1\, I_k(v_t) + (1 - c_1)\, I_k(v_{ref}). \tag{3.14}$$
>
> Étant données les propriétés des invariants choisis (complétude et stabilité), la courbe $v'_t$ obtenue par reconstruction de $I_k(v'_t)$ est une forme intermédiaire entre $v_t$ et $v_{ref}$. En considérant que $\arg C_{k_0}(v'_t) = \arg C_{k_0}(v_t)$ et que $\arg C_{k_1}(v'_t) = \arg C_{k_1}(v_t)$, on peut prétendre reconstruire $v'_t$ dans la même pose que $v_t$ (même orientation et même point initial de description des contours). De même, nous supposons que le barycentre et la taille de $v'_t$ sont identiques à ceux de $v_t$.

(a) Modèle (b) (c)

Figure 3.10 – Résultats de détection sur la forme en « U » pour (b) une translation et un facteur d'échelle, (c) un facteur d'échelle et une rotation, prenant (a) comme *a priori* de forme. Résultats extraits de [54].

**2– Calcul de l'énergie de forme $E_{prior}$.** On définit la force additionnelle contenant l'information *a priori* sur le modèle par

$$E_{prior} = c_2 \left(v'_t - v_t\right). \tag{3.15}$$

Le coefficient $c_2$, fixé de manière empirique, permet de contrôler l'influence de cette force par rapport aux énergies internes et externes intervenant dans le modèle.

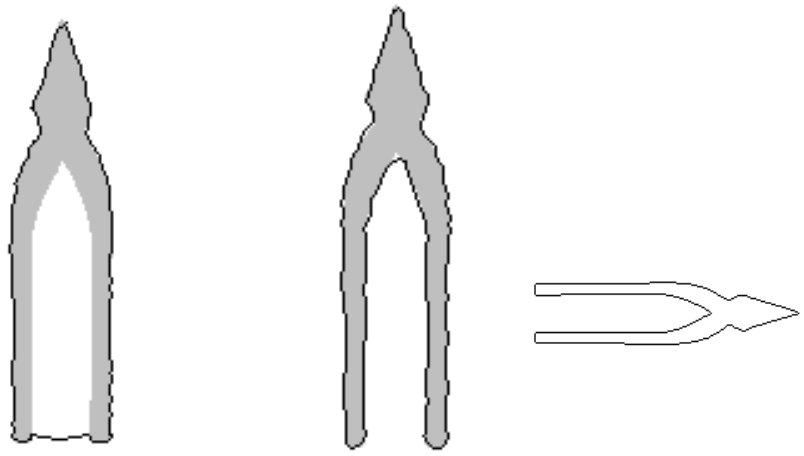
### 3.3.3 Illustration du comportement du *snake* avec *a priori*

L'objectif de cette section est de présenter quelques résultats expérimentaux qui illustrent la capacité du modèle à résoudre certaines lacunes du modèle des *snakes*, à l'aide de l'information supplémentaire donnée par l'*a priori*.

**Invariance de la force *a priori* et concavités** L'évolution de la courbe dans les zones concaves reste un handicap pour les *snakes*. La figure 3.10 illustre les résultats de segmentation de la forme « U » par la nouvelle méthode dans deux cas de transformations en prenant (a) comme image de référence. La méthode est capable de gérer la concavité quelque soit la transformation euclidienne qui est appliquée.

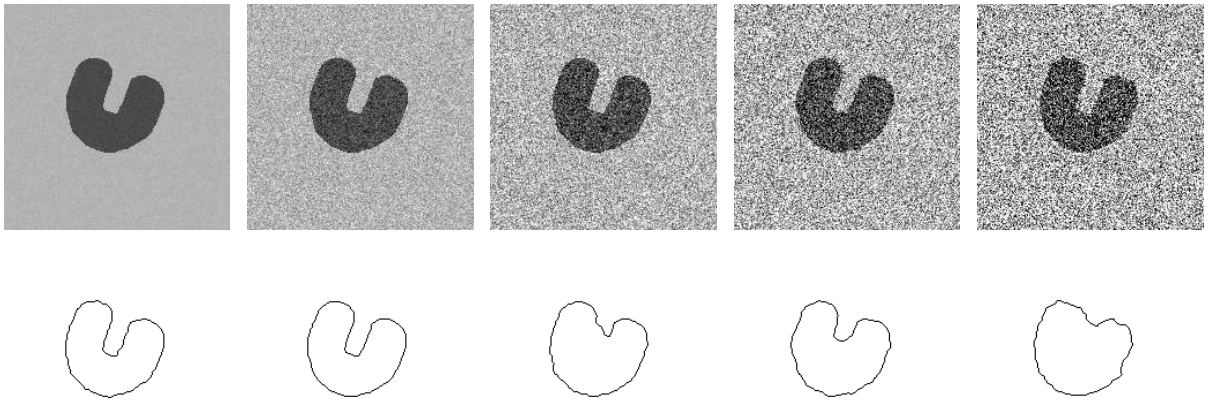
**Évolution dans de fortes concavités** Le modèle « GVF » est connu dans les littérature pour remédier au problème d'évolution des *snakes* classiques dans les concavités. Par contre, nous avons découvert qu'il le remédie que partiellement puisqu'en présence de concavités très profondes, l'algorithme n'évolue pas comme attendu, contrairement au modèle que nous proposons, *cf.* figure 3.11.

**Robustesse au bruit** Afin d'évaluer la robustesse de la méthode au bruit, nous avons effectué une série d'expériences. L'une des expériences consistait a restaurer la frontière d'une image de la lettre « U » bruitée avec deux gaussiennes telles que leurs moyennes et écart-types vérifient : $m_2 - m_1 = 180$ et $\sigma = \sigma_1 = \sigma_2$. La figure 3.12 présente les résultats obtenus pour différentes valeurs de $\sigma$. Comme attendu, plus $\sigma$ augmente et plus la détection se dégrade. Néanmoins, en doublant la valeur des poids $c_1$ et $c_2$ associés à la force additionnelle –*cf.* équations (3.14) et (3.15)– nous avons obtenu des résultats de détection très nettement meilleurs.



(a) (b)

Figure 3.11 – Évolution dans les zones fortement concaves par le modèle « GVF » (a) et par le modèle avec *a priori* (b). Résultats extraits de [54].



(a) $\sigma = 5$ (b) $\sigma = 25$ (c) $\sigma = 45$ (d) $\sigma = 65$ (e) $\sigma = 85$

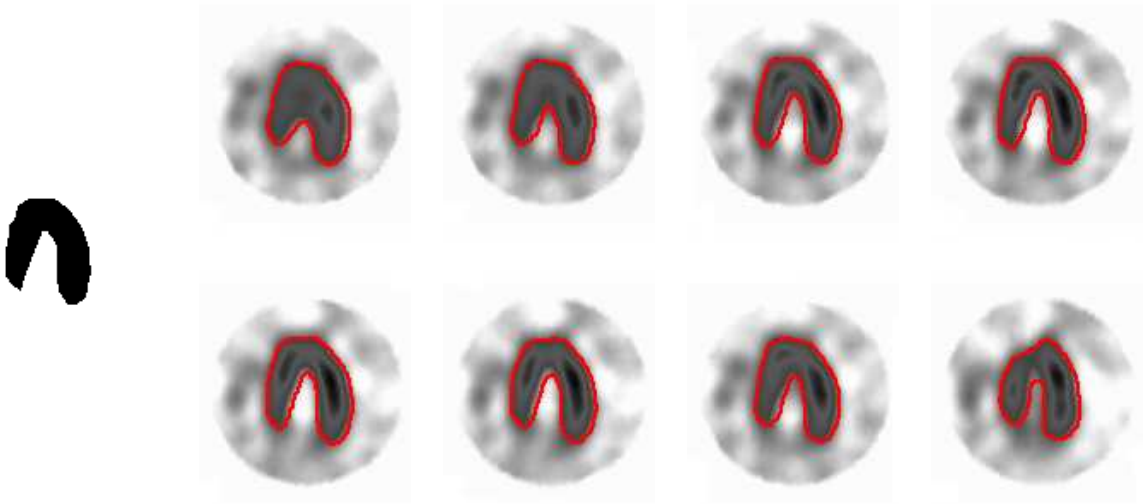
Figure 3.12 – Résultats de détection pour différentes valeurs de $\sigma$, avec $m2 - m1 = 180$, $c_1 = 0.12$, $c_2 = 0.25$ et $c_3 = 1$. Résultats extraits de [35].

### 3.3.4 Application à des séquences d'images

À titre d'illustration, nous présentons l'application du modèle de *snake* avec *a priori* dans deux contextes différents : la segmentation de coupes scintigraphiques du cœur et la poursuite d'un objet en mouvement dans une séquence vidéo.

**Segmentation de coupes scintigraphiques du myocarde**

Nous avons appliqué la méthode des *snakes* avec *a priori* de forme à la segmentation d'images scintigraphiques. La scintigraphie myocardique d'effort est un examen qui permet de visualiser les cellules fonctionnelles du cœur. Il consiste à injecter dans une veine du bras un produit faiblement radioactif qui est absorbé par les cellules cardiaques vivantes qui sont convenablement irriguées par leurs artères nourricières. Des images prises avec des caméras spéciales appelées



(a) Modèle, (b) Huit coupes scintigraphiques du ventricule gauche.

Figure 3.13 – Application des *snakes* avec *a priori* à la segmentation de coupes scintigraphiques du myocarde. Résultats de détection (b) à partir de l'*a priori* (a). Résultats extraits de [35].

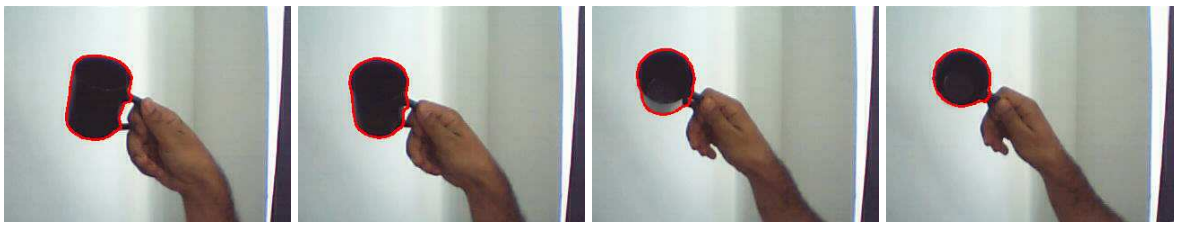
« gamma caméras » comptent alors la radioactivité présente au niveau du cœur. Plus la radioactivité est importante dans une zone, plus le nombre de cellules est important et plus la zone apparaît en rouge sur la photographie. Par contre, moins il y a de cellules, plus la zone devient bleue sombre. Pour aider au repérage des zones altérées, la photographie représente le cœur sous différentes coupes.

L'objectif est de segmenter le myocarde sur toutes les coupes. Ce type de résultat peut servir pour le diagnostic d'ischémie en comparant le myocarde au repos et à l'effort. Pour cela, nous avons appliqué le modèle des *snakes* avec *a priori*, en utilisant comme référence pour les huit images traitées la même forme approximative du myocarde, *cf.* figure 3.13. Nous avons comparé ces résultats avec ceux obtenus par une méthode de lignes de partage des eaux, corrigés et validés par un expert. Ils ont été jugés satisfaisants par l'expert.

**Poursuite d'un objet en mouvement**

La deuxième application concerne l'utilisation de l'*a priori* de forme pour le suivi d'un objet en mouvement dans une séquence vidéo (monochrome). L'objectif est d'évaluer la robustesse de la méthode pour un type de transformation plus général que les transformations euclidiennes que nous considérons dans le modèle.

Le principe est le suivant : le *snake* de la trame $t-1$ est utilisé comme *a priori* de forme pour faire converger le *snake* de la trame $t$. Dans ces conditions, nous faisons l'hypothèse que le mouvement de l'objet est faible par rapport à la cadence d'acquisition. Ainsi, l'objet garde presque la même forme entre deux trames consécutives. La figure 3.14 montre le résultat de suivi d'une main dans une séquence vidéo acquise par une caméra de type Webcam (faible résolution). Dans cette application, nous n'utilisons pas de méthode d'estimation de mouvement ; l'utilisation de méthodes de prédiction, comme le filtre de Kalman, pourrait permettre d'améliorer les résultats de suivi (en particulier pour les mouvements plus rapides).



(a) Trame 35 (b) Trame 68 (c) Trame 139 (d) Trame 191

Figure 3.14 – Application des *snakes* avec *a priori* au suivi d'une tasse en mouvement à partir d'une caméra fixe. $c_1 = 0.1$, $c_2 = 0.25$ and $c_3 = 1$. Résultats extraits de [35].

## 3.4 Mélange non-gaussien et poursuite d'objets dans une vidéo (thèse de W. Ketchantang)

Le dernier thème abordé dans ce chapitre concerne la poursuite d'un objet d'intérêt dans une séquence animée. Cette problématique a été motivée par le sujet de thèse de W. Ketchantang [104] qui a porté sur le développement d'une méthode « grand public » de reconnaissance d'iris, pour des applications embarquées d'authentification comme le déverrouillage d'un téléphone portable.

