



HAL
open science

Sur quelques problèmes mathématiques en analyse d'images et vision stéréoscopique

Andrés Almansa

► **To cite this version:**

Andrés Almansa. Sur quelques problèmes mathématiques en analyse d'images et vision stéréoscopique. Traitement du signal et de l'image [eess.SP]. Université René Descartes - Paris V, 2005. tel-00011765

HAL Id: tel-00011765

<https://theses.hal.science/tel-00011765>

Submitted on 6 Mar 2006

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

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

UNIVERSITÉ PARIS V RENÉ DESCARTES U.F.R. MATHÉMATIQUES ET INFORMATIQUE

RAPPORT SCIENTIFIQUE

présenté par
Andrés ALMANSA

pour obtenir le diplôme d'
HABILITATION À DIRIGER DES RECHERCHES
(arrêté du 23 novembre 1988)

Sur quelques problèmes mathématiques en analyse d'images et vision stéréoscopique

Soutenue le 1 Décembre 2005 devant le Jury composé de

Gilles AUBERT	Rapporteur
Vicent CASELLES	Examineur
Antonin CHAMBOLLE	Rapporteur
Yves MEYER	Président
Lionel MOISAN	Examineur
Jean-Michel MOREL	Examineur
Françoise PRÊTEUX	Rapporteur
Bernard ROUGÉ	Examineur

A Sofia

Table des matières

Table des matières	5
1 Analyse et reconnaissance d'empreintes digitales	9
1.1 Augmentation locale de contraste de la zone d'intérêt à bruit de fond contrôlé .	12
1.2 Débruitage directionnel adaptatif et détection de points caractéristiques	14
1.3 Mise en correspondance élastique	16
2 Théorie de Shannon	18
2.1 Modèles de formation et acquisition d'images	18
2.1.1 Le modèle de "caméra à trou d'épingle" (pin hole camera)	18
2.1.2 Le modèle de "balayage" (push-broom)	19
2.1.3 De l'image physique à l'image numérique	19
2.1.4 Échantillonnage régulier carré et hexagonal	20
2.1.5 Repliement spectral et théorème de Shannon	23
2.1.6 Systèmes d'échantillonnage à haute résolution	24
2.2 Échantillonnage régulier	26
2.2.1 Déconvolution de Wiener pour les images aliasées	27
2.2.2 Extrapolation spectrale par minimisation de la variation totale	29
2.2.3 Adaptation de la cellule réciproque par maximisation de l'information mutuelle	31
2.3 Échantillonnage irrégulier et restauration	32
2.3.1 Échantillonnage perturbé de fonctions à bande limité et restauration . .	33
2.3.2 Vers une restauration basée sur un modèle linéaire complet de formation d'image	35
3 Interpolation non shannonienne	40
4 Détection <i>a contrario</i> de structures géométriques	46
4.1 Détection de points de fuite	47
4.2 Détection de rayures dans les films ¹	49
4.3 Détection de taches	50
4.4 Validation expérimentale du principe de Helmholtz dans des images dégradées	51
4.5 Détection de logos	51

5	Photogramétrie à faible B/H en milieu urbain	55
5.1	Méthode variationnelle	57
5.2	Méthodes affines par régions	60
5.3	Segmentation de textures	61
5.4	Conclusions partielles	62
6	Conclusions et Perspectives	63
A	Publications et encadrement	66
	Bibliographie	71

Introduction

Le but ultime de la vision par ordinateur et le traitement d'images est de développer des programmes pour permettre à un ordinateur de reconnaître dans les images numériques (qui lui sont présentés par l'intermédiaire d'une caméra, un scanner ou autre instrument d'acquisition d'images) les objets qui sont présents dans la scène visuelle, les relations spatiales entre eux, et toute information utile à une certaine tâche que nous voulons accomplir dans l'espace physique représenté par cette scène.

Malgré les travaux fondateurs de Marr [Mar82] dans les années 70, et les nombreuses recherches qui ont suivi, notre communauté scientifique est encore loin d'attendre ce but ultime, pourtant si quotidien et banal pour l'être humain qui saisit sans aucun effort un objet en se servant de ses yeux. Mais il est indéniable que des grands progrès ont été faits. Pour ceci il a été nécessaire d'avoir recours à presque toutes les branches des mathématiques appliquées (analyse de Fourier, géométrie, EDPs, analyse numérique, probabilités et statistiques, théorie de l'information, etc.) comme en témoigne ce document de synthèse. Cette contribution des mathématiques au traitement d'images n'a pas été sans retour pour les mathématiques elles mêmes. L'utilisation de modèles mathématiques très sophistiqués dans des buts applicatifs très précis a, elle aussi, nourri le développement des branches mathématiques concernées, en suggérant des extensions des théories disponibles afin de mieux s'adapter à la réalité physique, ainsi qu'en proposant des conjectures, dont certaines pas encore démontrés (voir [Des00] pour un exemple).

L'ordinateur ne dispose pour son interprétation de la scène que d'une (ou plusieurs) images discrètes qui ne sont autre chose qu'un tableau bidimensionnel de valeurs de gris qui représentent la luminosité de la scène visuelle dans une direction donnée. Une partie des recherches en analyse d'image et *pattern recognition* se concentrent dans le traitement de tels tableaux en espérant d'en déduire des informations sur les objets de la scène. C'est dans une certaine mesure ce que j'essaye de faire dans le chapitre 4, en nous inspirant de la théorie de la Gestalt de Wertheimer, Kanizsa, Metzger [Kan80, Wer23] et de sa formalisation mathématique proposé par Desolneux, Moisan, Morel [DMM02a].

Mais un tel programme est condamné à l'échec si nous ne tenons pas compte des relations : (i) entre l'image numérique (le tableau de valeurs mesurés), et l'image physique idéale, composée de photons qui incident dans la rétine ou plan image ; et (ii) entre l'image physique idéale et le monde physique qu'elle représente. Dans le deuxième cas il y a une évidente réduction de dimension de l'espace physique (à trois dimensions) à sa projection sur la rétine (bidimensionnelle). Mais dans le premier cas il y a aussi souvent une forte dégradation, ainsi qu'une réduction d'information entre l'image physique idéale et l'image numérique transmise à l'ordinateur. Dans

les chapitres 2 et 3 je décris mes contributions à la recherche de la meilleure approximation possible de l'image physique à partir du tableau discret, tout en gardant une estimation de la quantité d'information qui a été préservée de l'image physique idéale.

Enfin, le chapitre 5 retourne à la question de la relation entre le monde tridimensionnel et ses projections bidimensionnelles dans un cas très simplifié mais de grande utilité pratique : celui de l'estimation automatique de modèles numériques d'élévation (MNE) par photogrammétrie à partir de paires stéréo d'images satellitaires. Ce problème et les recherches menées au sein du CNES pour développer des systèmes précis et performants de calcul de MNE ont été l'une de nos principales motivations pour travailler sur les sujets décrits dans les chapitres 2, 3 et 4, et ce chapitre 5 met en évidence les liens très étroits qui existent entre les différentes étapes d'un système complet de vision par ordinateur. En effet, le calcul des disparités entre les deux images d'une paire stéréo ne peut atteindre les niveaux de précision requis en pratique que si les deux images discrètes sont débruitées, déflouées, rééchantillonnées et interpolées correctement en conformité avec les images physiques qu'elles représentent. De même, afin de déterminer d'une façon fiable (avec un faible nombre de faux positifs et négatifs) si une correspondance entre deux points homologues est significative ou simplement due aux nombreuses dégradations de l'image, il est essentiel de savoir combien d'information l'image restaurée garde encore de l'image physique.

Cette description porte sur mes recherches menées depuis 1999 à ce jour. Mon activité de recherche a commencé en 1995-1996 comme assistant en analyse numérique de l'économiste Carlo Graziani. Dans cette collaboration j'ai développé des méthodes numériques pour résoudre des équations linéaires aux dérivées partielles (d'intérêt en macroéconomie) dont certaines conditions initiales sont substituées par des conditions d'équilibre à l'infini [P8,P9,P10,P23]. S'agissant de recherches sur un sujet qui a très peu de liens avec ma spécialité actuelle, j'ai décidé de laisser cette partie de mes travaux en dehors de ce document de synthèse, et de renvoyer le lecteur aux publications dont je fournis la plus significative en pièce jointe.

Mon intérêt pour le traitement d'images a été motivé dans un premier temps par le problème de l'analyse et la reconnaissance d'images d'empreintes digitales en collaboration entre l'Université de la République, la Cour Suprême de Justice, le Service d'Identification Civile (Uruguay), et le Royal Institute of Technology (KTH, Suède). S'agissant d'un domaine applicatif assez différent, nécessitant des modèles mathématiques aussi différents, j'ai séparé la synthèse de ces recherches, menées entre 1995 et 2000, dans le chapitre 1.

Chapitre 1

Analyse et reconnaissance d'empreintes digitales

Les empreintes digitales sont utilisées depuis plus d'un siècle comme moyen d'identification des personnes, car le patron de sillons dont elles sont composées est unique à chaque individu et inchangé le long de sa vie. Bien avant l'apparition des ordinateurs, des procédures manuelles ont été développées afin de classifier des grandes quantités d'empreintes digitales dans un archive et de pouvoir chercher ensuite si une nouvelle empreinte y est présente ou pas. Ces méthodes de classification utilisent le nombre et position relative des points *core* et *delta* (voir figure 1.1) qui peuvent être caractérisées comme des singularités du champ d'orientations des sillons ayant un indice de Poincaré de $+\pi$ pour un *core* et $-\pi$ pour un *delta* (ailleurs, le champ d'orientations étant régulier, l'indice de Poincaré est nul).

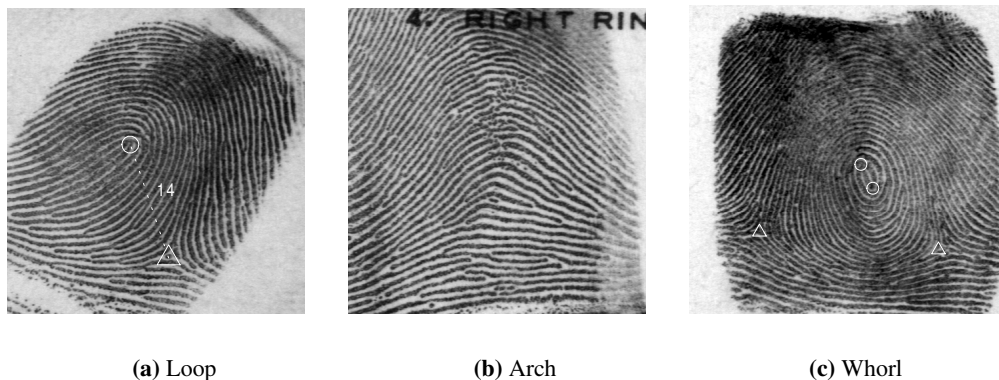


FIG. 1.1: Some fingerprint classes. The corresponding *core* and *delta* points are represented as white circles and triangles respectively. Note that the loop has a single core and a single delta, the arch has no core and no delta, and the whorl has two cores and two deltas.

Mais ces "macro-singularités" ne servent qu'à classifier l'archive. Pour identifier une per-

sonne il faut trouver des correspondances à un niveau plus fin, à savoir, des bifurcations et terminaisons des sillons (voir figure 1.2) qui constituent les *points caractéristiques* de l'empreinte appelées *minutiae* en anglais. Selon la législation des différents pays entre 12 et 25 de telles correspondances de points caractéristiques peuvent être requises pour identifier une personne en plus de la correspondance de la classification.

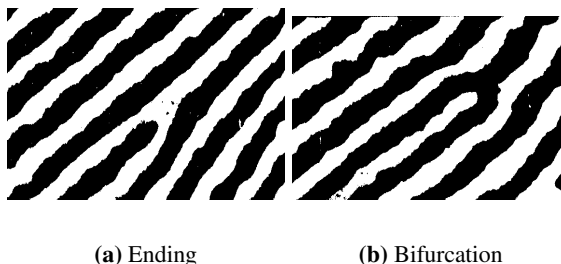


FIG. 1.2: *Some types of minutiae.* The black lines represent the ridges, whereas white represents the background. Note the subtle difference between forks and bifurcations : Both consist of a branching point in the ridge structure. In the case of a fork, however, the two branching ridges form a very small sharp angle, whereas in a bifurcation, they join the third ridge forming a rounded shape.

Avec l'avenant d'ordinateurs suffisamment puissants, des algorithmes automatiques pour la comparaison et recherche d'empreintes digitales ont été développés. Ces derniers émulent dans une certaine mesure la procédure manuelle que l'expert en identification fait quotidiennement en se servant de ses yeux pour reconnaître les détails qui servent à la classification ou l'identification. La figure 1.3(a) montre schématiquement les différents modules d'un tel système d'identification, appelé *AFIS* en anglais pour Automatic Fingerprint Identification System.

Mon travail sur les empreintes digitales au sein d'une équipe de l'Université de la République en Uruguay a été motivé par le souci de fournir le Département National d'Identification Civile (DNIC) et la Cour Suprême de Justice (SCJ) des certaines des fonctionnalités d'un tel système d'identification automatique. Il existait déjà à l'époque (1995-2000) des AFIS commerciales mais à un prix inabordable pour l'Etat uruguayen. Il faut distinguer ici un AFIS d'un système de contrôle d'accès qui est énormément plus abordable et vendu presque "dans la rue". Un système de contrôle d'accès ne gère d'habitude que quelques centaines d'utilisateurs, et sa seule fonction est de vérifier que l'utilisateur est bien celui qu'il dit être. Un AFIS, au contraire, doit gérer une base de données beaucoup plus grande et répondre à des questions du type "est-ce que cette personne est déjà dans la base ?", ce qui implique soit une classification très fiable, soit une recherche exhaustive, le plus souvent une combinaison des deux.

La DNIC gère un archive d'empreintes digitales de cinq millions d'individus (toute la population du pays) qui doit être consulté chaque fois que quelqu'un demande une carte d'identité pour la première fois, afin d'éviter des problèmes de double identité. La SCJ gère un archive indépendant de quelques centaines de milliers d'individus qui doit être consulté chaque fois qu'un suspect est arrêté, afin de trouver s'il a un dossier criminel ou pas. La réponse doit arriver dans certains délais légaux, car le procès qui suivra (avec ou sans prison) dépend des antécédents

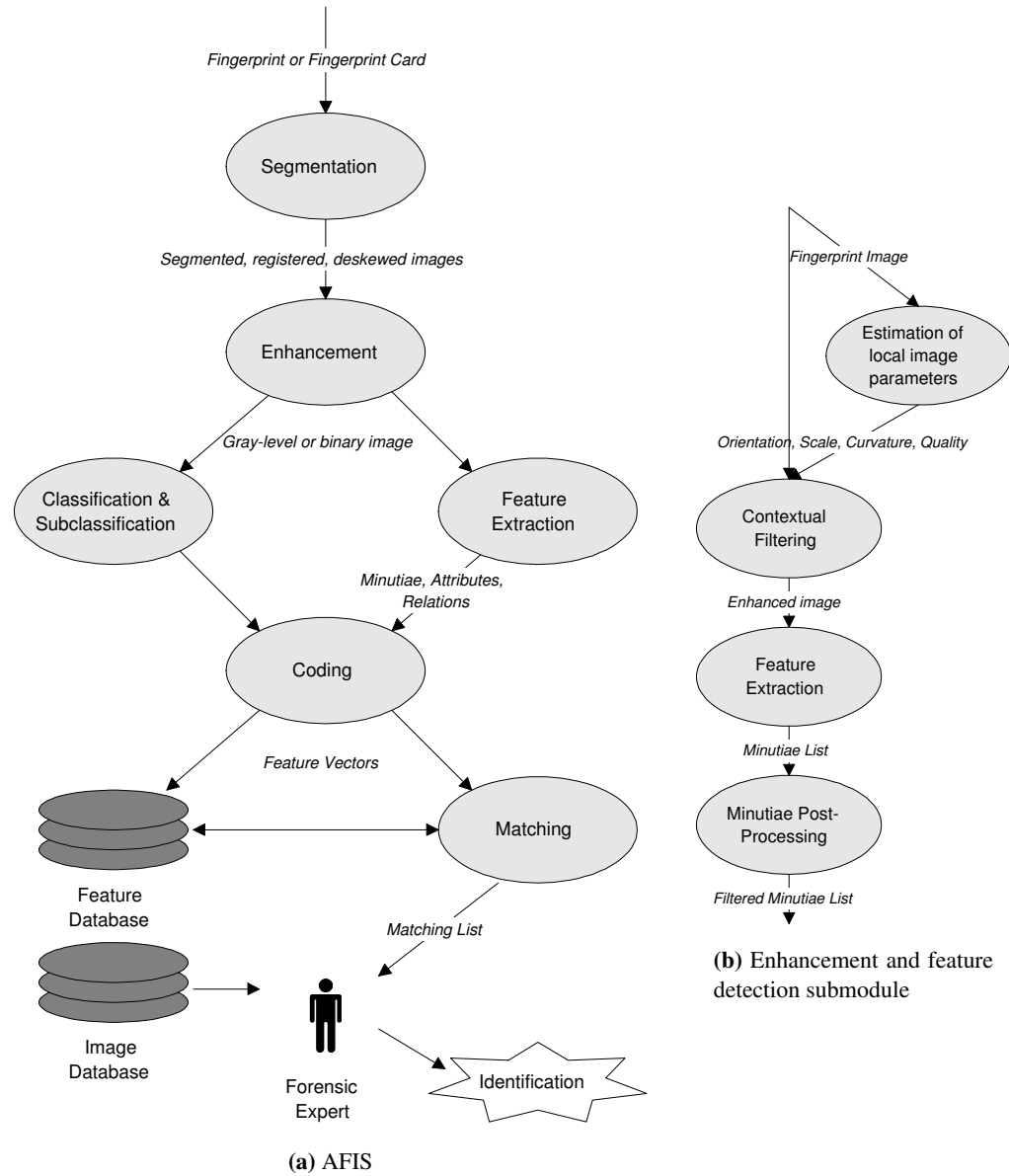


FIG. 1.3: Modules composing a complete Automatic Fingerprint Identification System (AFIS).

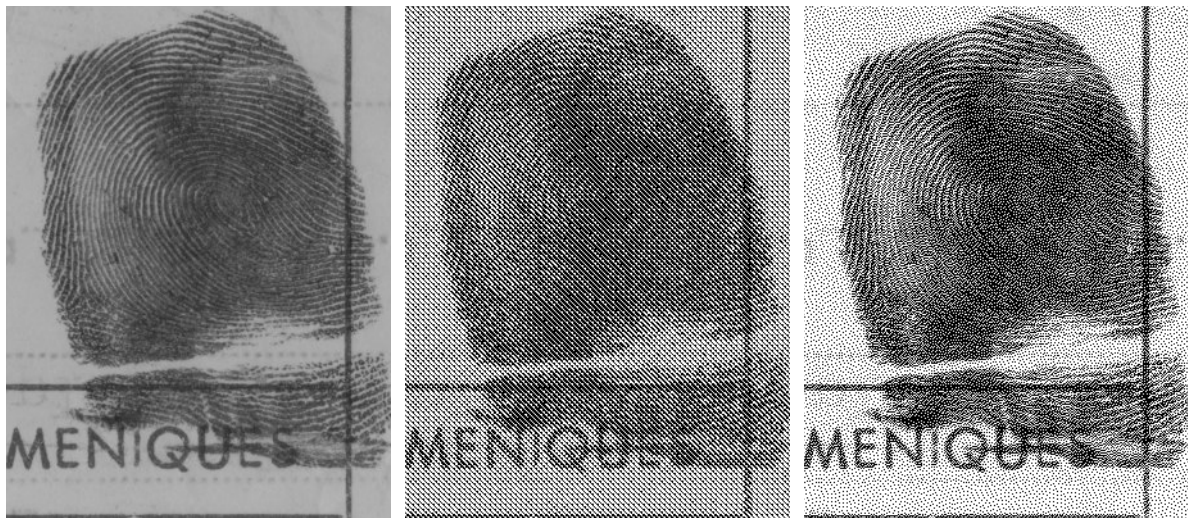
criminels du suspect.

1.1 Augmentation locale de contraste de la zone d'intérêt à bruit de fond contrôlé

La première tâche qui nous a été confiée par la SCJ (en 1995) a été de raccourcir les délais entre la détention d'un suspect et son identification en utilisant des moyens électroniques pour envoyer l'empreinte digitale de l'endroit de la détention à l'endroit où se trouve l'archive pour son identification. Le problème principal auquel nous nous sommes confrontés (mise à part la compression d'images et leur transmission surmontant les difficultés d'une internet encore en naissance) a été de faire une impression lisible et à taille naturelle de l'empreinte sur papier, afin de permettre sa recherche manuelle dans l'archive, l'agrandissement n'étant pas acceptable pour l'identification. Le défi était important sachant que ces images peuvent être très bruitées, à contraste très variable d'une zone à l'autre de la même image, et que le rapport entre les détails de l'image et la résolution des imprimantes courantes à l'époque était très faible (voir figure 1.4(a)).

En fait l'écart moyen entre deux sillons est de l'ordre de 0.015 pouces mais très variable (entre 45 et 90 sillons par pouce), tandis que les imprimantes laser disponibles étaient d'une résolution de 300 points par pouce en noir et blanc. Une simple binarisation de l'image pour amener l'image en niveaux de gris à une image imprimable en noir et blanc ne montre les détails nécessaires à l'identification que dans certaines régions assez contrastées. Un imprimante de 300 ppp peut aussi simuler quelques niveaux de gris avec l'utilisation d'un tramage de taille 5x5. Mais ce tramage entraîne une perte de résolution, qui fait que toute fréquence supérieure à environ 60 lignes par pouce soit distordue par un effet de Moiré à cause de son interférence avec le tramage. C'est bien ce que l'on observe dans la figure 1.4(b). L'effet de Moiré peut être réduit si le tramage géométrique est substitué par une diffusion aléatoire de l'erreur de binarisation telle que l'algorithme proposé par Floyd et Steinberg [FS75] (voir figure 1.4(c)), mais encore la plupart des détails ne sont pas visibles à cause de la grande variabilité de contraste. Une égalisation globale de contraste (figure 1.4(d)) ne suffit pas à corriger l'erreur et une égalisation locale (figure 1.4(e)) rehausse trop le bruit de fond et la texture du papier, tout en produisant des artifices dans la transition figure-fond qui rendent difficile l'interprétation de cette partie de l'image. La solution que nous avons envisagée avec Bergengruen, Curbelo et Drets [P20] consiste à séparer la zone d'intérêt M du fond par une combinaison de seuils et filtres morphologiques. Ensuite nous faisons une transition douce entre une égalisation globale sur le fond et une égalisation locale dans la zone d'intérêt M . Pour éviter les artefacts dans la zone de transition, l'égalisation locale, n'utilise que l'information intérieure à M pour calculer l'histogramme (figure 1.4(f)).