Les systèmes de reconnaissance d'iris sont considérés parmi les plus fiables lorsqu'on les compare à d'autres systèmes de biométrie physique basés sur les empreintes digitales, la forme d'un visage ou d'une main [159, 46] [7] car la texture de l'iris est particulièrement stable au cours de la vie et qu'elle est difficile à falsifier. Avec l'expiration du brevet de Flom and Safir [78] sur la reconnaissance d'iris en 2005, et à cause des enjeux commerciaux/sécuritaires très importants auxquels ils sont à l'origine, de nombreux algorithmes de détection, localisation et reconnaissance d'iris ont été proposés ces dernières années [16]. Pour comparer objectivement ces méthodes, des compétitions à très grande échelle sont régulièrement organisées. Nous pouvons citer le projet *Iris Challenge Evaluation* (ICE [8]) proposé par le *National Institute of Standards and Technology* (NIST).

Par contre, pour être performant, ces systèmes de reconnaissance d'iris nécessitent une coopération stricte de l'utilisateur pour l'acquisition de l'image de leur iris. Il n'est pas rare que plusieurs prises de vues soient nécessaires avant d'obtenir une image exploitable par le système de reconnaissance. En effet, de nombreux artéfacts viennent entacher l'image comme

1. le flou lié au mouvement de l'utilisateur,
2. les occlusions partielles de l'iris par les cils et les paupières,
3. la défocalisation de l'iris par rapport à l'optique,

auxquels s'ajoutent les problèmes inhérents à la qualité des optiques des capteurs grand-publics de type webcam : faible résolution et présence importante de bruits. L'objectif des recherches actuelles sur la reconnaissance d'iris est de rendre le système le plus robuste possible à tous ces artéfacts, tout en le rendant moins contraignant pour l'utilisateur. Pour cela, dans le cadre de sa thèse développée en collaboration avec la société ST Microelectronics, William Ketchantang [104] a proposé de remplacer l'appareil photographique par une caméra vidéo sensible à l'éclairage infrarouge et de sélectionner les meilleurs images d'iris au fur et à mesure du flux vidéo. L'acquisition est stoppée lorsqu'un certain nombre d'images ont été jugées exploitables par le système.

7. — D'autres exemples concernent les biométries *physiologique* comme l'électroencéphalogramme ou l'électrocardiogramme et *comportementale* comme la posture, la voix ou une signature manuscrite.

8. — `http://iris.nist.gov/ice/`

Dans ce cadre, nous avons proposé un nouvel algorithme « temps réel » de suivi d'objets en couleur [102]. Le système est basé sur une caractérisation de l'objet d'intérêt par un mélange de couleurs [116, 115, 137], et son suivi par un filtre de Kalman pour améliorer la détection à chaque image et réduire les temps de calculs. L'originalité de l'approche repose principalement sur la modélisation du mélange de couleurs par des copules gaussiennes (*cf.* paragraphe 2.2.3, page 31) et des marges dans le système de Pearson. Ainsi, ce n'est pas tant la forme de l'objet qui est poursuivie, mais une configuration de couleurs présentes dans la scène, indépendamment de l'organisation spatiale des couleurs entre-elles.

Nous commençons par présenter le principe de la méthode de suivi, valable pour tout type d'objet couleur, puis nous présentons son application pour la reconnaissance d'iris à l'aide de séquences vidéo.

### 3.4.1 Suivi colorimétrique d'un objet

L'idée principale du suivi consiste à caractériser la distribution de couleurs de l'objet d'intérêt par un mélange de lois, puis de mettre à jour les paramètres du mélange afin de minimiser l'erreur de prédiction sur la position de la cible dans la trame suivante, et ainsi de réduire les temps de calcul.

Plusieurs approches ont été proposées pour caractériser la distribution de couleurs d'un objet. Parmi celles-ci, on peut citer N. S. Peng *et al* [140] qui proposent d'utiliser l'algorithme des *mean shift* pour modéliser l'histogramme de l'objet, ou le modèle de mélange gaussien utilisé par S. J. McKena *et al* [116], M. Störring *et al* [168], G. Xiong *et al* [173]. La plupart du temps, des lois gaussiennes sont utilisées pour caractériser la distribution de couleurs. Ce choix est motivé par les hypothèses communément acceptées sur les scènes naturelles, le bruit des capteurs et les caractéristiques de l'espace couleur RGB. Un fait important qui promeut l'utilisation de gaussiennes est leur facilité d'utilisation et leur faible temps de calcul. Mais, cette hypothèse peut parfois apparaître trop restrictive. En effet, elle n'est pas adaptée à certains capteurs, comme les caméras infra-rouge ou certaines modalités d'images médicales. Il est également habituel de convertir les données d'un format couleur vers un autre pour rendre la poursuite plus robuste aux changements d'illumination. Ainsi, certaines transformations non linéaires entre les espace couleurs peuvent rendre la modélisation gaussienne inadaptée aux données.

Nous avons donc proposé de modéliser la distribution de couleurs par un mélange de $K$ copules gaussiennes $M$-dimensionnelles ($M$ désignant le nombre de bande de l'espace couleur) avec des marges appartenant au système de Pearson :

$$\begin{aligned}\forall\, \mathbf{x} = (x_1, \ldots, x_M)^t \in \mathbb{R}^M, \quad p\left(\mathbf{x}\,|\Theta\right) &= \sum_{k=1}^{K} \alpha_k \, f_k\left(\mathbf{x}\,|\theta_{k,1}, \ldots, \theta_{k,M}, \boldsymbol{\rho}_k\right) \\ &= \sum_{k=1}^{K} \alpha_k \prod_{m=1}^{M} f_{k,m}\left(x_m\,|\theta_{k,m}\right)\; c\left(x_1, \ldots, x_M; \boldsymbol{\rho}_k\right),\end{aligned}$$

où l'ensemble $\theta_{k,m}$ contient les quatre moments définissant la $m^{\text{e}}$ marge $f_{k,m}$ de la $k^{\text{e}}$ densité multidimensionnelle $f_k$ du mélange ; $c(\ldots, \boldsymbol{\rho}_k)$ désignant la copule gaussienne de matrice de corrélation $\boldsymbol{\rho}_k$).

Les paramètres du modèle sont estimés à l'aide de l'algorithme SEM, en considérant que les pixels sont indépendants. Ils sont mis à jour automatiquement au fur et à mesure du flux vidéo, de

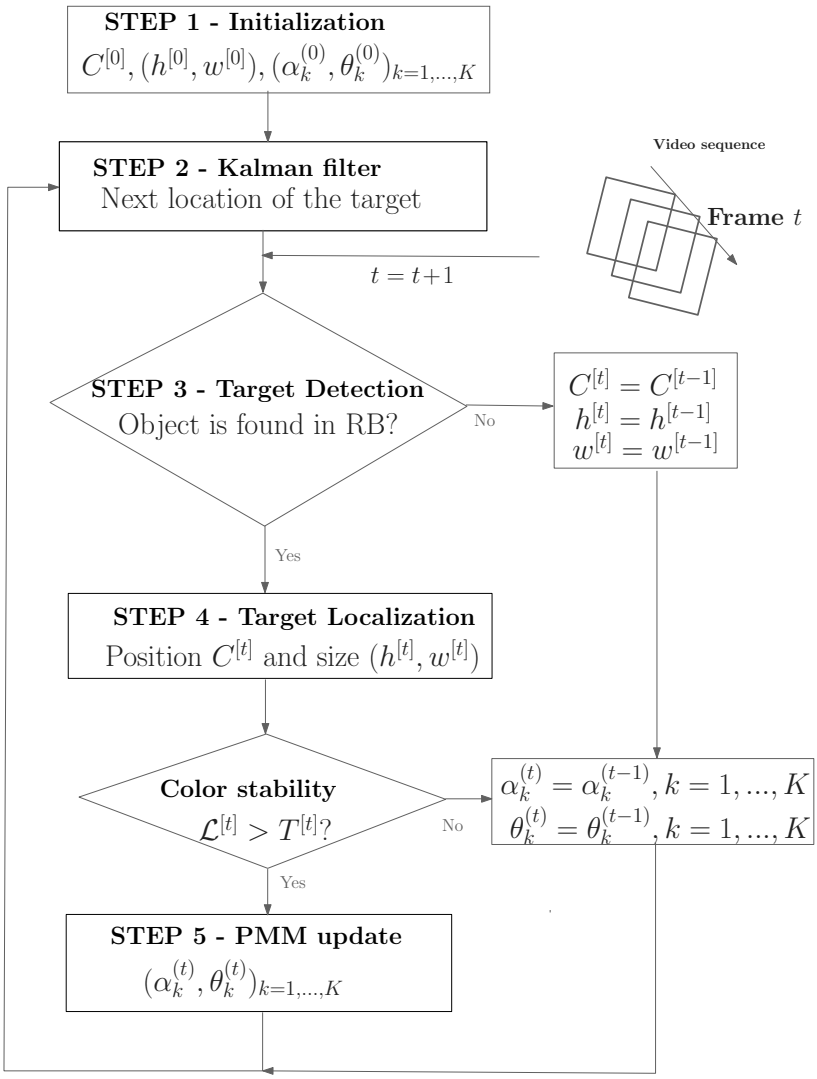
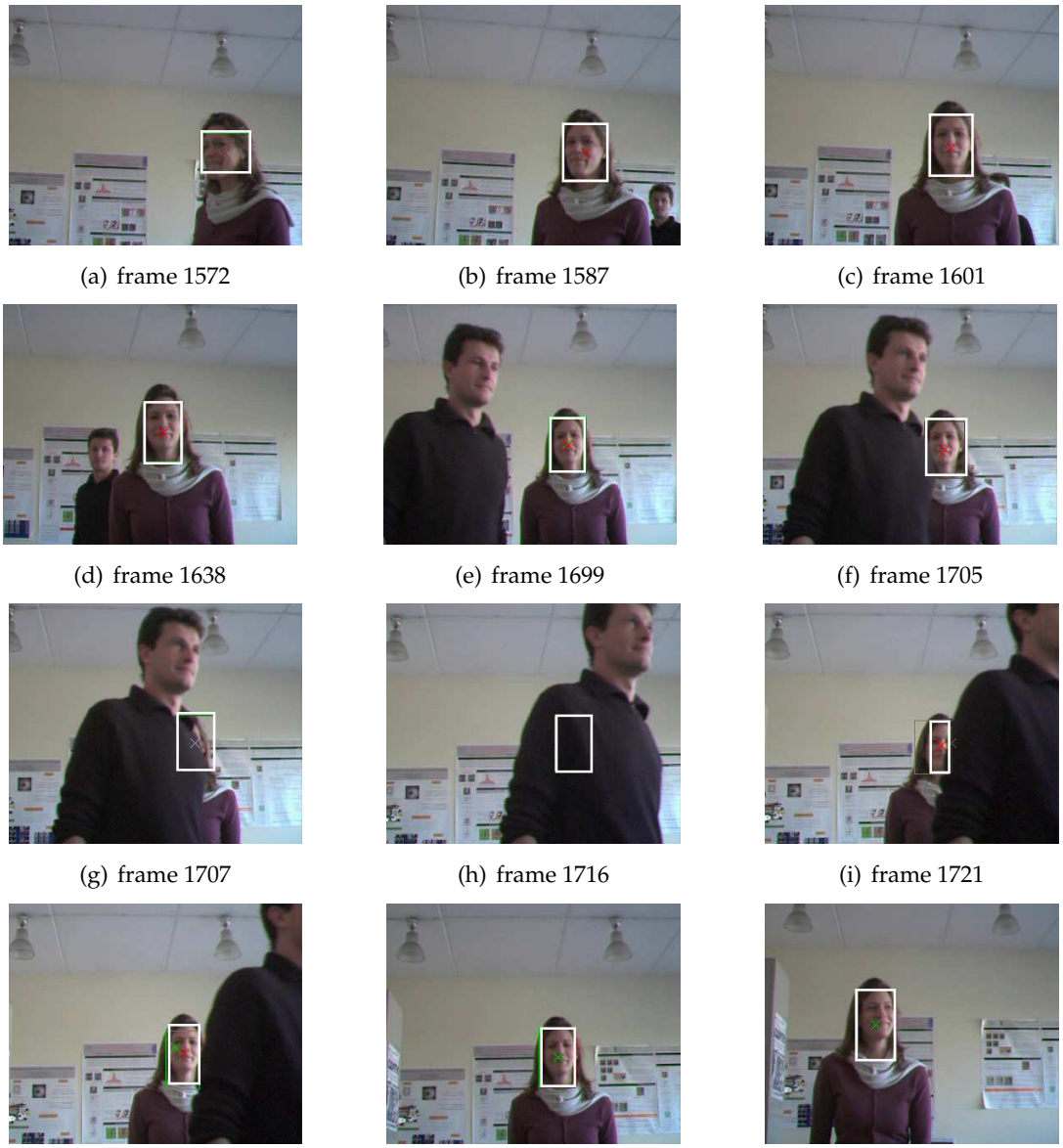



Figure 3.15 – Synopsis général de l'algorithme de poursuite de cible couleur. Les lettres $h$, $w$ et $C$ réfèrent respectivement à la hauteur, la largeur et le centre de la fenêtre de localisation de la cible. La log-vraisemblance du modèle $\mathcal{L}$ est comparée à un seuil estimé dynamiquement à chaque itération. Résultats extraits de [102].

telle sorte que si la vue apparente de l'objet sur une trame est modifiée (apparition de parties cachées de l'objet par exemple), le modèle continue à suivre l'objet en s'adaptant automatiquement aux légers changements de couleurs. Si le changement d'une trame à une autre est trop important, le système considère avoir perdu l'objet (occlusion totale ou mouvement trop brusque de la cible) et se stabilise dans une position d'attente, prêt à continuer la poursuite dès que l'objet réapparaît [9]. Le squelette général de l'algorithme de suivi est présenté dans la figure 3.15.