A ce jour beaucoup d'autres solutions bien meilleures sont possibles. Ce travail a cependant, à mon sens, le mérite d'avoir donnée une solution économique et pertinente à un problème réel. Nos collègues à la SCJ nous sont toujours très reconnaissants d'avoir contribué à réduire le temps d'attente de certains criminels qui ne l'étaient pas ! [P22, P21].



(a) image original

(b) binarisée avec une filigrane de 120lpp

(c) binarisée par diffusion d'erreur (Floyd-Steinberg)



(d) égalisation globale de contraste, puis FS

(e) égalisation locale de contraste, puis FL

(f) solution proposée dans [P20]

FIG. 1.4: Binarisation d'une empreinte digitale pour son impression à taille réelle dans une imprimante à 300 ppp sans perte de détails. La taille des images est doublée ici par rapport à la taille réelle, afin d'observer les défauts de chaque méthode et pas ceux de votre imprimante.

1.2 Débruitage directionnel adaptatif et détection de points caractéristiques

Ensuite je me suis intéressé (1996-1998) à la partie d'un AFIS qui consiste à débruiter l'image (*image enhancement*) pour détecter ensuite les points caractéristiques (*feature detection* ou *minutiae extraction*) qui servent à l'identification (voir figure 1.3(b)). Si l'image a été correctement débruitée la détection de points caractéristiques se fait très simplement par une binarisation locale suivie du calcul de son squelette et l'application d'opérateurs logiques locaux pour trouver les bifurcations et terminaisons de lignes. L'étape cruciale est donc le débruitage qui se fait, dans la quasi-totalité de la littérature spécialisée sur le sujet, par l'application de filtres locaux (tels que des filtres de Gabor) pour rehausser des patrons dans une certaine direction et plage de fréquences, au détriment de toute autre information qui est considérée comme du bruit et éliminée. La pertinence de ce filtrage local dépend donc d'une bonne estimation de l'orientation et fréquence des sillons, et de la cohérence entre la vraie image et le modèle de sillons périodiques et parfaitement parallèles sous-entendu par l'application des filtres directionnels. Néanmoins dans la plupart des travaux, l'application des ces filtres directionnels, ainsi que l'estimation de leurs paramètres, était faite d'une façon très discrète, parfois globale, dans les meilleurs des cas en traitement par blocs, ce qui produit parfois des artefacts sur l'image de sortie.

Ma principale contribution au sujet [P7] ¹ avec mon directeur de thèse de mastère Tony Lindeberg (KTH, Suède) a été de fournir une approche plus systématique au choix de paramètres de filtrage, d'en vérifier la cohérence avec l'image d'entrée, et d'éviter plusieurs artefacts dus au traitement par blocs en passant à la limite, c'est à dire en substituant le filtrage directionnel par une équation de diffusion directionnelle :

$$\begin{cases} \frac{\partial u}{\partial t}(x; t) = \nabla \cdot (\Sigma(x) \nabla u) \\ u(x; 0) = u_0(x) \end{cases} \quad (1.1)$$

où l'image originale u_0 est filtrée jusqu'à une certaine échelle t à estimer. ² Au lieu d'un coefficient de diffusion scalaire comme dans l'équation de Laplace, cette équation utilise un tenseur de diffusion $\Sigma(x)$ (une matrice 2x2 définie positive aussi à estimer) permettant de faire une diffusion plus forte dans une direction (celle du vecteur propre de Σ associé à son plus grand valeur propre) que dans la direction orthogonale, ou bien de faire une diffusion isotrope classique dans le cas où les deux valeurs propres sont égaux.

Notons au passage que l'équation de *diffusion anisotrope* de Perona-Malik [PM90], plus connue dans la communauté de traitement d'images pour ses bonnes propriétés de débruitage, consiste à utiliser un coefficient de diffusion scalaire $\Sigma(x) = \frac{1}{\nabla u(x,t)}$ ce qui rend l'équation non linéaire. Cette équation a donné lieu à des variantes améliorées comme le *mouvement par courbure moyenne* ainsi que le *scale-space morphologique affine* [AGLM93, ST93] qui agissent sur les lignes de niveau de l'image en les rendant plus régulières, et éventuellement les réduisant à une ellipse, puis un petit cercle, puis un point jusqu'à sa disparition. En plus toutes les lignes de niveau sont régularisées simultanément mais en préservant leurs relations d'inclusion originales.

¹ voir aussi [P11,P18] pour des versions préliminaires et [P33] pour une version plus étendue.

² On notera aussi la solution de cette équation à l'échelle t par $\mathcal{D}_{\Sigma,t}u_0 := u(\cdot; t) = \mathcal{D}_{t,\Sigma,1}u_0$.

Cette propriété qui est très utile pour l'élimination de bruit impulsionnel en préservant les bords contrastés des objets, n'est pas si utile dans le cas des empreintes digitales pour la simple raison qu'elle empêche un ensemble de petites taches (ensembles de niveau) très proches entre elles d'être réunies dans une seule. Or, c'est bien des opérations de regroupement de ce type qui permettent à l'oeil de reconnaître la continuité des sillons qui ne sont pas entourés par aucune ligne de niveau, mais composés d'un pointillé de petites taches d'encre noire, à peine un peu plus denses sur le sillon que sur les arêtes voisines.

L'équation de scale-space gaussien affine (1.1) que nous avons utilisée permet au contraire de faire ce genre de regroupements et dérive plutôt des idées de Witkin [Wit83] et Koenderink [Koe84] et les recherches de Lindeberg [Lin94] et Weickert [Wei98] qui ont suivi. Weickert a été le premier à proposer une EDP du type (1.1) pour le débruitage d'empreintes digitales [Wei95]. Dans son schéma appelé *coherent enhancing diffusion* l'orientation de la diffusion est dérivée du tenseur de structure

$$\Sigma(x) \approx (\mu(x; t) + \varepsilon I)^{-1} \quad \text{où} \quad \mu(\cdot; t) = g_s * (\nabla u(\cdot; t) \nabla u(\cdot; t)^T) \quad (1.2)$$

c'est à dire une version lissée par une gaussienne de variance s du produit tensoriel du gradient de l'image avec lui même. Cette opération permet d'avoir une estimation semi-locale (à l'échelle s) de l'orientation des sillons ainsi que de leur cohérence avec un modèle de sillons parallèles. Notre travail a consisté à compléter celui de Weickert en fournissant des estimations locales et optimales des paramètres d'échelle d'intégration s et de diffusion t . Cette sélection d'échelles a été basée sur les travaux précédents de Lindeberg [Lin96] qui prône l'utilisation des maxima locaux (le long des échelles t) d'un opérateur différentiel adapté à la structure géométrique cherchée. Quant à l'échelle d'intégration, elle doit être proportionnelle à l'échelle de diffusion $s = \beta t$ avec une constante de proportionnalité $\beta > 1$ qui dépend de la qualité de l'image. Dans des régions peu bruitées une petite échelle permettra une bonne estimation plus locale de l'orientation dominante, tandis que dans des régions plus bruitées une échelle s plus grande sera nécessaire. Nous avons trouvé qu'un indice $\mathcal{Q} \in [0, 1]$ de qualité image peut être déduit du tenseur de structure μ , car ce tenseur est plus contrasté et excentrique dans les régions cohérentes avec un modèle de sillons parallèles. Une fois calculé, cet indice est utile pour contrôler le choix de l'échelle d'intégration s , mais aussi pour pondérer le calcul de μ en sorte que les régions incohérentes avec un modèle de sillons parallèles (souvent des cicatrices) ne soient pas prises en compte. Nous pouvons interpréter cette procédure comme une sorte d'estimation robuste de l'orientation des sillons. Enfin nous avons proposé un algorithme qui combine toutes ces idées en cherchant un point fixe des fonctions d'échelle s , t et du tenseur de structure μ telles que

$$\begin{aligned} u(\cdot; t') &= \mathcal{D}_{\Sigma, t'} u_0 \\ t(x) &= \operatorname{argmax}_{t'} \mathcal{R}u(x; t') \\ s(x) &= \beta(\mathcal{Q}(x))t(x) \\ \mu(x) &= g_{s(x)} * (\nabla u(x; t(x)) \nabla u(x; t(x))^T \mathcal{Q}(x)) / g_{s(x)} * \mathcal{Q}(x) \\ \Sigma &= (\mu + \varepsilon I)^{-1} \lambda_{\max}(\mu) \end{aligned} \quad (1.3)$$

Une fois trouvé le point fixe, l'image débruitée $u(x; t(x))$ peut être binarisée pour trouver les points caractéristiques. A nouveau, dans les régions où l'image est incohérente avec le filtrage effectué (ce qui est indiqué par une petite valeur de \mathcal{Q}) nous risquons de faire de fausses détections.

Pour cette raison les points caractéristiques trouvés dans un point x tel que $Q(x) < 0.45$ sont rejetées, la valeur du seuil ayant été apprise pour faire la meilleure séparation possible de bonnes et fausses détections.

Des expériences sur quelques dizaines d'images sur lesquelles les points caractéristiques ont été marqués à l'avance ont montré l'efficacité de cette technique donnant un nombre de fausses détections sensiblement inférieur aux techniques précédentes. Ce qui a empêché la méthode de s'imposer dans des systèmes complets d'identification est sans doute la grande quantité des ressources computationnelles requises. Ceci ne devrait plus être un problème à ce jour étant donné les progrès dans la vitesse des processeurs, ainsi qu'en ce qui concerne les méthodes numériques pour résoudre des EDPs comme (1.1). A nouveau c'est Weickert [WtHRV98] qui a proposé une méthode numérique implicite et performante pour résoudre (1.1) en respectant la version discrète des axiomes de scale-space qui conduisent à cette EDP. Cette méthode, en plus de diminuer sensiblement le nombre d'itérations par rapport à la méthode explicite que j'ai utilisée, permet d'être combinée avec des schémas multigrid et de calcul en parallèle pour accélérer encore la convergence [WZtHRN97].

1.3 Mise en correspondance élastique

Enfin dans [P16] je m'intéresse à la mise en correspondance de points caractéristiques. Dans la plupart des travaux sur le sujet deux points caractéristiques de deux empreintes différentes sont considérés en correspondance si : (i) leur vecteurs de caractéristiques (orientation, type, distance à des points voisins, etc.) se ressemblent et (ii) ils sont proches après recalage affine entre les deux images. Or, le recalage affine est souvent insuffisant pour bien modéliser les déformations élastiques que subit la peau des doigts lors de l'acquisition de l'empreinte. Notre approche pour déterminer si deux ensembles de points caractéristiques (p_1, \dots, p_m) (détectés dans une image) et (q_1, \dots, q_n) (détectés dans la deuxième) consiste à trouver une déformation ϕ qui minimise l'énergie de déformation sous contraintes

$$\int (\Delta\phi)^2 + \lambda \sum_i g(\inf_j |\phi(p_i) - q_j|) \quad (1.4)$$

où $g(x) = \min(d^2, x^2)$ impose un seuil de distance d au delà de laquelle $\phi(p_i)$ et q_j ne sont plus considérés en correspondance. Une déformation ϕ de trop grande énergie ou avec un faible nombre de correspondances (p_i, q_j) indique que les deux images ne correspondent pas à la même personne.

Malgré les apparences, le problème de trouver un minimum local de ce problème variationnel non convexe, n'exige pas un énorme temps de calcul pour deux raisons :

- Il n'est pas nécessaire de calculer ϕ explicitement, il suffit d'utiliser une base de m fonctions base centrées sur les points p_i et ne faire que des calculs linéaires sur les m coefficients représentant ϕ dans cette base [Boo89] (m est d'habitude de l'ordre de quelques dizaines, beaucoup plus petite que l'image composée d'un quart de millions de pixels).
- La recherche d'un minimum local d'un tel problème non-convexe peut être ramenée à la recherche alternée de minima locaux de deux fonctionnelles convexes, comme il a été montré par Laurent Cohen [Coh96].