À titre d'illustration, la figure 3.16 présente le résultat de suivi d'un visage sur une séquence couleur de type webcam (résolution $352 \times 288$ pixels). L'objet d'intérêt est modélisé par un mélange de deux lois bi-dimensionnelles. Les occlusions partielles et totales du visage par un objet masquant en mouvement est géré convenablement par le système qui continue à poursuivre la cible dès lors qu'elle est présente dans la scène. Le modèle classique utilisant des gaussiennes

9. — sous l'hypothèse que celui-ci réapparaît approximativement à l'endroit de sa disparition.



(a) frame 1572 (b) frame 1587 (c) frame 1601

(d) frame 1638 (e) frame 1699 (f) frame 1705

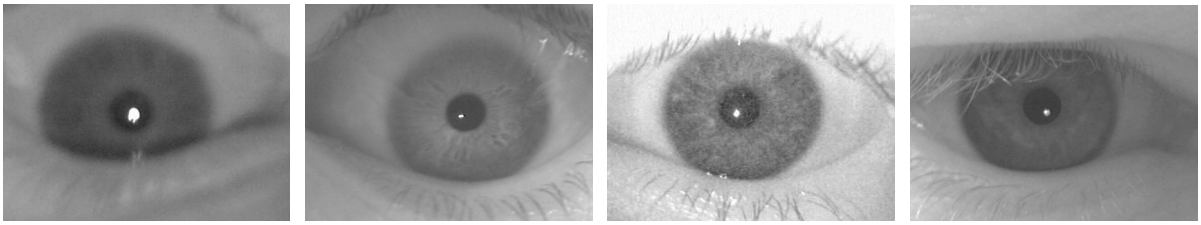
(g) frame 1707 (h) frame 1716 (i) frame 1721

(j) frame 1722 (k) frame 1748 (l) frame 1863

Figure 3.16 – Poursuite d'un visage partiellement et totalement masqué. Les rectangles qui apparaissent sur chacune des images modélisent les fenêtres de prédiction et de localisation de l'objet d'intérêt. Résultats extraits de [104].

multidimensionnelles ne permet pas d'obtenir un suivi aussi robuste.

Cette méthode présente l'avantage, d'une part, d'être indépendante du choix de la cible car cette dernière est caractérisée par des attributs statistiques et non géométriques et, d'autre part, d'être relativement robuste aux occlusions partielles de la cible, aux variations de taille de l'objet dans l'image, aux faibles mouvements de la caméra et aux mouvements rapides de la cible. Le temps de traitement dépend essentiellement de la taille de l'objet suivi. Pour la séquence présentée en exemple, la cible est suivie à la cadence de 12 images par seconde (PC avec Pentium IV à 2.6 GHz). De nombreuses expérimentations ont permis de valider l'algorithme pour des objets noyés dans des scènes complexe, et pour des objets de forme et de chrominance variées.



(a) Floue (b) Focus (c) Bruit (d) Occlusion

Figure 3.17 – Quatre cas d'iris dégradée lors de l'acquisition. Résultats extraits de [100].

Nous avons appliqué ce modèle au système de reconnaissance d'iris basé vidéo développé par W. Ketchantang. Les images du système vidéo étant monochromes (infra-rouge), l'iris est caractérisé par une mélange de lois monodimensionnelles issues du système de Pearson.

### 3.4.2 Reconnaissance de l'iris pour l'authentification des personnes

La division AST de la société ST Microelectronics a conçu une plate-forme logicielle et matérielle complète d'acquisition et d'analyse d'iris. Malgré les bonnes performances des algorithmes de reconnaissance, l'acquisition d'images d'iris de bonne qualité avec une caméra à faible coût reste un problème difficile à résoudre. En effet, la qualité des images est un facteur important qui diminue notablement les performances du système avec notamment un taux de faux rejet (« FRR [10] ») important. La figure 3.17 illustre quatre cas typiques d'artéfacts rencontrés pendant l'acquisition. Dans chacun des cas, un spot blanc de taille et de forme différentes apparaît au niveau de la pupille. Ils proviennent de la réflexion spéculaire de l'éclairage infra-rouge. Les séquences vidéo montrent généralement

- un mouvement aléatoire de la tête de l'utilisateur dans un plan parallèle au plan optique de la caméra, résultant en de larges et rapides translations de la pupille ;
- une translation de la tête de l'utilisateur le long de l'axe optique, résultant en (i) des variations importantes de la taille de la pupille, et (ii) une alternance de séries d'images focalisées et défocalisées ;
- des mouvements de rotation de faible ampleur de la tête.

Aussi est-il nécessaire de savoir mesurer la qualité de l'iris au niveau de la pupille pour réaliser une sélection des meilleurs images, c'est-à-dire des iris nettes et non obstruées par les paupières. La solution adoptée jusqu'à présent consistait à enregistrer une vidéo de l'iris de l'utilisateur désirant s'authentifier pendant un temps donné (20 secondes par exemple) et de post-traiter les images pour déterminer lesquelles seront exploitées par le système de reconnaissance. Cette solution manque de souplesse et nous avons développé un module de sélection des meilleurs images d'iris au fur et à mesure du flux, à la cadence vidéo (*cf.* figure 3.18). Cette solution permet de stopper l'acquisition lorsque le nombre d'images jugées de bonne qualité est atteint ($M$), avant de procéder à la reconnaissance. Les travaux nous ont conduits à enregistrer deux brevets [103, 114]. Pour juger de la qualité de l'iris, il est nécessaire de savoir repérer dans chaque trame la pupille.

10. — FRR : False Rejection Rate.

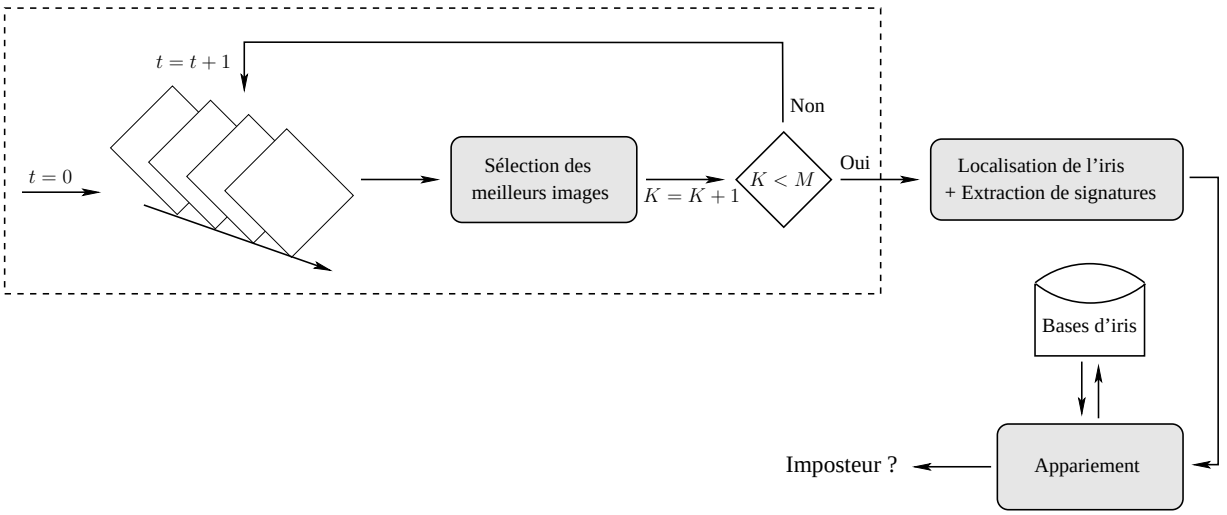



Figure 3.18 – Squelette de l'algorithme de sélection des meilleurs images d'iris au fur et à mesure du flux vidéo, avant la phase d'authentification.

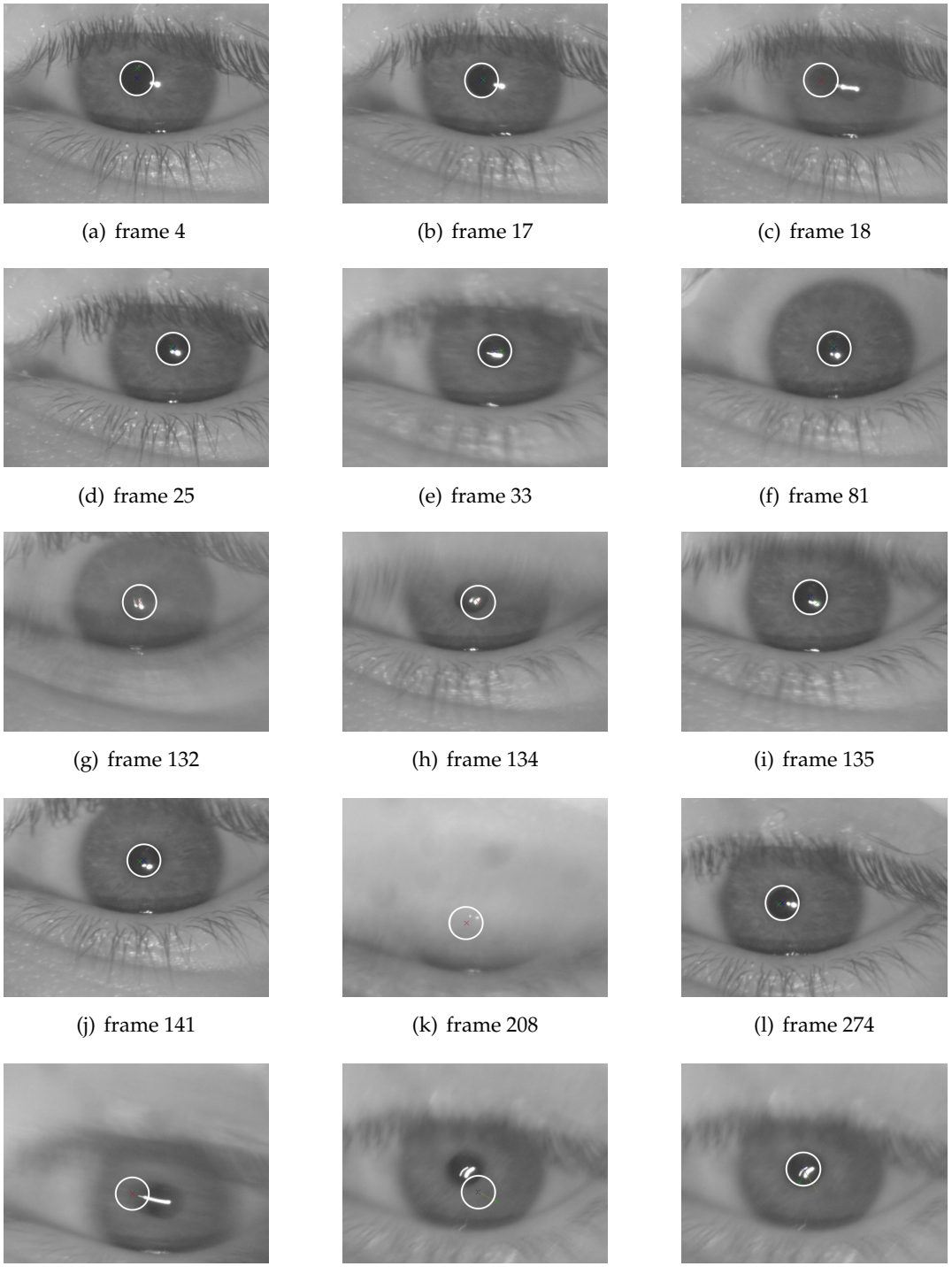
Cela est réalisé à l'aide du modèle présenté dans la section précédente, en intégrant des informations sur la forme approximativement circulaire de la pupille pour améliorer la robustesse du système [100]. Un exemple de poursuite est montré dans la figure 3.19. La méthode de sélection des images en fonction de la netteté de l'iris est décrite dans [101]. L'analyse de netteté et la décision d'acceptation ou de rejet prennent $20ms$ en moyenne.

Une fois que le système a sélectionné $M$ images d'iris de bonne qualité, le module de reconnaissance est activé (*cf.* figure 3.18). Le principe de reconnaissance d'une image d'iris à partir d'un base pré-enregistrée rejoint celui de l'indexation de bases d'objets exposé dans la section 3.2. Différents types de signature caractéristique de l'iris ont été proposés dans la littérature, depuis les travaux initiaux de J. Daugman [44], qui a proposé en 1992 son fameux *iris code* basé sur la transformée de Gabor. Le système actuel de reconnaissance développé à ST Microelectronics est basé sur les travaux de thèse de C.-L. Tisse [170], qui utilise la concept d'image analytique, introduit par J.-P. Havlicek *et al* [92] par extension de la transformée de Hilbert aux images, et le codage des fréquences émergentes de l'iris. Dans ce schéma, nous avons proposé de remplacer la transformée de Fourier par la transformée de Fourier-Mellin (*cf.* section 3.1) mieux adaptée à la prise en compte des dilatations de l'iris et des rotations de la tête.