Les résultats présentés dans [P16] ne sont pas encore satisfaisants d'un point de vue pratique face à d'autres méthodes non variationnelles qui permettent de modéliser également des déformations élastiques (voir par exemple [JHB97]). La raison principale est que nous n'utilisons que la position (x, y) (la distance euclidienne $|\phi(p_i) - q_j|$) pour mesurer la ressemblance de deux points caractéristiques, tandis que plein d'autres caractéristiques sont disponibles et utilisées dans la littérature. Le but de ce travail était simplement de montrer la faisabilité d'une approche variationnelle pour modéliser les déformations dans la mise en correspondance dactilaire. Une fois montré cette faisabilité, la technique peut être généralisée pour inclure des mesures plus élaborées de ressemblance entre deux points. Une telle voie de recherche a été suivie très récemment dans le groupe d'Anil Jain [CDRJ05] en utilisant les orientations en plus des coordonnées spatiales pour mesurer la ressemblance entre deux points.

Dans le chapitre 3 je reprends cette même énergie de déformation dans le but d'interpoler des modèles de terrain.

Chapitre 2

Théorie de Shannon sur la formation, la restauration et l'interpolation d'images numériques

Les recherches que je présente dans ce chapitre ont été largement inspirées par le travail de Bernard Rougé [Rou97] sur la théorie de Shannon et la conception de systèmes imageurs à haute résolution, et par l'originale façon de Jean-Michel Morel de présenter l'analyse de Fourier et des généralisations peu connues de cette théorie dans ses cours [ML98] à l'École Polytechnique et à l'ENS de Cachan. Je présente un résumé de ces résultats, un peu à ma façon, dans la section 2.1. Mes propres contributions concernent l'exploitation de ces outils théoriques pour fournir des mesures objectives de la qualité d'un système imageur et leur application au développement de nouvelles méthodes de restauration d'images (sections 2.2 et 2.3). L'ensemble constitue un nouveau regard sur la théorie de Shannon que nous sommes en train de mettre sous la forme de notes de cours [P1] que j'utilise à présent comme bibliographie de base pour mon cours d'*Analyse de Fourier pour le Traitement d'Images* à l'Université de la République en Uruguay.

2.1 Modèles de formation et acquisition d'images

2.1.1 Le modèle de "caméra à trou d'épingle" (pin hole camera)

ou *camera oscura* comme il était connu dans l'antiquité consiste à faire une abstraction de l'appareil d'acquisition d'images comme s'il consistait d'une boîte noire ne laissant rentrer de la lumière que par un trou infiniment petit (le *point focal* O) situé à une distance F du *plan image* (ou la rétine dans le cas de l'oeil). La lumière se reflétant en ligne droite dans un objet de la scène photographiée, passe par le point focal et se projette sur le plan image, en reproduisant une projection bidimensionnelle de la lumière reflétée par les objets de la scène observée par l'appareil. Ainsi, un objet M situé dans l'espace dans les coordonnées $(X, Y, Z) \in \mathbb{R}^3$ et ne trouvant aucune occlusion sur la ligne MO sera projeté sur le point m de coordonnées (x, y) appartenant à l'intersection du plan image et la ligne de vue MO . C'est à dire que la *projection*

perspective s'écrit

$$P : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \quad P(X, Y, Z) = (x, y) = \left(-F \frac{X}{Z}, -F \frac{Y}{Z}\right)$$

dans la géométrie simplifiée, où l'origine des coordonnées (X, Y, Z) est le point focal O , l'axe OZ est perpendiculaire au plan image, et passe par son origine o , et les axes ox et oy du plan image sont parallèles aux axes OX et OY . Cette transformation projective (qui devrait s'écrire strictement $P : \mathbb{P}_3 \rightarrow \mathbb{P}_2$ pour inclure les points à l'infini) a en général 11 degrés de liberté qui dépendent de la construction interne de l'appareil (5 paramètres "intrinsèques") et de sa position et orientation (6 paramètres "extrinsèques"). La procédure qui consiste à calculer ces paramètres pour un appareil donné s'appelle d'habitude *calibration*.

2.1.2 Le modèle de "balayage" (push-broom)

diffère du précédent par le fait que l'on substitue la projection sur le plan image par une famille de projections sur une ligne. A chaque instant nous ne photographions que les points se trouvant sur un plan (disons $Y = Y_0$)

$$P_Y : \mathbb{R}^2 \rightarrow \mathbb{R} \quad P(X, Z) = x = -F \frac{X}{Z}$$

On reconstitue le plan image en faisant "balayer" ces projections unidimensionnelles le long de la dimension Y obtenant ainsi dans un cas le très simplifié

$$P : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \quad P(X, Y, Z) = (x, y) = (P_Y(X, Z), Y) = \left(-F \frac{X}{Z}, Y\right)$$

Dans ce mode chaque ligne (\cdot, y) est capturée dans un moment différent, et le point focal O avance avec Y . C'est le mode habituel d'acquisition dans les satellites d'observation terrestre, car ils capturent ses images à fur et à mesure qu'ils avancent dans leurs orbites, balayant ainsi le sol.

Ce que nous appelons *image physique*, *image idéale*, ou *paysage* est la fonction

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}$$

telle que $f(x, y)$ est la quantité d'énergie lumineuse (photons) qui s'est accumulé sur le point (x, y) du plan image pendant le temps de capture. Elle représente aussi la quantité de lumière réfléctée pendant ce temps par un certain objet M de coordonnées (X, Y, Z) , dans la direction MO , et tel que $P(X, Y, Z) = (x, y)$, sous l'hypothèse qu'aucun objet n'a occulté le passage de la lumière sur la droite OM .

2.1.3 De l'image physique à l'image numérique

Les dispositifs physiques sont en effet un peu plus complexes... Le flux de photons passe d'abord par plusieurs ouvertures généralement circulaires. Ces ouvertures peuvent être le diaphragme d'un appareil photographique, l'orifice d'un télescope ou d'un zoom, la pupille de

l'oeil. Comme ces ouvertures ne peuvent pas se réduire à un point minuscule car nous n'aurions pas assez de lumière rentrant dans l'appareil, les concepteurs des systèmes imageurs sont obligés d'ajouter un complexe système optique de lentilles ou miroirs visant à focaliser une certaine plage de profondeurs de champ sur le plan image. Les objets en dehors de cette plage restent donc plus flous car hors champ focal. Le fait que la lumière passe par une ouverture engendre un phénomène de diffraction qui rend aussi l'image plus floue. Si l'appareil bouge lors du temps d'obturation l'image sera dégradée dans la direction du mouvement.

Enfin, dernier point à ne pas négliger, les photons sont comptés au moment de l'impact par des capteurs tapissant la rétine ou le plan focal. Dans la rétine humaine, ces capteurs forment un réseau hexagonal en nid d'abeille. La plupart des capteurs CCD sont approximativement carrés et disposés en matrice. Dans tous les cas, on peut modéliser l'opération d'échantillonnage comme un décompte des photons effectué par chaque capteur dans un temps donné (le "temps d'obturation" pour un appareil photographique). Cette opération d'échantillonnage a deux temps du point de vue mathématique : l'intégration du champ photonique sur la surface du capteur, puis l'attribution de cette valeur au centre du capteur. La bonne nouvelle dans toute cette complexité physique est que l'effet de toutes ces dégradations sur le paysage peut-être bien modélisée mathématiquement par une convolution avec un noyau h suivie d'un échantillonnage sur le réseau régulier Γ sur laquelle se trouvent les capteurs

$$\begin{aligned} u &= f * h \\ g &= S(u) \end{aligned} \quad (2.1)$$

Le noyau de convolution h , dont la transformée de Fourier $\hat{h} = H$ s'appelle "Fonction de transfert en modulation" ou FTM, est une combinaison de noyaux dus aux différents phénomènes physiques

$$h = h_{\text{capteur}} * h_{\text{optique}} * h_{\text{mouvement}} \quad (2.2)$$

chacun pouvant être calculé en fonction des caractéristiques physiques de l'appareil. Le noyau qui correspond au décompte de photons dans le capteur par exemple, n'est autre chose que la fonction indicatrice de la forme du capteur. Par exemple, pour un capteur carré de taille c

$$h_{\text{capteur}}(x, y) = \mathbb{1}_{|x| \leq \frac{c}{2}, |y| \leq \frac{c}{2}} \quad H_{\text{capteur}}(u, v) = \text{sinc}(uc) \text{sinc}(vc) \quad (2.3)$$

2.1.4 Échantillonnage régulier carré et hexagonal

L'échantillonnage du paysage convolé correspond en général à la multiplication par le *peigne de Dirac*

$$S(u) = \Delta_{\Gamma} \cdot u \quad (2.4)$$

$$\Delta_{\Gamma} = \sum_{\gamma \in \Gamma} \delta_{\gamma} \quad (2.5)$$

composé de fonctions delta sur les points d'un réseau régulier

$$\Gamma = \{ne_1 + me_2 : n, m \in \mathbb{Z}\} \quad (2.6)$$

c'est à dire les multiples entiers d'une base $\{e_1, e_2\}$ de \mathbb{R}^2 . dans le cas le plus courant des capteurs CCD (comme ceux d'un appareil photo numérique), composé de capteurs carrés de taille c disposés en matrice, le réseau d'échantillonnage est simplement $c\Gamma_4 = c\mathbb{Z}^2$. La presque intégralité des traités d'analyse d'image ne présentent l'analyse de Fourier et la théorie de l'échantillonnage que sur un réseau carré comme celui-ci. Néanmoins il existe des raisons importantes pour considérer d'autres réseaux d'échantillonnage réguliers, notamment le réseau hexagonal, qui correspond à prendre comme base de \mathbb{R}^2 deux vecteurs e_1 et e_2 faisant un angle de 60° au lieu de 90° :

$$\Gamma_4 = \mathbb{Z} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathbb{Z} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (\text{réseau carré})$$

$$\Gamma_6 = \mathbb{Z} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathbb{Z} \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix} \quad (\text{réseau hexagonal}).$$

Mon travail de stage de DEA avec Morel et Rougé [P30, P26 appendice A] a commencé par étudier les propriétés de tous les réseaux d'échantillonnage raisonnables, en nous n'avons retenu que le réseau carré et l'hexagonal, sachant que l'hexagonal a certains avantages sur le carré, notamment :

- *Densité d'échantillonnage.* Dans des conditions isotropes le réseau hexagonal nécessite 13.4% moins d'échantillons qu'un réseau d'échantillonnage carré (ceci découle du théorème d'échantillonnage de Shannon et du théorème de paquetage optimale de sphères.
- *La connexité discrète* peut être définie sur le réseau hexagonal plus naturellement de façon cohérente avec une version discrète du théorème de la courbe de Jordan avec un seul concept de 6-connexité, tandis que le réseau carré exige l'alternance entre les concepts de 4-connexité et 8-connexité pour rester cohérent avec ce théorème.
- Sur le réseau hexagonal il est possible de définir des familles d'*ondelettes plus isotropes* avec une meilleure sélectivité d'orientations [Jaf89, Mey90, SA90b, SA90a, SL97, LS94, CS93, KS98, DGSS99], des transformées distance plus précises, des filtres isotropes et 3-séparables, etc.

Les principales difficultés qui ont empêché l'hexagonal de s'imposer sont probablement d'ordre technique, dû au fait que le matériel d'acquisition et affichage reste majoritairement carré, que les méthodes disponibles de passage d'un réseau à l'autre ne sont pas simples ou entraînent une déformation de l'image ou une perte d'information, et principalement que les outils pour faire de l'analyse de Fourier en hexagonal sont mal connus dans la communauté du traitement d'images, ou inutilement compliqués.