Actuellement des tests de validation et de fiabilité de grande ampleur sont en cours de réalisation sur plusieurs bases d'images :

- Une base propre de séquences (appelée STIRIS) acquises lors d'une campagne réalisée en décembre 2005 : 145 vidéo d'iris (500 images) provenant de 37 personnes (œil droit / œil gauche).
- Les bases d'images fixes d'iris CASIA, UBIRIS et BATH qui servent de référence pour comparer les différents algorithmes de reconnaissance proposés dans la littérature.

Après avoir déterminé les valeurs optimales des paramètres de sélection à partir de bases d'apprentissage, le système complet d'identification a permis d'atteindre, sur la première base d'iris, les résultats d'authentification suivants :



(a) frame 4 (b) frame 17 (c) frame 18

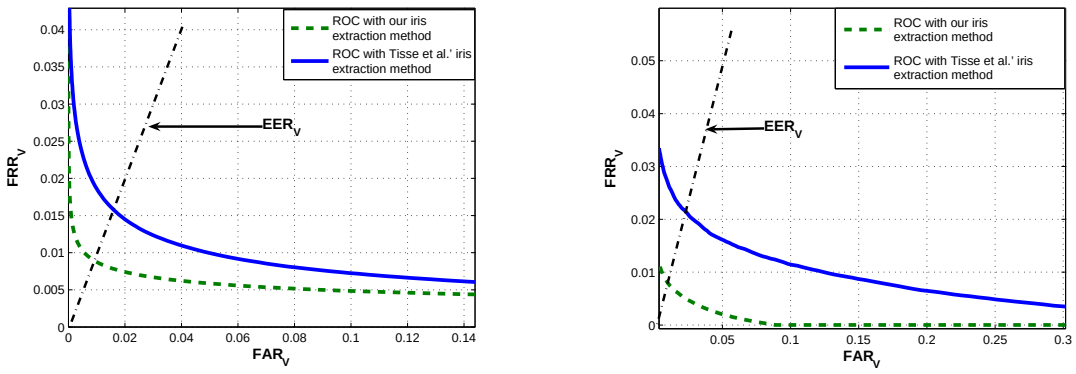
(d) frame 25 (e) frame 33 (f) frame 81

(g) frame 132 (h) frame 134 (i) frame 135

(j) frame 141 (k) frame 208 (l) frame 274

(m) frame 307 (n) frame 327 (o) frame 328

Figure 3.19 – Résultat de suivi d'une pupille illustrant le comportement de l'algorithme dans différentes situations (occlusions, flou, mouvement rapide, ...). L'algorithme perd temporairement la pupille, puis la retrouve systématiquement. Résultats extraits de [100].




(a) STIRIS (b) BATH

Figure 3.20 – Courbes ROC montrant l'évolution du taux de fausse alarme (FAR) en fonction du taux de faux rejet (FRR) pour deux bases d'iris. EER désigne le taux d'égale erreur (FAR=FRR). Résultats extraits de [104].

- ✓ Taux de fausse alarme d'identification : $0.023\%$.
- ✓ Taux de faux rejet d'identification : $3.2\%$.

Ces taux améliorent très nettement les résultats obtenus précédemment (*cf.* figure 3.20) et sont tout à fait compétitifs au niveau international. Certaines améliorations restent encore à réaliser avec, par exemple, l'utilisation d'un modèle de rupture en conjonction avec le filtre de Kalman [10] pour améliorer la modélisation du mouvement saccadé de l'œil lors de l'acquisition. Également, l'intégration de la dernière génération de capteurs vidéo CMOS offrant une résolution de 3 Mégapixels devrait permettre d'améliorer la qualité des images et rendre ainsi la sélection des iris plus facile et plus rapide.

## 3.5 Conclusion & Perspectives

L'ensemble des travaux présentés dans ce chapitre a trait à la représentation et à la caractérisation des formes 2D, avec deux aspects principaux : la reconnaissance et la détection de formes d'une part, et le suivi d'objets en mouvement d'autre part. Les outils mis en œuvre relèvent de la théorie de l'analyse harmonique et des moments algébriques, pour des applications d'indexation de bases d'objets, de poursuite d'objets d'intérêt et de reconnaissance biométrique. À titre de conclusion, il nous paraît maintenant utile de revenir sur quelques points importants et de mettre en valeur quelques perspectives à court terme que l'on peut envisager pour la poursuite des travaux engagés.

Le thème de l'invariance est central en traitement d'image et notre contribution au sujet concerne la définition de nouvelles familles complètes de descripteurs de formes 2D et leur utilisation dans de nouveaux contextes.

Dans les deux premiers sujets abordés dans ce chapitre, il s'agit de l'invariance par similitudes pour des applications de reconnaissance de formes 2D et d'indexation de bases d'objets à niveaux de gris isolés. Des études systématiques concernant leurs comportement numérique ont été menées. Ces résultats sont à l'origine de travaux en tatouage d'images à niveaux de gris [160] ou en

reconnaissance de cellules bio-médicales [72]. Leur utilisation sur des images satellitaires à très haute résolution, pour des applications de reconnaissance de structures constitue un projet de recherche que je compte aborder dans le cadre du programme franco-italien ORFEO (« *Optical and Radar Federated Earth Observation* »).

Dans le troisième sujet, c'est l'invariance par similitudes des formes-contours qui est utilisée pour détecter et suivre un objet d'intérêt noyé dans une scène complexe, à partir d'un modèle *snakes* avec *a priori* géométrique de forme. Il est tout a fait envisageable d'étendre le principe en incorporant un *a priori* géométrique

- ✓ contour, invariant aux transformations affines, en utilisant les familles complètes de descripteurs proposées dans [31] ;
- ✓ à niveaux de gris, invariant aux similitudes planes, en utilisant l'une des deux représentations originales exposées ci-dessus.

Nous poursuivons actuellement ces travaux prometteurs pour la détection d'un objet dans un couple d'images stéréoscopiques calibrées. Le principe consiste à utiliser le *snake* résultat de l'image de gauche pour initialiser celui de l'image de droite, en utilisant les lignes épipolaires des images rectifiées, conduisant à des résultats de détection plus robustes que si les détections avaient été réalisées indépendamment sur chaque image. Ensuite, on peut envisager une poursuite d'un objet sur une séquence stéréo en utilisant simultanément l'*a priori* de forme entre le couple d'images stéréoscopiques rectifiées de la trame $i$ et entre les vues gauches des trames $i$ et $i+1$.

La quatrième et dernière thématique utilise les principes de l'invariance de la distribution de couleurs d'un objets pour le suivre au cours du temps, utilisant des méthodes statistiques présentées dans le chapitre 2. Ce dernier point, à l'intersection entre modélisation statistique et reconnaissance géométrique de formes est l'un des projets de rechercheque nous abordons maintenant dans la conclusion générale de ce document.

CHAPITRE 4

# Conclusion Générale et Projets de Recherche

AU TERME de la rédaction de ce mémoire de HDR, nous pouvons dresser un bilan synthétique des différentes activités de recherche développées depuis neuf années, dont sept passées au sein de l'équipe « Groupe Signaux Multidimensionnels » de l'Institut Fresnel. Les outils que nous avons proposés ont bien souvent été développés dans le cadre de collaborations académiques (ENST Bretagne, INT, ENSI, CNES) et industrielles (ST MicroElectronics, Shaktiware) et par le biais d'encadrements doctoraux (R. Mezhoud, C. Carincotte et W. Ketchantang) ; je souhaite souligner ici l'importance de ces contributions dans les résultats présentés. Nous élaborons ensuite les projets de recherche envisagés, notamment les sujets combinant segmentation d'images et reconnaissance de formes, dans le cadre applicatif des images très haute resolution des futurs capteurs optique et radar.

## 4.1 Synthèse des résultats et projets en cours

Cette synthèse est l'occasion de dresser nos principales contributions et de mettre en évidence certaines perspectives pour lesquelles des travaux en cours tentent d'apporter une réponse.

### Segmentation statistique d'images

La première partie de ce document expose plusieurs extensions du modèle des chaînes de Markov cachées (CMC) pour la segmentation d'images par le biais du parcours de Hilbert-Peano. Ces extensions portent sur des modifications soit de la modélisation des données observées avec le modèle de *chaîne de Markov vectorielle* et des données cachées avec le modèle de *chaîne de Markov floue*, soit de la topologie de la chaîne –et donc des hypothèses de dépendance statistique sous-jacentes–, aboutissant aux modèles appelés *chaîne de Markov d'ordre supérieur* et *chaîne de Markov couple*. Pour tous ces modèles, la démarche originale reste essentiellement la même, en adaptant la définition des probabilités *avant-arrière* à chaque structure. Elle permet de conduire les calculs, pour l'estimation comme pour la segmentation, évitant ainsi les méthodes par simulations que l'on rencontre dans les modèles par champs de Markov. Ainsi, pour chacun d'eux nous avons proposé une méthode d'estimation des paramètres basée sur le principe de ECI.

Ces modèles sont évalués dans le cadre de la segmentation d'images radar et de la détection de changements lors de catastrophes naturelles. Mais leur très grande généralité ne les limite pas au domaine du traitement d'images, et devrait bénéficier aussi aux champs classiques d'utilisation des modèles de Markov (économie, sciences de la santé, ...) ; en particulier à la reconnaissance de la parole, précurseur dans la communauté du traitement signal, et à l'origine de nombreuses

extensions du modèle original [155, 91]. Notons toutefois que le traitement d'images satellitaires procure un cadre particulièrement riche pour l'évaluation de ces modèles, d'une part, pour la complexité et la nature fortement bruitée des données et, d'autre part, pour la diversité des applications proposées, toujours renouvelées par des progrès technologiques continus.

Nous avons évalué nos modèles sur des images de dimensions modérées ou sur des extraits de plus grandes images (typiquement $512 \times 512$ pixels). Or, la prochaine génération de capteurs produira des images de dimensions supérieures à $10000 \times 10000$ pixels qui ne pourront pas être traitées par ces modèles du fait de la quantité de mémoire nécessaire au fonctionnement des algorithmes. En effet, ils nécessitent la mémorisation de tableaux bidimensionnels de probabilités, dont la taille est proportionnelle au nombre de pixels de l'image traitée. À ce problème d'ordre technique, s'ajoute celui plus fondamental lié à l'hypothèse de stationnarité. Celle-ci s'avère en effet parfois trop restrictive ou inadaptée pour décrire les scènes à segmenter, comme c'est le cas d'une coulée de lave le long des pentes d'un volcan ou d'une inondation.

Pour répondre à ces deux verrous méthodologiques, dans le cadre de son stage de master recherche, Z. Bouyahia [15] a proposé très récemment une solution pragmatique qui consiste à utiliser une fenêtre glissante sur l'image, de dimension faible devant les dimensions de l'image. Dans ce modèle, l'image est scannée pixel par pixel selon le parcours de Hilbert-Peano. Pour chaque pixel, une sous-chaîne est extraite, sur laquelle nous utilisons le modèle des CMC pour l'estimation des paramètres. La décision de classification selon le critère MPM est prise uniquement sur le pixel d'intérêt. Ainsi, en déplaçant la fenêtre sur tous les pixels de l'image, nous obtenons une carte de segmentation non stationnaire car le jeu de paramètres est ré-estimé sur chaque fenêtre [14].

### Reconnaissance de formes 2D

La seconde partie de ce document traite de la reconnaissance de formes 2D, avec pour thème centrale l'invariance géométrique. Les outils mis en œuvre relèvent de la théorie de l'analyse harmonique et des moments algébriques, pour des applications d'indexation de bases d'objets, de poursuite d'objets d'intérêt et de reconnaissance biométrique.

Dans un premier temps, nous avons proposé de nouvelles familles complètes de descripteurs de forme invariants aux similitudes. Celles-ci sont issues de la *transformée de Fourier-Mellin analytique* et des *moments complexes* et évaluées dans le cadre d'applications de reconnaissance de formes et d'indexation de bases d'objets à niveaux de gris. Ces travaux ont permis de proposer un algorithme de détection des axes de symétrie par réflexion et par rotation d'un objet.

Ensuite, c'est une description des formes-contours, invariante aux similitudes planes, que nous avons utilisée pour proposer une méthode de détection d'un objet d'intérêt dans une image. Celle-ci repose sur un modèle de *snakes* dont l'évolution est contrainte par un *a priori* géométrique, constitué de la forme approximative de l'objet recherché. Nous poursuivons actuellement ces travaux prometteurs pour la détection d'un objet dans un couple d'images stéréoscopiques calibrées [34]. Le principe consiste à utiliser le *snake* résultat de l'image de gauche pour initialiser celui de l'image de droite, en utilisant les lignes épipolaires des images rectifiées. Ensuite, on peut envisager la poursuite d'un objet dans une séquence stéréo en combinant simultanément l'*a priori* de forme entre le couple d'images stéréoscopiques rectifiées de la trame courante et entre les vues gauches des trames courante et suivante.

La poursuite d'objet est également le thème abordé dans la dernière étude. Celle-ci a été mo-

(a) Toulouse (urbain, péri-urbain) (b) Roujan (péri-urbain, vignes)

Figure 4.1 – Simulations Pléiades (images PELICAN), panchromatique $0,70m$, Taille : $1460 \times 1460$. Programme d'accompagnement d'*ORFEO*, CNES, 2004.

tivée par une application de poursuite d'un iris en temps réel, pour un projet de reconnaissance biométrique grand-public. Ainsi, dans ce travail, ce n'est pas tant la forme de l'objet que nous exploitons mais sa distribution de couleurs. Celle-ci est modélisée par un mélange non gaussien issu des travaux sur les copules que nous avons présentées dans la première partie. Ce modèle s'adapte automatiquement aux déformations perçues de l'objet dans le plan de la caméra lorsqu'il se déplace dans l'espace.