Mais en suivant les travaux de Morel et Ladjal [ML98] nous nous apercevons que l'analyse de Fourier sur n'importe quel réseau régulier n'est pas plus compliquée que sur un réseau carré. De fait, dans le cas des Séries de Fourier de fonctions périodiques, il suffit de supposer l'image périodique par rapport à un sous-réseau $\Gamma_N = N\Gamma$ (avec $N \in \mathbb{Z}$) du réseau d'échantillonnage pour que l'algorithme de la DFT sur Γ soit identique à l'algorithme pour la DFT habituelle (la FFT). Dans la plupart de la littérature [SL97, PN82, Mer79, MS83] un algorithme spécialisé (HFFT) doit être utilisé, à cause de l'insistance à utiliser une périodisation différente. De plus, dans [P26 appendice A] nous fournissons un algorithme performant et précis permettant le passage d'un réseau à l'autre, sous l'hypothèse que les images soient conformes aux hypothèses du théorème d'échantillonnage de Shannon.

L'un des points clés qui simplifient la théorie dans le cas d'un réseau régulier quelconque est que la transformée de Fourier d'un peigne de Dirac sur un réseau Γ généré par les vecteurs e_1, e_2

$$\hat{\Delta}_\Gamma = S^* \Delta_\Gamma^* \quad (2.7)$$

est proportionnelle à un peigne de Dirac sur le *réseau dual* Γ^* , qui est généré par les vecteurs e_1^*, e_2^* satisfaisant les conditions suivantes

$$\langle e_i, e_j^* \rangle = 2\pi \delta_{ij} \quad (2.8)$$

de biorthogonalité avec la base primale. Dans le cas carré le réseau dual devient tout simplement un réseau carré de taille inversement proportionnelle au réseau primal :

$$\Gamma_4^* = 2\pi \left(\mathbb{Z} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathbb{Z} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \quad (\text{réseau carré dual})$$

$$\Gamma_6^* = 2\pi \frac{2}{\sqrt{3}} \left(\mathbb{Z} \begin{pmatrix} \sqrt{3}/2 \\ -1/2 \end{pmatrix} + \mathbb{Z} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \quad (\text{réseau hexagonal dual})$$

mais dans le cas hexagonal les axes du réseau subissent une rotation de 30° comme il est montré dans la figure 2.1.

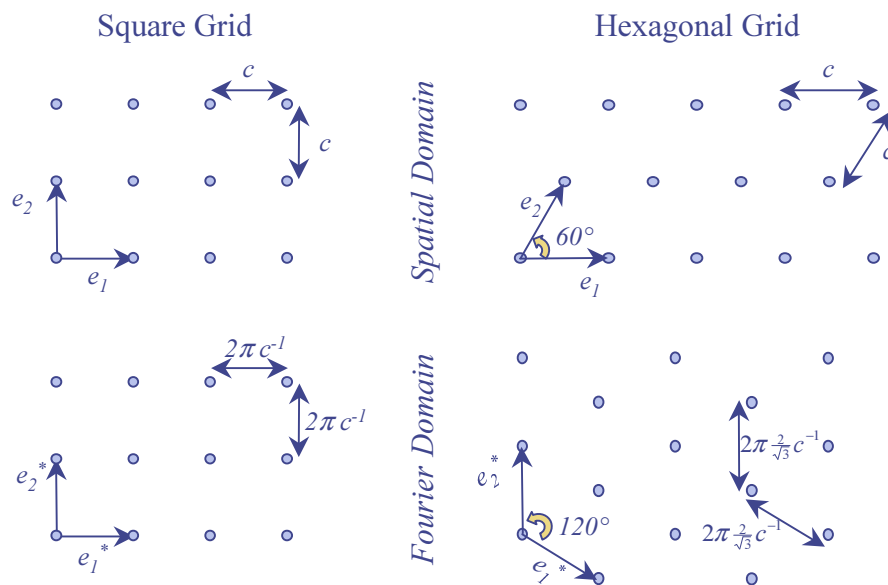


FIG. 2.1: *Le deux réseaux d'échantillonnage le plus courants et leurs réseaux duaux.* Quand une image est échantillonnée sur un réseau Γ , sa transformée de Fourier est périodisée par le réseau dual Γ^*

2.1.5 Repliement spectral et théorème de Shannon

En raison de l'équation (2.7) et du fait que la transformée d'une convolution $\mathcal{F}(f \cdot g) = \frac{1}{(2\pi)^2} \hat{f} * \hat{g}$, l'échantillonnage d'une image $u = f * h$ (le paysage convolé par le noyau de l'instrument) sur un réseau Γ entraîne la périodisation de sa transformée de Fourier

$$\hat{g} = \mathcal{F}(\Delta_\Gamma \cdot u) = \frac{S^*}{(2\pi)^2} \Delta_{\Gamma^*} * \hat{u} = \frac{S^*}{(2\pi)^2} (\hat{u} + \hat{u}_{\text{alias}}) \quad (2.9)$$

$$\hat{u}_{\text{alias}}(\omega) = \sum_{\gamma^* \in \Gamma^* \setminus 0} \hat{u}(\omega + \gamma^*) \quad (2.10)$$

¹ Cette périodisation mélange dans le cas général la "transformée pure" \hat{u} avec son repliement spectral \hat{u}_{alias} . Si le support spectral $\text{supp}(\hat{u})$ est contenu dans une cellule D^* ne se recouvrant pas par des translations dans Γ^* , c'est à dire si (Γ^*, D^*) est un paquetage au sens de la définition suivante, alors nous pouvons les isoler et reconstruire u à partir des échantillons g .

Definition 1 (Recouvrement, paquetage, pavage). Soit Γ un réseau régulier et $D \subseteq \mathbb{R}^2$ un sousensemble borné du plan. Nous disons que

- (Γ, D) est un recouvrement², si $\bigcup_{\gamma \in \Gamma} (D + \gamma) = \mathbb{R}^2$
- (Γ, D) est un paquetage², if $\forall \gamma \in \Gamma, \gamma \neq 0 \implies D \cap (D + \gamma) = \emptyset$
- (Γ, D) est un pavage, s'il est à la fois un recouvrement et un paquetage.

En effet si $\text{supp}(\hat{u}) \subseteq D^*$ et (Γ^*, D^*) est un paquetage (ou à la limite un pavage) alors $\mathbb{1}_{D^*} \hat{u}_{\text{alias}} = 0$, mais $\mathbb{1}_{D^*} \hat{u} = \hat{u}$. Donc

$$\frac{1}{|D^*|} \mathbb{1}_{D^*} \cdot \hat{g} = \frac{1}{(2\pi)^2} \hat{u}$$

$$\frac{1}{|D^*|} \overline{\mathcal{F}}(\mathbb{1}_{D^*}) * g = u$$

ce qui montre formellement le théorème de Shannon sur l'échantillonnage

Theorem 1 (Shannon-Whittaker). Considérons la fonction $f \in L^2(\mathbb{R}^d)$, un réseau d'échantillonnage d -dimensionnel Γ et une cellule réciproque $D^* \subset \mathbb{R}^d$. Si les conditions suivantes sont satisfaites

- (S1) (Γ^*, D^*) est un paquetage
- (S2) $\text{supp}(\hat{u}) \subseteq D^*$ (i.e. u est à bande limitée),

alors u peut être complètement restauré à partir de ses échantillons sur Γ , i.e. à partir de

$$g = \Delta_\Gamma \cdot f. \quad (2.11)$$

En fait, $g \in l^2(\Gamma)$, et la reconstruction est donnée par la convolution suivante

$$u(x) = \sum_{\gamma \in \Gamma} g_\gamma s(x - \gamma) \quad (2.12)$$

¹ Pour clarifier la notation, nous allons souvent faire abstraction en ce qui suit de la constante de normalisation $\frac{S^*}{(2\pi)^2}$ qui devient égale à 1 quand le taux d'échantillonnage est égal à 1.

² Ici l'égalité d'ensembles est au sens faible "presque partout", plus précisément $A = B$ si les deux ensembles $A \setminus B$ et $B \setminus A$ ont une mesure nulle.

de l'image échantillonnée g avec le noyau sinc généralisé $s = \frac{1}{S^*} \overline{\mathcal{F}}(\mathbb{1}_{D^*})$. L'équation de reconstruction est valable en général seulement dans le sens de la norme $L^2(\mathbb{R}^d)$. Si en plus $g \in l^1(\Gamma)$, alors l'égalité est point à point et la convergence uniforme.

Il existe des nombreuses variantes de ce théorème qui sont utiles aussi bien pour l'analyse mathématique que pour l'ingénieur en traitement d'images. Nous en donnons une résumée dans [P4 app. A, P26 app. A du chap. 1].

Bien que la condition de paquetage est assez flexible (nous en tirons profit à fond dans la section 2.2), le pavage le plus régulier possible, à savoir la cellule de Voronoï de Γ^* , est la plus utilisée, voir le seul considéré dans l'énorme majorité des travaux en analyse de Fourier pour traitement d'images.

$$D_{\text{Vor}}^* := \{\omega \in \mathbb{R}^2 : \forall \gamma^* \in \Gamma^*, \gamma^* \neq 0 \Rightarrow |\omega| \leq |\omega + \gamma^*|\} \quad (2.13)$$

Cette cellule devient un carré régulier de coté $2\pi/c$ pour un réseau Γ carré, un hexagone régulier si Γ est hexagonale, et un hexagone irrégulier dans tout autre cas (voir figure 4 dans [P4]). En raison de son caractère canonique nous appellerons cette cellule de Voronoï aussi la cellule réciproque R_Γ du réseau d'échantillonnage. Aussi $S^* = |R_\Gamma|$ désignera la surface de cette cellule réciproque.

2.1.6 Systèmes d'échantillonnage à haute résolution

La condition de bande limitée du paysage convolé $u = f * h$ exigée par le théorème de Shannon pour pouvoir faire une reconstruction parfaite de u est très difficile à atteindre en pratique. Le spectre du paysage f étant par règle générale infini, seulement la fonction transfert H peut être responsable du fait que le spectre de u soit à support compact assez petit. Or, le seul terme de H ayant un support spectral borné est celui dû à la diffraction. Mais dans plusieurs cas (y compris en imagerie satellitaire) l'optique est souvent de si bonne qualité que le support de H_{opt} est beaucoup plus large que celui de toute cellule réciproque D^* associée à un taux d'échantillonnage réalisable avec une matrice CCD de capteurs.

D'autre part la FTM associée au comptage de photons des capteurs carrés de taille c (H_{capteur} , voir équation (2.3)) bien que à support non-borné, devient quasiment négligeable après son premier passage par zéro à la fréquence $\pm\pi c$. Mais la cellule réciproque R_Γ du réseau d'échantillonnage $\Gamma = c\mathbb{Z}^2$ de la matrice CCD, n'arrive qu'aux fréquences $\pm\pi c/2$ (voir figure 2.2(a)).

Étant donné que les capteurs ne peuvent pas se recouvrir dans une barrette (mode push-broom) ou dans une matrice CCD (mode pin-hole) beaucoup de systèmes imageurs produisent des images assez aliasées. Pour d'autres raisons ceci est le cas des films en DVD tel qu'il a été observé récemment par Moisan et Morel [MM01]. Pourtant les ingénieurs du CNES [Rou97, LRS01] ont inventé deux variantes du mode *push-broom* permettant d'augmenter le taux d'échantillonnage (voir aussi figure 2.3)

1. Ajouter une deuxième barrette décalée par rapport à la première d'un demi pixel ($c/2$), comme on l'indique dans la figure 2.3(b). Cette configuration permet d'obtenir les systèmes Hipermode et Supermode.