## 4.2 Projets de recherche

Après avoir abordé les pistes de recherche à court terme concernant chacune des deux thématiques, nous présentons maintenant les projets à moyen terme qui concernent simultanément plusieurs domaines abordés dans ce mémoire. Celles-ci sont motivées par le contexte thématique de la télédétection et, en particulier, par l'exploitation opérationnelle des futures images du programme de coopération *ORFEO* [1] : le système dual d'observation de la Terre à résolution métrique, dont Pléiades (France) est la composante optique et Cosmo-Skymed (Italie) est la composante radar. À titre d'exemple, deux simulations Pléiades sont proposées dans la Fig. 4.1. Des enjeux particulièrement motivants apparaissent eu égard à ces nouvelles données Très Haute Résolution (THR).

Les perspectives ouvertes par le système d'observation *ORFEO* concernent plusieurs champs de l'analyse des images et requièrent des compétences encore non exploitées en imagerie satellitaire, en particulier détection et la reconnaissance de formes. En effet, les résolutions submétriques entraînent l'apparition de nouveaux objets visibles dans les images –bâtiments, arbres, véhicules, animaux, voies de circulation...–, qu'une fois détectés il serait intéressant de reconnaître pour l'exploitation thématique des images. Mais l'amélioration de la résolution est à l'ori-

1. — Initialement franco-italien, ce programme accueille maintenant la Suède, la Belgique, l'Espagne et l'Autriche.

gine de difficultés nouvelles, puisque, d'une part, les dimensions des images explosent et, d'autre part, de nouvelles formes de bruits –liés à la surface elle-même, comme des objets mobiles, et non au capteur– apparaissent désormais. Ainsi, pour exploiter convenablement ces données, il devient primordiale de développer des méthodes de segmentation orientées objets/primitives plutôt que pixels.

Nous avons proposé des éléments de réponse dans le cadre du stage de master recherche de J.-N. Braun [17], dans un contexte très différent puisqu'il s'agissait de segmenter des images de mosaïques anciennes. Ainsi, ce ne sont pas directement les pixels qui sont pris en compte mais les tesselles de la mosaïque, par le biais de primitives les caractérisant (couleur moyenne, forme…). Il en résulte des segmentations sémantiquement plus satisfaisantes. Par exemple, les travaux réalisés par à l'ENST concernant les champs de Markov sur des graphes de primitives (graphes de régions, segments, objets) [171] est une piste intéressante que nous pourrions adapter au modèle des chaînes de Markov.

Une des exploitations possibles de ces données concerne la mise à jour automatique de cartes *IGN*[2] à partir de données THR, qu'elles soient optique et/ou radar. D'une certaine manière, il s'agit de détecter les changements entre une carte et une image, c'est à dire deux représentations très différentes de la même scène. C'est le sujet de H. Dong [74] pour lequel nous comptons exploiter les modèles markoviens entre les objets qui apparaissent, disparaissent ou subsistent dans la carte. Notamment, nous comptons développer une méthode de segmentation orientée-objet des images THR conditionnée par les objets de la carte.

Plus particulièrement, la mise à jours de cartes IGN à partir de données optique et radar THR dans la thèse de Dong Han [74]. Ainsi, apparition et disparition de nouveaux objets pourra être modélisé par probabilités conditionnelles et réaliser des segmentations des images optiques et radar contraintes par les objets de l'ancienne carte IGN pour la mettre à jour. Plusieurs stratégies sont possibles. Reprendre le sujet de thèse pour le CNES et Thèse Han Dong. MAJ de carte : Segmentation orientée objet, conditionnée par une carte à une date plus ancienne que l'image qui va servir pour la MAJ.

Ainsi, les promesses de ces nouveaux capteurs, qui rendent accessible des objets à des résolutions sub-métriques, seront une très bonne opportunité pour faire converger les deux thématiques décrites dans ce manuscrit car l'expérience acquise tant en segmentation qu'en reconnaissance de formes pourra être mise à profit pour l'exploitation de ces nouvelles données THR.

Enfin, nous envisageons une autre forme de collaboration très différente entre les deux thématiques par le biais de la définition et la caractérisation d'un système de Pearson 2D. Fondamentalement le système de Pearson est caractérisé par l'invariance aux homothéties du *skewness* et du *kurtosis*, construits par normalisation des moments statistiques d'ordre 3 et 4. C'est ainsi que J.-M. Nicolas [135] a utilisé la transformée de Mellin pour résoudre plus simplement l'équation différentielle à l'origine du système. Si on souhaite étendre le système au cas de lois bidimensionnelles, on peut rechercher un système caractérisé par l'invariance des moments bi-dimensionnels aux rotations du plan. Nous pouvons pour cela utiliser les invariants des moments complexes proposés dans la section 3.2, page 69. Les lois ainsi construites seront à symétrie de révolution, comme l'est la loi gaussienne 2D à matrice de covariance identité.

2. — *IGN* : Institut Géographique National.

# Liste des figures


1.1 Acquisition ASAR d'Envisat effectuée lors du naufrage du Prestige en 2002. *cf.* section 2.4, page 48. . . . . 6

1.2 Images ERS-2 PRI d'une région agricole anglaise, avant et pendant une inondation.*cf.* section 2.2, page 25. . . . . 8

1.3 Indexation et recherche par l'exemple. *cf.* section 3.1, page 60. . . . . 9

1.4 Poursuite d'une tasse. *cf.* section 3.3, page 74. . . . . 10

2.1 Construction du parcours de Hilbert pour une image $8 \times 8$. À chaque pixel de coordonnées $(l, c)$ est associée une abscisse $n$ indiquant le numéro d'ordre dans le parcours de la courbe, depuis la position initiale $(1, 1) \rightarrow n = 1$ jusqu'à la position finale $(1, 8) \rightarrow n = 64$. . . . . 16

2.2 Représentation de la structure de dépendance d'une CMC à bruit indépendant. . . 19

2.3 Squelette général de l'algorithme de segmentation non supervisée. . . . . 24

2.4 Résultats de segmentation d'une image RadarSat de mer glacée, en mode « ScanWide » (*cf.* texte). ©IFRTP - $512 \times 512$ (24 octobre 1996 - Mer Dumont d'Urville, large de la Terre d'Adélie). Résultats extraits de [60]. . . . . 26

2.5 À gauche : description des données vectorielles par le parcours de Hilbert (3 bandes). À droite : représentation de la structure de dépendance d'une CMC vectorielle (3 bandes). . . . . 28

2.6 (a) Acquisition ASAR d'Envisat effectuée lors du naufrage du Prestige en 2002, ©ESA. (b) Segmentation par CMC directement sur l'image. (c) Segmentation par CMC vectorielle le long d'une décomposition multi-échelle de l'image. Résultats extraits de [57]. . . . . 30

2.7 Exemples de densités bi-variées pour les trois copules avec des marges gaussiennes (1re colonne) et des marges gamma (2e colonne). Résultats extraits de [62]. . . . . 34

2.8 Une image multicomposantes de quatre bandes. Taille : $256 \times 256$. . . . . 35

2.9 Résultats de segmentation. Les mêmes couleurs ont été choisies pour les classes les plus similaires. Résultats extraits de [62]. . . . . 36

2.10 Hypothèses d'indépendance d'un modèle $\text{CMC}_{3-2}$. Les lignes en pointillés représentent les nouvelles connections obtenues en relâchant l'hypothèse **H3**. . . . . 38

2.11 Données simulées : (a) Image originale à deux classes, (b) Image bruitée et (c) son histogramme. . . . . 41

2.12 Résultats de la segmentation non supervisée de l'image de la figure 2.11 pour différentes valeurs de $R$ et $S$. $\tau$ est le taux d'erreur. $\mathcal{M}$ est un indice sur la quantité de mémoire nécessaire à l'estimation de paramètres par rapport au modèle $\text{CMC}_{1-1}$. $t$ est le temps de calcul exprimé en secondes. Résultats extraits de [11]. . . . . 42

2.13 Extrait ($512 \times 512$) d'une image ERS-RSO d'une nappe de pétrole au large des côtes Égyptiennes, ©ESA. . . . . 43

2.14 Résultats de la segmentation non supervisée de l'image de nappe de pétrole de la figure 2.13 pour différentes valeurs de $R$ et $S$. Résultats extraits de [53]. . . . . 43

2.15 Représentation de la structure de dépendance d'une chaîne de Markov couple. . . . 44

2.16 Résultats de segmentation non supervisée de l'image simulée de la figure 2.11 par le modèle CMCc. (a) Copule produit ; (b) copule gaussienne ; (c) copule logistique de Gumbel. . . . . 46

2.17 Lois bidimensionnelles estimées à l'aide de la copule logistique de Gumbel avec marges gaussiennes, aboutissant à l'image segmentée 2.16(c). . . . . 47

2.18 Segmentation non supervisée de l'image radar de nappe de pétrole présentée page 43, pour différentes copules (marges gaussiennes). . . . . 48

2.19 Objectif de la segmentation statistique floue : trouver (c) plutôt que (b) à partir de l'observation (a). Extrait de [28]. . . . . 49

2.20 (a) Image originale et (b) image floue correspondante. Le tableau (c) présente les paramètres des bruits gaussiens utilisés dans la segmentation des images floues de la figure 2.21. . . . . 52

2.21 La ligne (1) présente les images bruitées obtenues à partir des paramètres du tableau 2.20(c). La ligne (2) montre les résultats de segmentation avec l'algorithme des CMCs floues (critère MPM) pour $F = 4$ sous-intervalles flous. La ligne (3) illustre le comportement de l'algorithme, sur l'image bruitée (a), pour un nombre croissant de sous-intervalles flous. Résultats extraits de [28]. . . . . 53

2.22 Images ERS-2 PRI (taille : $512 \times 512$) d'une région agricole près de Gloucester, Angleterre, avant (a) et pendant (b) une inondation. © ESA, distribution Eurimage. . . 55

2.23 Cartes de changements obtenues avec les modèles de CMC (a) et CMC floue (b) avec une modélisation non linéaire des deux classes floues et avec $F = 2$ pour les deux mesures floues. Résultats extraits de [26]. . . . . 56

3.1 Reconstruction de l'image d'un papillon en fonction du nombre d'harmoniques utilisées (algorithme TFMA-F) . . . . . 62

3.2 Résultats de reconnaissance pour deux images requêtes (situées en haut et à gauche) à l'aide de la famille définie par les paramètres de normalisation de l'équation (3.7) et de l'approximation TFMA-F. Les images sont triées en fonction de la distance croissante à la requête, de gauche à droite et de bas en haut. Résultats extraits de [55]. 64

3.3 (a) et (b) Deux images d'une même carte électronique. (c) La courbe 3D représente **E**. Celle-ci montre un minimum pour les paramètres donnés par la différence de taille et d'orientation entre (a) et (b). Résultats extraits de [56]. . . . . 66

3.4 Une image d'étoile de mer (a) et dessin des fonctionnelles pour l'estimation des axes de symétrie par rotation (b) et réflexion (c) de l'image. Résultats extraits de [88]. 67

3.5 La ligne (a) montre l'évolution de la reconstruction de la lettre 'E' avec un nombre croissant de MC. La ligne (b) montre la version binarisée de ces images reconstruites (seuil à $0.5$). La ligne (c) montre la différence pixel à pixel entre l'image originale et l'image reconstruite. Résultats extraits de [88]. . . . . 71
3.6 Architecture générique d'un système d'indexation de bases d'images. . . . . 72
3.7 Résultats d'une recherche séquentielle sur l'image requête Obj16_69 en utilisant (a) les invariants complets des MC, (b) les invariants complets de la TFMA et (c) les invariants de Hu. Résultats extraits de [128]. . . . . 73
3.8 Résultats d'une recherche pour les invariants complets. Résultats extraits de [128]. . 74
3.9 Stratégie d'introduction des descripteurs invariants au modèle des *snakes*. . . . . 77
3.10 Résultats de détection sur la forme en « U » pour (b) une translation et un facteur d'échelle, (c) un facteur d'échelle et une rotation, prenant (a) comme *a priori* de forme. Résultats extraits de [54]. . . . . 78
3.11 Évolution dans les zones fortement concaves par le modèle « GVF » (a) et par le modèle avec *a priori* (b). Résultats extraits de [54]. . . . . 79
3.12 Résultats de détection pour différentes valeurs de $\sigma$, avec $m2-m1 = 180$, $c_1 = 0.12$, $c_2 = 0.25$ et $c_3 = 1$. Résultats extraits de [35]. . . . . 79
3.13 Application des *snakes* avec *a priori* à la segmentation de coupes scintigraphiques du myocarde. Résultats de détection (b) à partir de l'*a priori* (a). Résultats extraits de [35]. . . . . 80
3.14 Application des *snakes* avec *a priori* au suivi d'une tasse en mouvement à partir d'une caméra fixe. $c_1 = 0.1$, $c_2 = 0.25$ and $c_3 = 1$. Résultats extraits de [35]. . . . . 81
3.15 Synopsis général de l'algorithme de poursuite de cible couleur. Les lettres $h$, $w$ et $C$ réfèrent respectivement à la hauteur, la largeur et le centre de la fenêtre de localisation de la cible. La log-vraisemblance du modèle $\mathcal{L}$ est comparée à un seuil estimé dynamiquement à chaque itération. Résultats extraits de [102]. . . . . 83
3.16 Poursuite d'un visage partiellement et totalement masqué. Les rectangles qui apparaissent sur chacune des images modélisent les fenêtres de prédiction et de localisation de l'objet d'intérêt. Résultats extraits de [104]. . . . . 84
3.17 Quatre cas d'iris dégradée lors de l'acquisition. Résultats extraits de [100]. . . . . 85
3.18 Squelette de l'algorithme de sélection des meilleurs images d'iris au fur et à mesure du flux vidéo, avant la phase d'authentification. . . . . 86
3.19 Résultat de suivi d'une pupille illustrant le comportement de l'algorithme dans différentes situations (occlusions, flou, mouvement rapide, ...). L'algorithme perd temporairement la pupille, puis la retrouve systématiquement. Résultats extraits de [100]. . . . . 87
3.20 Courbes ROC montrant l'évolution du taux de fausse alarme (FAR) en fonction du taux de faux rejet (FRR) pour deux bases d'iris. EER désigne le taux d'égale erreur (FAR=FRR). Résultats extraits de [104]. . . . . 88