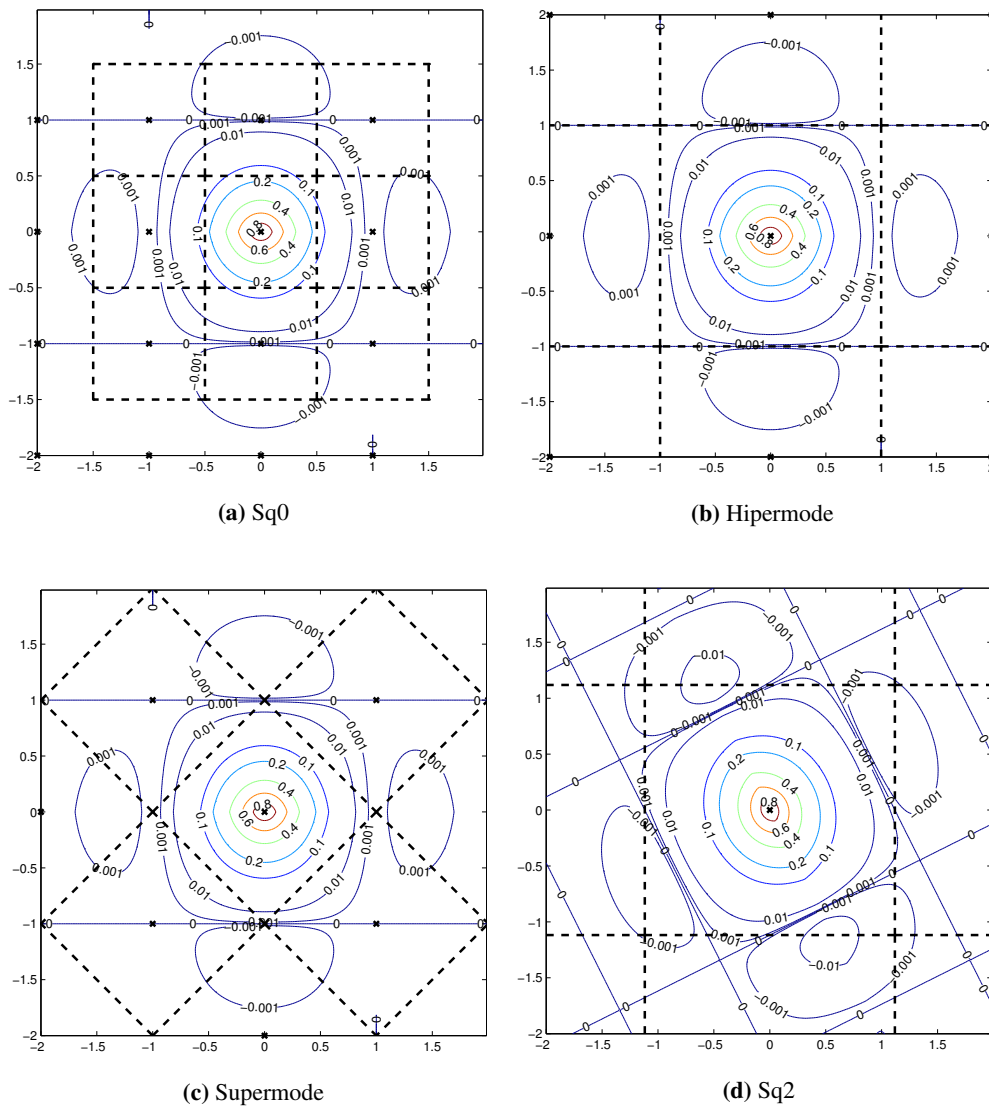


FIG. 2.2: Fonction de Transfert en Modulation et cellule réciproque R_{Γ} associée pour quatre systèmes imageurs satellitaires. (a) Sq0 - une seule barrette ou une matrice CCD ne suffit pas pour réduire assez le repliement spectral ; (b) Hipermode - avec deux barrettes satisfait presque parfaitement le théorème de Shannon, mais (c) Supermode - en jetant un échantillon sur deux de Hipermode représente un meilleur compromis qualité/coût de transmission (d) Sq2 - avec une barrette très inclinée nous arrivons aussi à satisfaire très bien les hypothèses de Shannon. *Légende* : Les points en forme de X représentent le réseau dual Γ^* au réseau d'échantillonnage. Les lignes pointillées représentent la cellule réciproque R_{Γ} , ainsi que ses translats par le réseau dual. Les autres lignes sont les lignes de niveau de la FTM de l'instrument.

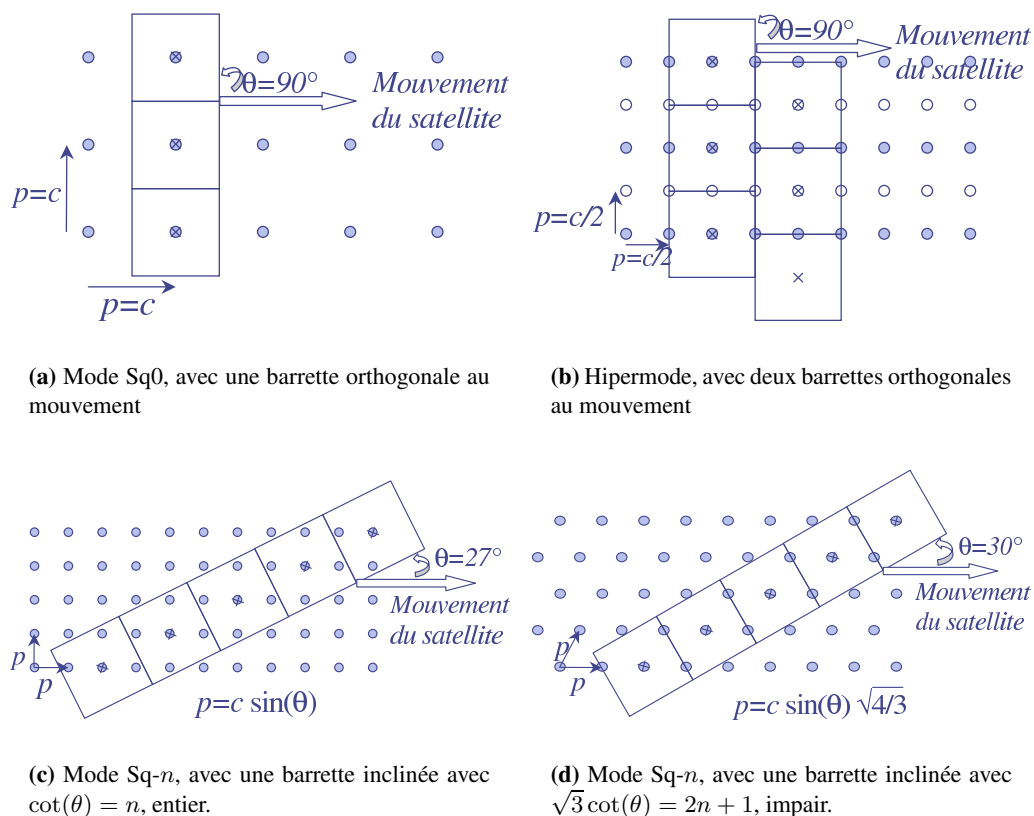


FIG. 2.3: Représentation schématique des différents modes d'échantillonnage à balayage. (a) Le mode classique avec une seule barrette, orthogonale au vecteur mouvement ; (b) L'un des modes à barrette multiple, toujours orthogonale au vecteur mouvement ; (c) Échantillonnage carré à barrette inclinée ; (d) Échantillonnage hexagonal à barrette inclinée.

2. Incliner la barrette et régler le taux d'échantillonnage de telle sorte que l'on obtienne un réseau régulier. Ceci est possible lorsque la cotangente de l'angle θ entre la barrette CCD et le vecteur vitesse du satellite est un nombre entier (donnant lieu à un réseau d'échantillonnage carré comme celui de la figure 2.3(c)), ou bien quand cette cotangente est multiple de $\frac{2n+1}{\sqrt{3}}$, où n est un nombre entier (donnant lieu à un réseau d'échantillonnage hexagonal comme celui de la figure 2.3(d)).

2.2 Échantillonnage régulier, mesures de résolution et restauration

Nous avons vu dans la section précédente à quel point les images numériques peuvent être contaminés par l'effet de repliement spectral (ou aliasage) en raison de la procédure physique d'acquisition, empêchant ainsi d'utiliser directement la formule d'interpolation fournie par le

théorème de Shannon. Quand j'ai fait une recherche bibliographique sur le sujet (voir introduction dans [P4,P26 ch 1]) j'ai été frappé par la quasi absence d'études systématiques pour analyser et corriger l'effet d'aliasage avec une perte minimale d'information. La pratique courante de l'ingénieur, consistant à utiliser une approximation plus régulière de formule d'interpolation (un sinc apodisé ou un spline cardinal d'ordre pas très élevé), arrive à corriger certains erreurs d'aliasage d'une façon assez pratique, mais pas forcément d'une façon optimale quand nous avons à notre disposition une information assez précise des conditions d'acquisition telles que la FTM et le réseau d'échantillonnage.

D'autre part les ingénieurs du CNES ont développé des nouveaux systèmes d'acquisition permettant de réduire au maximum l'effet de l'aliasage, mais ces solutions ne sont pas la panacée non plus car ils supposent une augmentation très significative du taux d'échantillonnage, et donc du coût associé au stockage et transmission de données entre le satellite et la base terrestre. Passé un certain point, l'augmentation du taux d'échantillonnage devient inutile car les mesures sont affectées des erreurs et le bruit l'emporte sur l'information apportée par les nouveaux échantillons.

Il faut donc savoir quelle est la meilleure approximation de l'image physique que nous pourrions obtenir avec chaque système (bien évidemment par simulation, avant de construire le système), ce qui nous amène à mesurer l'information mutuelle entre l'image échantillonnée par un système donné et le paysage idéal afin de donner une "note" à priori à chaque système, mais aussi à développer des algorithmes permettant de restaurer au mieux possible l'image physique et analyser les caractéristiques qualitatives du résultat.

Tel est le programme qui est décrit dans [P4,P26 ch 1], dont je donne une synthèse dans cette section. De plus, sachant que ces nouveaux systèmes imageurs peuvent fournir un échantillonnage régulier quelconque (pas forcément carré), nous avons développé ce programme dans le cadre plus général que d'habitude en traitement d'image, d'un réseau engendré par une base quelconque de \mathbb{R}^2 .

2.2.1 Déconvolution de Wiener pour les images aliasées

L'équation de formation d'image (2.1) n'est pas complète car elle ne tient pas compte des erreurs de mesure, qui sont modélisées en général par l'ajout d'un terme de bruit n

$$g = \Delta_{\Gamma} \cdot (f * h) + n \quad (2.14)$$

qui est modélisé habituellement par un bruit blanc d'écart type σ . Quand il n'y a pas de bruit ($n=0$) et le support spectral de $u = f * h$ est contenu dans la cellule réciproque R_{Γ} nous pouvons appliquer directement la formule de Shannon pour retrouver u . Si en plus \hat{h} est non nul sur R_{Γ} nous pourrions même retrouver la partie du paysage $\hat{f}|_{R_{\Gamma}} = \hat{u}/\hat{h}$. Mais pour les images bruitées, aliasées, et floues avec un noyau non inversible (à cause des passages par zéro de la fonction transfert) une telle opération ne ferait qu'amplifier énormément les bruit et l'aliasage.

L'approche la plus classique pour restaurer des images floues et bruitées est l'utilisation d'un Filtre de Wiener. Bien qu'il ne s'applique en pratique qu'à des images numériques, la dérivation de ce filtre se fait dans un cadre continu faisant abstraction de l'échantillonnage. Le modèle de formation d'image est donc :

$$g = h * f + n. \quad (2.15)$$

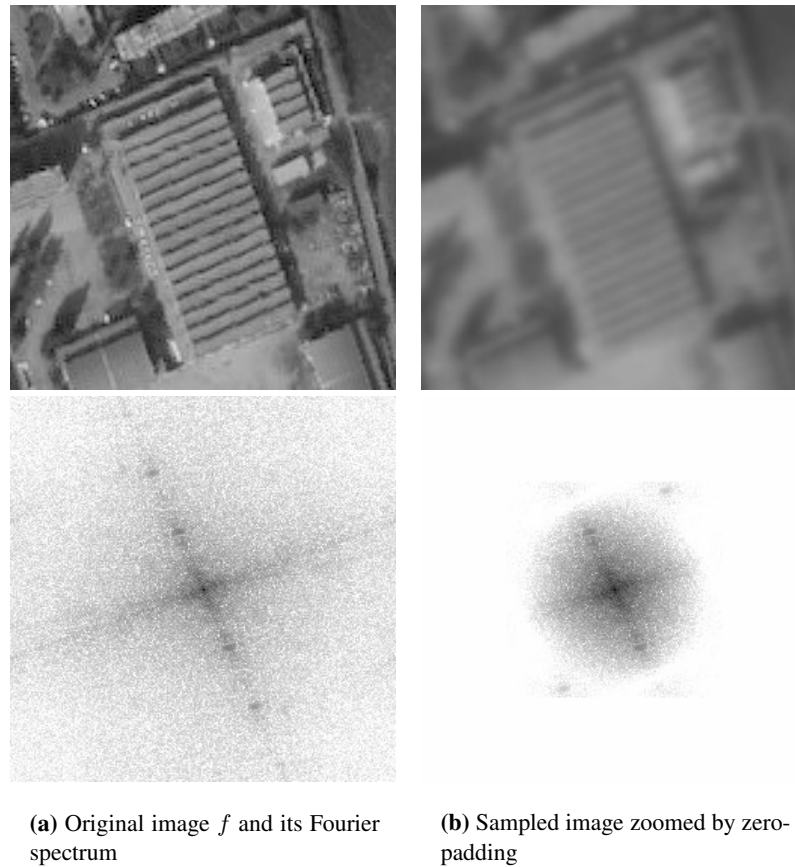


FIG. 2.4: Paysage original f et simulation de l'image échantillonnée selon le modèle (2.14)

La restauration \tilde{f} est trouvée par

$$\tilde{f} = k * g \quad \text{où} \quad k = \arg \min_k \mathbf{E}\{\|k * g - f\|^2\} \quad (2.16)$$

ou l'espérance est calculée par rapport aux variables aléatoires f et n , pour lesquelles des modèles statistiques Gaussiens colorés sont utilisés. Par exemple pour n nous supposons un bruit blanc, c'est à dire que tous ses coefficients de Fourier sont iid $\hat{n}(\omega) \sim N(0, \sigma)$, et pour f aussi tous les coefficients de Fourier sont supposés indépendants, mais avec des variances variables $\hat{f}(\omega) \sim N(0, F(\omega))$ avec $F(\omega) = (1 + |\omega|)^{-p}$ avec $p \in [1, 2]$. Ce modèle statistique d'image certes simpliste traduit bien le fait que l'image contient quelques discontinuités (les bords des objets) ce qui implique une décroissance des coefficients de Fourier inversement proportionnelle à la fréquence.

Avec ces hypothèses la solution est

$$\hat{k}(\xi) = \frac{\overline{H}(\xi)}{|H(\xi)|^2(1 + b^2(\xi))} \quad (2.17)$$

ou $b^2(\xi)$ est le rapport bruit-signal défini comme

$$b^2(\xi) = \frac{\text{Var}\{\hat{n}(\xi)\}}{\text{Var}\{\hat{u}(\xi)\}} = \frac{\sigma^2}{|HF|^2(\xi)} \quad (\text{rapport bruit-signal}) \quad (2.18)$$

Comme nous observons dans la figure 1.17.b dans [P4,P26 ch 1] cette restauration produit des effets catastrophiques, d'amplification de l'aliasage. Ceci est dû au fait que le modèle continu (2.15) ne tient compte de l'échantillonnage. Pour le corriger nous proposons dans [P4,P26 ch 1] d'appliquer cette même technique, mais en utilisant le modèle (2.14) qui en tient compte. la minimisation de l'équation (2.16) avec les mêmes supposés qu'avant donne

$$\hat{k}(\xi) = \frac{\overline{H}(\xi)}{|H(\xi)|^2(1 + a^2(\xi) + b^2(\xi))}. \quad (2.19)$$

où $a^2(\xi)$ est le rapport aliasage-signal défini comme

$$a^2(\xi) = \frac{\text{Var}\{\hat{u}_{\text{alias}}(\xi)\}}{\text{Var}\{\hat{u}(\xi)\}} = \frac{|HF|_{\text{alias}}^2(\xi)}{|HF|^2(\xi)} \quad (\text{rapport aliasage-signal}) \quad (2.20)$$

$$(2.21)$$

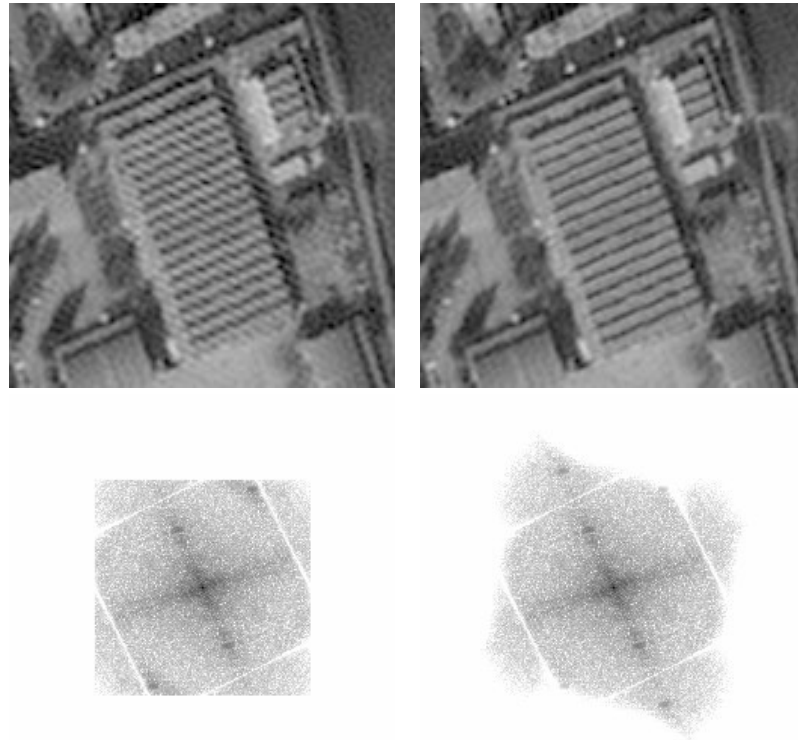
Une approche similaire a été proposée par Pratt [Pra78] il y a un quart de siècle, mais elle a été apparemment oubliée depuis. Néanmoins la figure 1.17.c dans [P4,P26 ch 1] montre bien que cette procédure très simple corrige correctement le phénomène d'aliasage.

Ce que cette procédure n'arrive pas à corriger est un phénomène de Gibbs (*ringing*) dû au fait que les coefficients de Fourier de f ne peuvent pas être restaurés du tout dans certains endroits très précis où la FTM de l'instrument est proche de zéro. Le filtre de restauration \hat{k} fait donc une coupure fréquentielle brutale à ces endroits ce qui se traduit en spatial par des oscillations artificielles autour des bords de l'image.

2.2.2 Extrapolation spectrale par minimisation de la variation totale

Pour corriger des phénomènes de Gibbs de ce genre (ainsi que ceux qui arrivent quand nous réduisons une image par coupure fréquentielle brutale suivi de souséchantillonnage), ce qui a été proposé par Malgouyres et Guichard [Mal99, MG01] est de remplir les régions où le spectre est inconnu par des coefficients de Fourier visant à réduire ces oscillations artificielles dues au phénomène de Gibbs, tout en gardant les coefficients restaurés dans la région connue. Un bon critère pour choisir ces coefficients consiste à minimiser la Variation Totale ($J(u) = \int |\nabla u|$) introduite par Rudin, Osher et Fatemi [ROF92, RLO03] dans le cadre du débruitage d'images. Ce choix est préférable à d'autres approches de régularisation comme celle de Tikhonov [Tik63] (qui consiste à minimiser $\int |Du|^2$) car la variation totale pénalise très fortement les oscillations (comme le phénomène de Gibbs) sans lisser les bords des objets présents dans l'image.³ La

³ Le principal inconvénient de la variation totale est qu'elle pénalise aussi fortement les régions texturées, car elles ne sont pas à variation totale bornée comme le montre Gousseau [GM01]. En conséquence Meyer [Mey02] a proposé un modèle d'image qui permet de séparer la géométrie (à variation totale bornée) de la texture (dans un espace dual à BV). Aussi Solé [BCRS03] a proposé une application plus locale de la minimisation de la variation totale permettant de mieux préserver les textures.



(a) Standard Wiener Restoration using the filter in equation (2.17)

(b) Wiener Restoration for sampled images using the filter in equation (2.19)

FIG. 2.5: Deconvolution of the image in figure 2.4 by two kinds of Wiener Restoration.

solution proposée par Malgouyres et Guichard pour extrapoler le spectre d'une image en dehors d'une région Ω est donc de minimiser la variation totale

$$\tilde{f} = \arg \inf_f \int |\nabla f(x)| dx + \lambda \int_{\Omega} |H\hat{f} - \hat{g}|^2 d\xi. \quad (2.22)$$

plus une attache aux données dérivée du modèle de formation d'image (2.14).⁴ Si l'on résout cette équation en utilisant comme domaine spectral de l'attache aux données $\Omega = R_{\Gamma} = D_{\text{Vor}}^*$ la cellule réciproque de Voronoï associée au réseau d'échantillonnage, comme il était proposé dans [Mal99, MG01], le résultat est à nouveau catastrophique quand on l'applique à des images fortement aliasées (voir figure 1.15.c dans [P4,P26 ch 1] pour un exemple). L'erreur a été bien

⁴ L'utilisation d'une attache aux données $\sum_{x \in \Gamma} |(h * f)(x) - g(x)|^2$ exprimée directement en spatial est aussi possible, mais il a été montré qu'elle conduit à un phénomène de pixelisation ne fournissant pas une bonne approximation du paysage original.

évidemment le choix de cellule réciproque. Mais comme j'ai souligné à la fin de la section 2.1.5 le théorème de Shannon nous laisse la liberté de choisir la forme de cette cellule pour l'adapter aux régions spectrales gardant la plus grande information du paysage original (face au bruit et au aliasage) à condition de respecter la condition de paquetage. Je reviendrai dans la section 2.2.3 sur le choix d'une telle cellule optimale $D_{R\text{-opt}}^*$. La figure 1.16.c dans [P4, P26 ch 1] montre que la minimisation de 2.22 sur $\Omega = D_{R\text{-opt}}^*$ la cellule optimale résout complètement les problèmes d'aliasage associés à l'utilisation de $\Omega = D_{\text{Vor}}^*$ la cellule de Voronoï, ainsi que les problèmes de *ringing* associés au filtre de Wiener (2.19).