4.1 Simulations Pléiades (images PELICAN), panchromatique $0,70m$, Taille : $1460 \times 1460$. Programme d'accompagnement d'*ORFEO*, CNES, 2004. . . . . 93

# Sigles & Abréviations

| | |
|---|---|
| ENSPM | École Nationale Supérieure de Physique de Marseille |
| EGIM | École Généraliste d'Ingénieurs de Marseille |
| ECM | École Centrale Marseille |
| ENSI | École Nationale des Sciences de l'Informatique, Tunis |
| GET | Groupe des École des Télécommunications |
| ENST-Bretagne | École Nationale Supérieure des Télécommunications de Bretagne |
| INT | Institut National des Télécommunications |
| ITI | Dpt Image et Traitement de l'Information, ENST-BRETAGNE |
| CITI | Département Communications, Images et Traitement de l'Information de l'INT |
| GRIF | Groupe de Recherche Image et Formes |
| GRIFT | Groupe de Recherche Image et Formes de Tunisie, Laboratoire Cristal |
| GSM | Groupe Signaux Multidimensionnel, Institut Fresnel |
| PRISM | Master recherche Protocoles, Réseaux, Images et Systèmes Multimédia de l'ENSI |

| | |
|---|---|
| CMC | Chaîne de Markov Cachée |
| CMCc | Chaîne de Markov Cachée Couple |
| CMCR | Chaîne de Markov Cachée d'ordre $R$ |
| ECI | Estimation Conditionnelle Itérative |
| EM | « *Estimation Maximisation* » |
| SEM | « *Stochastic EM* » |
| MAP | Maximum *A Posteriori* |
| MPM | Maximum *a Posteriori* des Marges |
| ACP | Analyse en Composantes Principales |
| ACI | Analyse en Composantes Indépendantes |
| TFM | Transformée de Fourier-Mellin |
| TFMA | Prolongement Analytique de la TFM |
| MC | Moments Complexes |

| | |
|---|---|
| MOSS | « *Multiscale Oil Slick Segmentation* » |
| SIR | Segmentation d'Images Radar |
| DéCA | Détection de Changement Abrupts par Télédétection RSO |
| CACHANT | CArtographie des CHangements Naturels par Télédétection (GDR-ISI) |
| | |
| THR | Très Haute Résolution |
| ORFEO | « *Optical and Radar Federated Earth Observation* » |
| RSO | Radar à Synthèse d'Ouverture |
| SPOT | Satellite Pour l'Observation de la Terre |
| IR | Infra-Rouge |

# Bibliographie


[1] Y. S. Abu-Mostafa and D. Psaltis. Image normalization by complex moments. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 7(1) :46–55, 1985.

[2] A. Amini, S. Terhani, and T. Weymouth. Using dynamic programming for minimizing the energy of active contours in presence of hard constraints. In *Proc. of the 2nd Int. Conf. on Computer Vision (ICCV'96)*, pages 95–99, Tampa, Florida, 1988.

[3] K. E. Avrachenkov and E. Sanchez. Fuzzy Markov chains. *Fuzzy Optimization and Decision Making*, 1(2) :143–159, June 2002.

[4] O. Aycard, J.-F. Mari, and R. Washington. Learning to automatically detect features for mobile robots using second-order HMMs. In *IJCAI - Proc. of the Workshop on Reasoning with Uncertainty in Robotics*, Acapulco, Mexico, July 2003.

[5] L. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 41 :164–171, 1970.

[6] Y. Bazi, L. Bruzzone, and F. Melgani. An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. *IEEE Trans. Geosci. Remote Sens.*, 43(4) :874–887, April 2005.

[7] D. Benboudjema and W. Pieczynski. Unsupervised image segmentation using triplet Markov fields. *Computer Vision and Image Understanding*, 99(3) :476–498, 2005.

[8] B. Benmiloud and W. Pieczynski. Estimation des paramètres dans les chaînes de Markov cachées et segmentation d'images. *Traitement du Signal*, 12(5) :433–454, 1995.

[9] Btissam Benmiloud. *Chaînes de Markov cachées et segmentation statistique non supervisée de séquences d'images*. PhD thesis, Université Paris VII, Institut National des Télécoms, CITI, Evry, France, December 1994.

[10] Farabi Bensalah. *Estimation du mouvement par vision active*. PhD thesis, Université de Rennes, July 1996.

[11] L. Benyoussef, C. Carincotte, and S. Derrode. Extension of higher-order HMC modeling with application to image segmentation. *Digital Signal Processing (Elsevier), accepté le 26/10/2007, doi : 10.1016/j.dsp.2007.10.010*, 2007.

[12] A. Berchtold. The double chain Markov model. *Commun. Statist. Theory Methods*, 28 :2569–2589, 1999.

[13] J. C. Bezdek. *Pattern recognition with fuzzy objective function algorithms*. Kluwer Academic Publishers, 1981.

[14] Z. Bouyahia, L. Benyoussef, and S. Derrode. Unsupervised SAR images change detection with hidden Markov chains on a sliding window. In *Proc. of the SPIE-RS'07 - Image and Signal Processing for Remote Sensing*, Florence (Italy), 17 - 21 September 2007.

[15] Zied Bouyahia. Détection de changements par chaînes de Markov localement stationnaire. Rapport de Mastère Recherche (10 mois), soutenu le 29 mai 2007, Lab. Cristal, Équipe GRIFT, ENSI, Campus Universitaire de la Manouba, Tunisie, 2007.

[16] K.W. Bowyer, K. Hollingsworth, and P.J. Flynn. Image understanding for iris biometrics : A survey. Technical report, Department of Computer Science and Engineering, U. Notre Dame, 2007.

[17] Jean-Noël Braun. Segmentation d'images de mosaïques par une approche orientée cellules. Rapport de Mastère Recherche (4 mois), soutenu le 19 juin 2007, Université Paul Cézanne d'Aix-Marseille III, Institut Fresnel (UMR 6133), Équipe GSM, Marseille, 2007.

[18] C. Brekke and A. H. S. Solberg. Oil spill detection by satellite remote sensing. *Remote Sensing of Environment*, 95(1) :1–13, March 2005.

[19] N. Brunel and W. Pieczynski. Unsupervised signal restoration using hidden Markov chains with copulas. *Signal Processing*, 85(2) :2304–2315, 2005.

[20] N. Brunel, W. Pieczynski, and S. Derrode. Copulas in vectorial hidden Markov chains for multicomponent image segmentation. In *Proc. of the IEEE Int. Conf. ICASSP'05*, Philadelphia (USA), 19-23 March 2005.

[21] Nicolas Brunel. *Sur quelques extensions des chaînes de Markov cachées et couples. Applications à la segmentation non-supervisée de signaux radar*. PhD thesis, Université Pierre et Marie Curie, Paris VI, Institut National des Télécoms, CITI, Evry, France, December 2005.

[22] H. Caillol, W. Pieczynski, and A. Hillion. Estimation of fuzzy Gaussian mixture and unsupervised statistical image segmentation. *IEEE Trans. Image Process.*, 6(3) :425–440, March 1997.

[23] O. Cappé, E. Moulines, and T. Rydén. *Inference in hidden Markov models*, volume 47 of *Springer Series in Statistics*. Springer-Verlag, 2005.

[24] J.-F. Cardoso. Blind signal separation : statistical principles. *Proc. IEEE*, 9(10) :2009–2025, October 1998.

[25] C. Carincotte, S. Derrode, and S. Bourennane. Multivariate fuzzy hidden Markov chains model applied to unsupervised multiscale SAR image segmentation. In *Proc. of the IEEE Int. Conf. FUZZ'05*, Reno (Nevada, USA), 22-25 May 2005.

[26] C. Carincotte, S. Derrode, and S. Bourennane. Unsupervised change detection on SAR images using fuzzy hidden Markov chains. *IEEE trans. on Geoscience and Remote Sensing*, 44(2) :432–441, 2006.

[27] C. Carincotte, S. Derrode, G. Sicot, and J.-M. Boucher. Unsupervised image segmentation based on a new fuzzy HMC model. In *Proc. of the IEEE Int. Conf. ICASSP'04*, Montreal (Canada), 17-21 May 2004.

[28] Cyril Carincotte. *Segmentation non supervisée d'images par chaîne de Markov floue - Application à la détection de changements multi-dates*. PhD thesis, Université Paul Cézanne d'Aix-Marseille III, Institut Fresnel (UMR 6133), Équipe GSM, Marseille, Bourse Ministérielle, soutenue le 14 novembre 2005. Co-encadrée avec S. Bourennane. Jury : J.-M. Chassery (P), J.-P. Haton (R), W. Pieczynski (R), C. Amra, S. Bourennane et S. Derrode.

[29] R. Casselles and R. Kimmel. Geodesic active contours. In *Proc. of the 1st Int. Conf. on Computer Vision (ICCV'95)*, pages 694–699, Cambridge, Massachusetts, 1995.

[30] G. Celeux and J. Diebolt. The SEM algorithm : a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Computational Statistics Quaterly*, 2 :73–82, 1985.

[31] F. Chaker, M.T. Bannour, and F. Ghorbel. A complete and stable set of affine-invariant Fourier descriptors. In *Proc. of the 12th International Conference on Image Analysis and Processing (ICIP'03)*, pages 578–581, Barcelona, Spain, 1988.

[32] T. J. Cham and R. Cipolla. Stereo coupled active contours. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR'97)*, pages 1094–1097, San Juan, Puerto-Rico, 1997.

[33] Jocelyn Chanussot. *Approches vectorielles ou marginales pour le traitement d'images multi-composantes*. PhD thesis, Université de Savoie, November 1998.

[34] M. A. Chermi, S. Derrode, and F. Ghorbel. Euclidean invariant snake for joint stereo segmentation. In *Proc. of the IEEE Int. Conf. ICTTA'06*, Damascus (Syria), 24-28 April 2006.

[35] M.-A. Chermi, S. Derrode, and F. Ghorbel. Fourier-based geometric shape prior for snakes. *Pattern Recognition Letters*, 29(7) :897–904, May 2008.

[36] Mohamed-Ali Chermi. *Snakes* avec *a priori* de forme. Applications à la vision stéréo et au suivi d'objets en mouvement. Rapport de Mastère Recherche (10 mois), Lab. Cristal, Équipe GRIFT, ENSI, Campus Universitaire de la Manouba, Tunisie, 2006.

[37] C. Chesnaud, P. Réfrégier, and V. Boulet. Statistical region snake-based segmentation adapted to different physical noise models. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 21(11) :1145–1157, November 1999.

[38] L. D. Cohen. On active contour models and balloons. *Graphical Models and Image Processing*, 53(2) :211–218, March 1991.

[39] L. D. Cohen and I. Cohen. Finite-elements methods for active contour models and balloons for 2-D and 3-D images. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 15(11) :1131–1147, November 1996.

[40] P. Common. Independant component analysis, a new concept ? *IEEE Trans. Signal Process.*, 36 :287–314, 1994.

[41] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathornas, and P. N. Yianilos. The Bayesian image retrieval system, PicHunter : Theory, implementation, and psychophysical experiments. *IEEE trans. on Image Processing*, 9(1) :20–37, January 2000.

[42] D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr. Diffusion snakes : Introducing statistical shape knowledge into the Mumford-Shah functional. *Int. J. of Computer Vision*, 50(3) :295–313, December 2002.

[43] T. R. Crimmins. A complete set of Fourier descriptors for two-dimensional shapes. *IEEE trans. on Systems, Man, and Cybernetics*, 12 :848–855, 1982.

[44] J. Daugman. The importance of being random : statistical principles of iris recognition. *Pattern Recognition*, 36(2) :279–291, 2003.

[45] P. J. Davis. Plane regions determined by complex moments. *J. of Approximation Theory*, 19 :148–153, 1977.

[46] R. de Luis-Garcia, C. Alberola-Lopez, O. Aghzout, and J. Ruiz-Alzola. Biometric identification systems. *Signal Processing*, 83(12) :2539–2557, December 2003.

[47] E. de Villiers and J. du Preez. The advantage of using higher order HMM for segmenting acoustic files. In *Proc. of the Annual Symp. of the Pattern Recognition Association of South Africa*, Franschhoek, South Africa, November 2001.

[48] Y. Delignon, A. Marzouki, and W. Pieczynski. Estimation of generalized mixtures and its application in image segmentation. *IEEE trans. on Image Processing*, 6(10) :1364–1375, October 1997.

[49] Y. Delignon and W. Pieczynski. Modeling non-Rayleigh speckle distribution in SAR images. *IEEE trans. on Geoscience and Remote Sensing*, 40(6) :1430–1435, 2002.

[50] J.-P. Delmas. An equivalence of the EM and ICE algorithms for the exponential family. *IEEE trans. on Signal Processing*, 45(10) :2613–2615, 1997.

[51] A. P. Dempster, N. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. of the Royal Statsitical Society, series B*, 39 :1–38, 1977.

[52] A. P. Demptser, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. of the Royal Statistical Society*, B(39) :1–38, 1977.

[53] S. Derrode, C. Carincotte, and S. Bourennane. Unsupervised image segmentation based on high-order hidden Markov chains. In *Proc. of the IEEE Int. Conf. ICASSP'04*, Montreal (Canada), 17-21 May 2004.

[54] S. Derrode, M.A. Chermi, and F. Ghorbel. Fourier-based invariant shape prior for snakes. In *Proc. of the IEEE Int. Conf. ICASSP'06*, Toulouse (France), 14-19 May 2006.

[55] S. Derrode and F. Ghorbel. Robust and efficient Fourier-Mellin transform approximations for invariant grey-level image description and reconstruction. *Computer Vision and Image Understanding*, 83(1) :57–78, 2001.

[56] S. Derrode and F. Ghorbel. Shape analysis and symmetry detection in gray-level objects using the analytical Fourier-Mellin representation. *Signal Processing*, 84(1) :25–39, 2004.

[57] S. Derrode and G. Mercier. Multiscale oil slick segmentation from SAR images using a vector HMC model. *Pattern Recognition*, 40(3) :1135–1147, March 2007.

[58] S. Derrode, G. Mercier, and C. Carincotte. Développement d'un algorithme de détection automatique de changements entre images SAR. Rapport de recherche, financé par le CNES, Centre spatial de Toulouse (32k€), ENST Bretagne & Institut Fresnel (UMR 6133), 2003.

[59] S. Derrode, G. Mercier, and C. Carincotte. Développement d'un algorithme de détection automatique de changements entre images SAR. Projet financé par le CNES, Centre spatial de Toulouse, ENST Bretagne & Institut Fresnel (UMR 6133), 2003.

[60] S. Derrode, G. Mercier, J.-M. Le Caillec, and R. Garello. Estimation of sea-ice SAR clutter statistics from Pearson's system of distributions. In *Proc. of the IEEE IGARSS'01*, Sydney (Australia), 9-13 July 2001.

[61] S. Derrode, G. Mercier, and W. Piecznski. Unsupervised change detection in SAR images using a multicomponent HMC model. In *Proc. of the 2nd Int. MultiTemp Workshop*, Ispra (Italy), 16-18 July 2003.

[62] S. Derrode, G. Mercier, and W. Piecznski. Unsupervised multicomponent image segmentation combining a vectorial HMC model and ICA. In *Proc. of the IEEE Int. Conf. on Image Processing*, Barcelona (Spain), 14-17 September 2003.

[63] S. Derrode, R. Mezhoud, and F. Ghorbel. Comparaison de deux familles complètes de descripteurs de formes pour l'indexation de bases d'objets 2D à niveaux de gris. *Annales des Télécommunications*, 55(3/4) :184–193, 2000.

[64] S. Derrode and W. Piecznski. Segmentation non supervisée d'images par chaîne de Markov couple. In *Traitement et Analyse de l'Information - Méthodes et Applications*, Hammamet (Tunisia), 26 septembre - 3 octobre 2003.

[65] S. Derrode and W. Pieczynski. SAR image segmentation using generalized pairwise Markov chains. In *SPIEs Int. Symp. on Remote Sensing (SPIE-RS'02)*, Crete (Greece), 22-27 September 2002.

[66] S. Derrode and W. Pieczynski. Unsupervised signal and image segmentation using pairwise Markov chains. *IEEE trans. on Signal Processing*, 52(9) :2477–2489, 2004.

[67] S. Derrode, E. Trouvé, G. Mercier, J. Inglada, J. Chanussot, A. Joannic-Chardin, and M. Gay. CArtographie des CHAngements Naturels par Télédétection (CACHANT). Projet d'études (9k€), Resp. : S. Derrode & E. Trouvé (LISTIC), GDR ISIS & CNES, 2006.

[68] S. Derrode, E. Trouvé, G. Mercier, J. Inglada, J. Chanussot, A. Joannic-Chardin, and M. Gay. CArtographie des CHAngements Naturels par Télédétection (CACHANT). Projet d'études, GDR ISIS & CNES, 2006.

[69] Stéphane Derrode. *Représentation de formes planes à niveaux de gris par différentes approximations de Fourier- Mellin analytique en vue d'indexation de bases d'images*. PhD thesis, Université de Rennes I, GET/ENST BRETAGNE, Département *ITI*, Soutenue le 15 décembre 1999. Encadrée avec F. Ghorbel (ENSI/GRIFT, Tunisie).

[70] P.A. Devijver. Baum's Forward-Backward algorithm revisited. *Pattern Recognition Letters*, 3 :369–373, 1985.

[71] J. Dieudonné. *Éléments d'analyse*. Tome II. Gauthier-Villars, Cahiers scientifiques, Paris, France, 3$^{e}$ edition, 1983.

[72] K. Djemal, C. Chettaoui, and H. Maaref. Shapes description for cells sickle illness recognition. In *IEEE Int. Conf. on Signals Systems & Devices (SSD 2005)*, Sousse, Tunisia, 21-24 May 2005.

[73] M. Do and M. Vetterli. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. Image Process.*, 11(2) :146–158, February 2002.

[74] Han Dong. *Map updating using very high resolution optical and radar satellite images*. PhD thesis, Université Paul Cézanne d'Aix-Marseille III / École Centrale de Marseille, Institut Fresnel (UMR 6133), Équipe GSM, Marseille, Bourse financée par le gouvernement chinois, soutenance prévue en octobre 2010. Co-encadrée avec S. Bourennane.

[75] Jean-Batiste Durand. *Modèles à structure cachée : inférence, estimation, sélection de modèles et applications*. PhD thesis, Université Joseph Fourier, Grenoble I, January 2003.

[76] Y. Ephraim and N. Merhav. Hidden Markov processes. *IEEE trans. on Information Theory*, 48(6) :1518–1569, June 2002.

[77] Z. Ghahramani et M. I. Jordan. Factorial hidden Markov models. *Machine Learning*, 29 :245–273, 1997.

[78] L. Flom and A. Safir. Iris recognition system. U.s. patent 4,641,349, 1987.

[79] J. Flusser and T. Suk. Construction of complete and independent systems of rotation moment invariants. In N. Petkov and M. A. Westenberg, editors, *Proc. of the Computer Analysis of Images and Patterns Conf. (CAIP'03)*, number 2756 in Lecture Notes in Computer Science, pages 41–48. Springer, Berlin, 2003.

[80] A. Foulonneau, P. Charbonnier, and F. Heitz. Affine-invariant geometric shape priors for region-based active contours. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 28(8) :1352–1357, August 2006.

[81] M. Gade, W. Alpers, H. Hühnerfuss, V. Wismann, and P. Lange. On the reduction of the radar backscatter by oceanic surface films : scattometer measurements and their theoretical interpretation. *Remote Sensing of Environment*, 66 :52–70, 1998.

[82] M. Gastaud, M. Barlaud, and G. Aubert. Combining shape prior and statistical features for active contour segmentation. *IEEE trans. on Circuits and Systems for Video Technology*, 14 :726–734, May 2004.

[83] M. Gelautz and D. Markovic. Recognition of object contours from stereo images : an edge combination approach. In *Proc. of the Int. Symp. on 3D Data Processing, Visualization, and Transmission (3DPVT'04)*, pages 774–780, Thessaloniki, Greece, 2004.

[84] D. Geman and S. Geman. Stochastic relaxation, gibbs distribution and the bayesian restoration of images. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 6 :721–741, 1984.

[85] F. Ghorbel. Stability of invariant Fourier descriptors and its inference in the shape classification. In *Proc. of the Int. Conf. in Pattern Recognition (ICPR'92)*, The Hague, Netherlands, 30 August - 3 September 1992.

[86] F. Ghorbel. A complete invariant description for gray-level images by the harmonic analysis approach. *Pattern Recognition Letters*, 15 :1043–1051, October 1994.

[87] F. Ghorbel. Towards a unitary formulation for invariant image description ; Application to image coding. *Annales des Télécommunications*, 53(3) :143–163, May 1998.

[88] F. Ghorbel, S. Derrode, R. Mezhoud, T. Bannour, and S. Dhahbi. Image reconstruction from a complete set of similarity invariants extracted from complex moments. *Pattern Recognition Letters*, 27 :1361–1369, September 2006.

[89] N. Giordana and W. Pieczynski. Estimation of generalized multisensor hidden Markov chain and unsupervised image segmentation. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 19(5) :465–475, 1997.

[90] F. Girard-Ardhuin, G. Mercier, F. Collard, and R. Garello. Operational oil-slick characterization by SAR imagery and synergistic data. *IEEE Journal of Oceanic Engineering*, 30(3) :487–495, July 2005.

[91] J.-P. Haton, C. Cerisara, D. Fohr, Y. Laprie, and K. Smaïli. *Reconnaissance automatique de la parole. Du signal à son interprétation*. UniverSciences. Dunod, 2006.

[92] J. P. Havlicek, J. W. Havlicek, and A. C. Bovic. The analytic image. In *Proc. of the IEEE Int. Conf. on Image Processing (ICIP'97))*, pages 446–449, Santa Barbara, CA, USA, 26-29 October 1997.

[93] F. M. Henderson and A. J. Lewis. *Manual of remote sensing. Principles & applications of imaging radar*. John Wiley & Sons, Inc., 3$^{e}$ edition, 1998.

[94] A. Hyvärinen and E. Oja. ICA : algorithms and applications. *Neural Networks*, 13 :411–430, 2000.

[95] J. Inglada and G. Mercier. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE trans. on Geoscience and Remote Sensing*, 19(5) :465–475, 2007.

[96] F. Jelinek. *Statistical Methods for Speech Recognition*. Language, Speech and Communication. MIT Press, Cambridge, MA, USA, 1997.

[97] H. Joe. *Multivariate models and dependence concepts*, volume 73 of *Monographs on Stat. and App. Prob.* Chapman and Hall, 1997.

[98] S. Bengio K. Weber, S. Ikbal and H. Bourlard. Robust speech recognition and feature extraction using HMM2. *Computer Speech & Language*, 17(2-3) :195–211, 2003.

[99] M. Kass, A. P. Witkin, and D. Terzopoulos. Snakes : Active contour models. *Int. J. of Computer Vision*, 1(4) :321–331, January 1988.

[100] W. Ketchantang, S. Derrode, S. Bourennane, and L. Martin. Video pupil tracking for iris based identification. In *Proc. of the Advanced Concepts for Intelligent Vision Systems (ACIVS'05)*, Antwerp (Belgium), 20-23 September 2005.

[101] W. Ketchantang, S. Derrode, L. Martin, and S. Bourennane. Nouveau descripteur local de qualité des images d'iris dans les séquences vidéos. In *21ème colloque GRETSI'07*, Troyes (France), 11 - 14 Septembre 2007.

[102] W. Ketchantang, S. Derrode, L. Martin, and S. Bourennane. Pearson-based mixture model for color object tracking. *Machine Vision and Applications (Hindawi), accepté le 06/12/2007, doi : 10.1007/s00138-008-0124-4*, 2008.

[103] W. Ketchantang, L. Martin, and S. Derrode. Nouvel algorithme de sélection en temps réel des images d'iris exploitables pour l'authentification des personnes. Brevet, Num. d'enreg. nat. INPI : 06/08284., Société ST MicroElectronics Rousset SAS & Université Paul Cézanne d'Aix-Marseille III, 2006.

[104] William Ketchantang. *Étude d'une nouvelle modalité vidéo pour le reconnaissance biométrique de l'iris*. PhD thesis, Université Paul Cézanne d'Aix-Marseille III, Bourse Cifre, soutenue le 29 janvier 2008. Co-encadrée avec S. Bourennane.

[105] S. H. Kim, J. H. Choi, H. B. Kim, and J. W. Jung. A new snake algorithm for object segmentation in stereo images. In *Proc. of the IEEE Int. Conf. on Multimedia and Expo (ICME 2004)*, Taipei, Taiwan, 27-30 June 2004.

[106] S. Kotz, N. Balakrishnan, and N. L. Johnson. *Continuous multivariate distributions*, volume 1, Models and Applications of *Wiley Series in Probability and Statistics, Applied probability and statistics section*. Wiley Interscience, John Wiley & Sons, Inc, 2$^{e}$ edition, June 2000.

[107] J. Lapuyade-Lahorgue and W. Pieczynski. Partially Markov models and unsupervised segmentation of semi-Markov chains hidden with long dependence noise. In *Proc. of the Int. Symp. on Applied Stochastic Models and Data Analysis (ASMDA'07)*, pages 648–655, Crete, Greece, 2007.

[108] F. Lecellier, S. Jehan-Besson, M. J. Jadili, G. Aubert, and M. Revenu. Statistical region-based active contours with exponential family of observations. In *IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP'06)*, Toulouse, France, 15-19 May 2006.

[109] S. E. Levinson. Continuously variable duration hidden Markov models for automatic speech recognition. *Comput. Speech Lang.*, 1(1) :29–45, 1986.

[110] S. Z. Li. *Markov random field modeling in computer vision*. Springer-Verlag, Tokyo, Japan, 2e edition, 2001.

[111] R. S. Lipster and A. N. Shiryaev. *Statistics of random processes*, chapter 13 – Conditionally Gaussian sequences : filtering and related problems, page 427. Springer-Verlag, New York (USA), 2e edition, 2001.

[112] R. MacKay Altmann. Assessing the goodness-of-fit of hidden Markov models. *Biometrics*, 60 :444–450, June 2004.

[113] J.-F. Mari, J.-P. Haton, and A. Kriouille. Automatic word recognition based on second-order hidden Markov models. *IEEE Trans. Speech Audio Process.*, 5(1) :22–25, January 1997.

[114] L. Martin, G. Petitjean, S. Derrode, and W. Ketchantang. Procédé et dispositif de localisation d'un iris humain dans une image. Brevet, Num. d'enreg. nat. INPI : 07/03039., Société ST MicroElectronics Rousset SAS & Université Paul Cézanne d'Aix-Marseille III, 2007.

[115] S. J. McKenna, S. Gong, and Y. Raja. Modelling facial colour and identity with Gaussian mixtures. *Pattern Recognition*, 31(12) :1883–1892, 1998.

[116] S. J. McKenna, Y. Raja, and S. Gong. Tracking colour objects using adaptive mixture models. *Image and Vision Computing*, 17(3/4) :225–231, March 1999.

[117] G. J. McLachlan and T. Khrishnan. *The EM algorithm and extensions*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York, USA, 1996.

[118] G. Mercier. Mesure de dépendance entre images RSO. Technical Report RR-2005003-ITI, GET/ENST Bretagne, October 2005.

[119] G. Mercier, S. Derrode, and M. Lennon. Hyperspectral image segmentation with Markov chain model. In *Proc. of the IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS'06)*, Toulouse, France, 2003.

[120] G. Mercier, S. Derrode, W. Piecynski, R. Fjørtoft, J.-M. Nicolas, R. Garello, and F. Tupin. Segmentation statistique d'images radar à l'aide modèles markoviens (SIR). Rapport de recherche, financé par le Groupement des écoles des Télécoms (GET-RE315), ENST Bretagne & Institut Fresnel (UMR 6133) & Inst. Nat. des Télécoms & ENST, 2002.

[121] G. Mercier, S. Derrode, W. Piecynski, R. Fjørtoft, J.-M. Nicolas, R. Garello, and F. Tupin. Segmentation statistique d'images radar à l'aide modèles markoviens (SIR). Projet financé par le Groupement des Écoles des Télécoms (GET-RE109), ENST Bretagne & Institut Fresnel (UMR 6133) & Inst. Nat. des Télécoms & ENST, 2001.

[122] G. Mercier, S. Derrode, W. Piecynski, R. Garello, and J.-M. Le Caillec. Multiresolution Oil Slick Segmentation (MOSS). Rapport de recherche, financé par le Groupement des écoles des Télécoms (GET-RE315), ENST Bretagne & Institut Fresnel (UMR 6133) & Inst. Nat. des Télécoms, 2002.

[123] G. Mercier, S. Derrode, W. Piecynski, R. Garello, and J.-M. Le Caillec. Multiresolution Oil Slick Segmentation (MOSS). Projet financé par le Groupement des Écoles des Télécoms (GET-RE315), ENST Bretagne & Institut Fresnel (UMR 6133) & Inst. Nat. des Télécoms, 2002.

[124] G. Mercier, S. Derrode, W. Piecynski, J.-M. Nicolas, and R. Garello. Détection de Changements Abrupts par télédétection RSO (DéCA). Rapport de recherche, financé par le Groupement des écoles des Télécoms (GET-RI022), ENST Bretagne & Institut Fresnel (UMR 6133) & Inst. Nat. des Télécoms & ENST, 2005.

[125] G. Mercier, S. Derrode, W. Piecynski, J.-M. Nicolas, and R. Garello. Détection de Changements Abrupts par télédétection RSO (DéCA). Projet financé par le Groupement des Écoles des Télécoms (GET-RI022), ENST Bretagne & Institut Fresnel (UMR 6133) & Inst. Nat. des Télécoms & ENST, 2005.

[126] G. Mercier, S. Derrode, and W. Pieczynski. Segmentation multi-échelle de nappes d'hydrocarbure. *Traitement du Signal*, 21(4) :329–346, 2004.

[127] G. Mercier, S. Derrode, W. Pieczynski, J.-M. Nicolas, A. Joannic-Chardin, and J. Inglada. Copula-based stochastic kernels for abrupt change detection. In *Proc. of the IEEE Int. Conf. IGARSS'06*, Denvers (Colorado, USA), July 31-August 4 2006.

[128] Rim Mezhoud-Khouaja. *Indexation de bases d'images planes à l'aide d'invariants complets d'images à niveaux de gris*. PhD thesis, Université de Tunis II, Tunisie, ENSI, Lab. Cristal, Équipe GRIFT, Campus Universitaire de la Manouba, Manouba, Tunisie, Soutenue le 2 novembre 2005. Co-encadrée avec F. Ghorbel (ENSI/GRIFT, Tunisie). Jury : F. Kammoun (P), A. Ben Azza (R), L. Chen (R), N. Ellouze, F. Ghorbel et S. Derrode.

[129] H. Ming-Kuel. Visual pattern recognition by moment invariants. *IRE trans. on Information Theory*, 8 :179–187, 1962.

[130] T. P. Minka and R. W. Picard. Interactive learning using a society of models. *Pattern Recognition*, 30(4) :565–581, 1997.

[131] M. A. Mohamed and P. Gader. Generalized hidden Markov models-Part I : Theorical frameworks. *IEEE Trans. Fuzzy Syst.*, 8(1) :67–81, February 2000.

[132] G. Moser, S. B. Serpico, and G. Vernazza. Unsupervised change detection from multichannel SAR images. *accepted for publication in the IEEE Geoscience and Remote Sensing Letters*, 2007.

[133] R. B. Nelsen. *An introduction to copulas*, volume 139 of *Lectures Notes in Statistics*. Springer-Verlag, 1998.

[134] S.A. Nene, S.K. Nayar, and H. Murase. Columbia Object Image Library (COIL-20). Technical Report CUCS-005-96, Columbia University, February 1996.

[135] J.-M. Nicolas. Étude de lois de probabilités définies sur R+ : apport des fonctions caractéristiques et des cumulants de deuxième espèce dans l'étude de mélanges multiplicatifs et additifs. In *18ème colloque GRETSI'01*, Toulouse, September 2001.

[136] S. Osher and R.P. Fedkiw. Level set methods : An overview and some recent results. *Journal of Computational Physics*, 169 :463–502, 2001.

[137] E. Ozyildiz, N. Krahnstoever, and R. Sharma. Adaptive texture and color segmentation for tracking moving objects. *Pattern Recognition*, 35(10) :2013–2029, October 2002.

[138] N. Paragios and M. Rousson. Shape priors for level set representations. In *Proc. of the European Conf. on Computer Vision (ECCV'02)*, pages 78–92, Copenhagen, Denmark, 27 May - 2 June 2002.

[139] E. Pardoux. *Processus de Markov et applications - Algorithmes, réseaux, génome et finance*. Science Sup. Dunod, 2007.

[140] N. S. Peng, J. Yang, and Z. Liu. Mean shift blob tracking with kernel histogram filtering and hypothesis testing. *Pattern Recognition Letters*, 26 :605–614, August 2005.

[141] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook : Tools for content-based manipulation of image databases. *Int. J. of Computer Vision*, 18(3) :233–254, 1996.

[142] W. Pieczynski. Statistical image segmentation. *Machine Graphics and Vision*, 1(1/2) :261–268, 1992.

[143] W. Pieczynski. Champs de Markov cachées et estimation conditionnelle itérative. *Traitement du Signal*, 11(2) :141–153, 1994.

[144] W. Pieczynski. Arbres de Markov couple ; pairwise Markov trees. *Comptes Rendus de l'Académie des Sciences - Mathématique, Série I*, 335 :79–82, 2002. *in French*.

[145] W. Pieczynski. Chaînes de Markov triplet ; Triplet markov chains. *Comptes Rendus de l'Académie des Sciences - Mathématique, Série I*, 335(3) :275–278, 2002.

[146] W. Pieczynski. Pairwise Markov chains. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 25(5) :634–639, 2003.

[147] W. Pieczynski. Modèles de Markov en traitement d'images. *Traitement du Signal*, 20(3) :255–278, 2004.

[148] W. Pieczynski. Multisensor triplet Markov chains and theory of evidence. *Int. J. of Approximate Reasoning*, 45(1) :1–16, 2007.

[149] W. Pieczynski, J. Bouvrais, and C. Michel. Estimation of generalized mixture in the case of correlated sensors. *IEEE Trans. Image Process.*, 9(2) :308–311, February 2000.

[150] W. Pieczynski, C. Hulard, and T. Veit. Triplet Markov chains in hidden signal restoration. In *SPIE International Symposium on Remote Sensing (SPIE-RS)*, Crete, Greece, 22-27 septembre 2002.

[151] W. Pieczynski and A.-N. Tebbache. Pairwise Markov random fields and segmentation of textured images. *Machine Graphics & Vision*, 9(3) :705–718, 2000.

[152] W. H. Press, S. A. Teukolsky, A. T. Vetterling, and B. P. Flannery. *Numerical recipes in C, the art of scientific computing*. Cambridge University Press, 2$^{e}$ edition, 1983.

[153] R. J. Prokop and A. P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *Graphical Models and Image Processing*, 54(5) :438–460, 1992.

[154] Jean-Noël Provost. *Classification bathymétrique en imagerie multispectrale SPOT*. PhD thesis, Université de Bretagne Occidentale, Brest, June 2001.

[155] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of the IEEE*, 77(2) :257–286, February 1989.

[156] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysman. Image change detection algorithms : a systematic survey. *IEEE Trans. Image Process.*, 14(3) :294–307, March 2005.

[157] M. Rochery, I. H. Jermyn, and J. Zerubia. Higher order active contours. *Int. J. of Computer Vision*, 69(1) :27–42, August 2006.

[158] T. Roncalli. Gestion des risques multiples. Technical report, Groupe de Rechercche Opérationnelle, Crédit Lyonnais, 2002.

[159] A. Ross and A. K. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13) :2115–2125, September 2003.

[160] S. Roy, E.-C. Chang, and K. Natarajan. A unified framework for resolving ambiguity in copy detection. In *Proc. of the ACM Int. Conf. on Multimedia (MM'05)*, pages 648–655, Singapore, 2005.

[161] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra. Relevance feedback : a power tool in interactive content-based image retrieval. *IEEE trans. on Circuits and Systems for Video Technology, special issue on Interactive Multimedia Systems for the Internet*, 8(5) :644–655, September 1998.

[162] H. Sagan. *Space filling curves*. Springer-Verlag, New York, USA, 1994.

[163] F. Salzenstein, C. Collet, S. Lecam, and M. Hatt. Non stationnary fuzzy Markov chain. *Pattern Recognition Letters*, 28(16) :2201–2208, December 2007.

[164] F. Salzenstein and W. Pieczynski. Parameter estimation in hidden fuzzy Markov random fields and image segmentation. *CVGIP : Graphical Models and Image Processing*, 59(4) :205–220, July 1997.

[165] G. Simons, Y. C. Yao, and G. Morton. Global Markov models for eukaryote nucleotide data. *J. Statist. Plann. Inference*, 130 :251–275, 2005.

[166] W. Skarbek. *Generalized Hilbert scan in image printing*, chapter Theoretical Foundations of Computer Vision. Akademie Verlag, Berlin, Germany, R. Klette and W. G. Kropetsh edition, 1992.

[167] L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 14(11) :1061–1075, November 1992.

[168] M. Störring, T. Kocka, H. J. Andersen, and E. Granum. Tracking regions of human skin through illumination changes. *Pattern Recognition Letters*, 24(11) :1715–1723, July 2003.

[169] M. R. Teague. Image analysis via the general theory of moments. *J. of the Optical Society of America*, 70(8) :920–930, 1980.

[170] Christel-Loïc Tisse. *Contributions à la vérification biométrique de personnes par reconnaissance de l'iris*. PhD thesis, Université de Montpellier II, ST Microelectronics, Rousset, France, October 2003.

[171] Florence Tupin. *Champs de Markov sur graphes pour le traitement des images radar*. Habilitation à Diriger des Recherches, Université de Rennes, March 2007.

[172] G. Winkler. *Image analysis, random fields and Markov chain Monte-Carlo methods - A mathematical introduction*, volume 27 of *Springer Series in Stochastic Modelling and Applied Probability*. Springer-Verlag, 2$^{e}$ edition, 2003.

[173] G. Xiong, C. Feng, and L. Ji. Dynamical Gaussian mixture model for tracking elliptical living objects. *Pattern Recognition Letters*, 27(7) :838–842, May 2006.

[174] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE trans. on Image Processing*, 7(3) :359–369, March 1998.

[175] P.-T. Yap, R. Paramesran, and O. Seng-Huat. Image analysis by Krawtchouk moments. *IEEE trans. on Image Processing*, 12(11) :1367–1377, November 2003.