2.2.3 Adaptation de la cellule réciproque par maximisation de l'information mutuelle

Supposons qu'une image physique f ait été capturée et échantillonnée sur un réseau régulier Γ par un système d'acquisition conforme au modèle de formation d'image de l'équation (2.14). Soit D^* une cellule quelconque telle que (Γ^*, D^*) forme un paquetage ou un pavage, en conformité avec le théorème de Shannon. Afin de mesurer la quantité d'information utile (hors bruit et aliasage) que nous pouvons espérer récupérer du paysage convolé $u = f * h$ dans cette cellule, nous considérons l'information mutuelle $I(\hat{g}, \hat{u}|_{D^*})$ entre l'image échantillonnée et le paysage convolé.

Dans la mesure où $\hat{g}(\xi)$ et $\hat{u}(\xi')$ sont indépendants pour tout $\xi, \xi' \in D^*$ avec $\xi \neq \xi'$ (ce qui est vrai pour le modèle gaussien indépendant utilisé dans la section 2.2.1), cette information mutuelle est additive. Nous définissons donc la *densité d'information effective*

$$I_{\text{eff}} := \frac{1}{(2\pi)^2} \int_{D^*} I(\hat{g}(\xi), G(\xi)) d\xi = I(\hat{g}, \hat{u}|_{D^*}) \quad (2.23)$$

comme mesure de fidélité de l'image échantillonnée au paysage original, car elle mesure le nombre moyen de bits d'information commune entre les deux images par unité de surface.

Bien que assez simpliste, le modèle statistique utilisé reste un outil pratique car il fournit une minoration globale de la densité d'information mutuelle, qui est relativement facile à calculer [P26 chap 1, P4] :

Theorem 2. *Soit g une image obtenue par échantillonnage du paysage convolé $u = h * f$ selon l'équation (2.14). Rappelons que $\hat{g} = \hat{u} + \hat{u}_{\text{alias}} + \hat{n}$ en notant $\hat{u}_{\text{alias}}(\omega) = \sum_{\gamma^* \in \Gamma^* \setminus \{0\}} \hat{u}(\omega + \gamma^*)$. Alors l'information mutuelle entre $\hat{g}(\xi)$ et $\hat{u}(\xi) = H(\xi)\hat{f}(\xi)$ est toujours minorée par :*⁵

$$I(\hat{g}(\xi), \hat{u}(\xi)) \geq \frac{1}{2} \log_2 \left(\frac{2^{2H(\hat{u}(\xi))} + 2^{2H(\hat{u}_{\text{alias}}(\xi))} + 2^{2H(\hat{n}(\xi))}}{2^{2H(\hat{u}_{\text{alias}}(\xi) + \hat{n}(\xi))}} \right). \quad (2.24)$$

Cette minoration est atteinte lorsque $\hat{u}(\xi)$, $\hat{u}_{\text{alias}}(\xi)$ et $\hat{n}(\xi)$ sont toutes des variables aléatoires Gaussiennes, indépendantes à moyenne nulle. Dans ce cas l'information mutuelle se réduit à l'expression suivante

$$I(\hat{g}(\xi), \hat{u}(\xi)) = \frac{1}{2} \log_2 \left(1 + \frac{1}{a^2 + b^2}(\xi) \right). \quad (2.25)$$

⁵ Observez que le symbole H est utilisé pour l'entropie tandis que H désigne la FTM.

La densité d'information effective I_{eff} fournit une mesure de fidélité de g par rapport à $u = (f * h)$, mais elle dépend du choix de la cellule réciproque D^* . Le plus naturel est donc de choisir D^* de sorte qu'elle maximise I_{eff} .

$$D_{\Gamma\text{-opt}}^* := \arg \sup_{\{R: (\Gamma^*, R) \text{ est un paquetage}\}} \int_R I(\hat{g}(\xi), \hat{u}(\xi)) d\xi \quad (2.26)$$

Dans [P26 chap 1, P4] je montre que, sous certaines hypothèses techniques assez faciles à vérifier sur les rapports entre la FTM et le réseau dual (voir Lemme 5 et Proposition 1), ce problème d'optimisation a une solution unique (à moins d'un ensemble de mesure nulle). En plus cette solution unique forme avec le réseau dual Γ^* un pavage du plan et elle peut se calculer simplement par l'expression

$$D_{\Gamma\text{-opt}}^* = \left\{ \xi \in \mathbb{R}^d : I(\xi) > I(\xi + \gamma) \text{ pour tout } \gamma \in \Gamma^*, \gamma \neq 0 \right\} \quad (2.27)$$

où $I(\xi) = I(\hat{g}(\xi), \hat{u}(\xi)) = \frac{1}{2} \log_2 \left(1 + \frac{1}{a^2 + b^2}(\xi) \right)$ est l'information mutuelle.

D'autres mesures de fidélité de $\hat{g}(\xi)$ par rapport à $\hat{u}(\xi)$ sont possibles (notamment une mesure de résolution effective r_{eff}) et je donne une axiomatique pour trouver des mesures raisonnables et les cellules optimales correspondantes. Néanmoins je n'ai trouvé aucune différence entre les cellules optimales obtenues par maximisation de r_{eff} et I_{eff} . Ces mesures de fidélité sont toujours à comparer à une la mesure "nominale" qui est la mesure optimale qui serait obtenue par le même réseau d'échantillonnage si le système optique était idéal (c'est à dire $H = \mathbb{1}_{D^*}$, donc pas d'aliasing). La résolution nominale r_{nom} , par exemple, donne simplement le nombre d'échantillons par unité de surface, tandis que I_{nom} indique le nombre de bits par unité de surface nécessaires à coder l'image. Des valeurs élevés de r_{eff} et I_{eff} indiquent une haute fidélité, tandis que des valeurs élevées de r_{nom} ou I_{nom} indiquent un coût de stockage et transmission élevé. Les rapports $r_{\text{eff}}/r_{\text{nom}}$ et $I_{\text{eff}}/I_{\text{nom}}$ (mesuré en bits d'information utile par échantillon) donnent une bonne indication de l'efficacité du système à trouver un bon compromis entre fidélité à l'original et coût de stockage et transmission. Du tableau 2.1 nous pouvons déduire que le système Supermode (en presque parité avec les systèmes Hex0 –barrette inclinée à 60° – et Sq1 –barrette inclinée à 45°) est parmi les plus efficaces dans ce sens, bien plus efficace que Hipermode, malgré le fait que Hipermode est un peu mieux résolu. Un raisonnement similaire a amené l'équipe du CNES à retenir Supermode plutôt que Hipermode pour le satellite SPOT5 lancé en 2002.

2.3 Échantillonnage irrégulier et restauration

Le modèle de formation d'image (2.14) décrit dans la section précédente n'est pas tout à fait complet, spécialement dans des systèmes d'acquisition d'images à balayage tels que les satellites, car en raison des vibrations inévitables du satellite lors de la capture de l'image, le réseau d'échantillonnage devient irrégulier. Un modèle mieux adapté à cette situation est le suivant

$$g = \Delta_\Lambda \cdot (h * f) + n \quad (2.28)$$

System	r_{nom} (échant/ c^2)	r_{eff} (échant/ c^2)	I_{nom} (bits/ c^2)	I_{eff} (bits/ c^2)	$r_{\text{eff}}/r_{\text{nom}}$ (%)	$I_{\text{eff}}/I_{\text{nom}}$ (bits/échant)
Sq0	1.00	0.76	6.66	3.91	76	3.91
Hex0	1.15	0.93	7.44	5.14	80	4.45
Supermode	2.00	1.63	9.51	8.76	81	4.38
Sq1	2.00	1.66	9.99	8.91	83	4.46
Hex1	3.46	2.22	11.40	11.35	64	3.28
Hipermode	4.00	1.99	10.30	10.27	49	2.57
Sq2	5.00	2.26	11.68	11.67	45	2.33
Hex2	8.08	2.28	11.82	11.82	28	1.46

TAB. 2.1: Évaluation quantitative de plusieurs systèmes imageurs ayant le même système optique et des capteurs de la même taille c . Seul le réseau d'échantillonnage (avec une ou plusieurs barrettes à des angles différents) et une partie de la FTM (due au mouvement pendant le temps d'exposition) changent de l'un à l'autre

où n reste un bruit gaussien de variance σ^2 , f l'image physique, h le noyau de convolution associé au système imageur, mais l'ensemble $\Lambda = \{\lambda_k\}_k$ de points sur lesquels l'image est échantillonnée par le peigne de Dirac

$$\Delta_\Lambda = \sum_{\lambda_k \in \Lambda} \delta(\cdot - \lambda_k).$$

n'est pas forcément un réseau régulier. Ceci est le *modèle général d'échantillonnage irrégulier*, mais je m'intéresse à un cas particulier, spécialement pertinent pour modéliser les vibrations d'un satellite. Nous l'appelons *échantillonnage perturbé* car Λ prend la forme particulière

$$\Lambda = \mathbb{Z}^2 + \varepsilon(\mathbb{Z}^2) \quad (2.29)$$

où $\varepsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est une fonction de perturbation "régulière et de faible amplitude" dans le sens suivant : L'amplitude A de la perturbation (mesurée comme le max ou l'écart type de ε) est de l'ordre de 1 pixel ou moins, et le support spectral de la perturbation $\text{supp}(\hat{\varepsilon}) \subseteq [-\frac{2\pi}{T}, \frac{2\pi}{T}]^2 = R_\varepsilon$ est beaucoup plus petit que la cellule réciproque $R_{\mathbb{Z}^2} = [-\pi, \pi]^2$ associée au réseau d'échantillonnage sans perturbation. Ceci dit la période de la fréquence maximale de vibration est $T > 2$.

2.3.1 Échantillonnage perturbé de fonctions à bande limité et restauration

Afin d'aborder ce problème je me suis intéressé dans un premier temps avec Bernard Rougé et Stéphane Jaffard [P14, P26 chap 2, P15] à un cas plus simple où la FTM $\hat{h} = \mathbb{1}_{[-\pi, \pi]^2}$ peut être bien approchée par une fonction porte idéale, assurant que le paysage convolé à échantillonner $u = h * f$ soit bien à bande limitée dans la cellule $R_{\mathbb{Z}^2} = [-\pi, \pi]^2$.

Dans ce cas, et en absence de bruit il existe une généralisation du théorème d'échantillonnage de Shannon, due à Kadec [Kad64] (voir aussi [You80, section 1.10, page 42]) et généralisée au cas bidimensionnel par Favier et Zalik [FZ95] et avec une borne moins serrée par Chui et Shi

[CS96]. Le théorème montre qu'il existe une formule de reconstruction exacte et stable de u à partir des échantillons g , pourvu que la perturbation de l'échantillonnage soit bornée

$$\sup_{\xi} |\varepsilon(\xi)| < 0.11$$

Jaffard [Jaf00] montre, lui, que dans certaines conditions raisonnables sur la forme de la perturbation cette constante de majoration peut être remplacé par 0.25, mais pas plus (on peut déjà trouver des contre-exemples avec $|\varepsilon(\xi)| \leq 0.25$). Ce résultat montre bien la difficulté du problème car, dans des cas réels l'amplitude de la perturbation est bien plus grande. En plus la formule de reconstruction que l'on peut déduire directement du théorème est très peu performante du point de vue algorithmique.

La seule façon de s'en sortir est d'augmenter le taux d'échantillonnage. Une généralisation possible de ce concept au cas de l'échantillonnage irrégulier est la *densité de Beurling*

$$D(\Lambda) = \lim_{r \rightarrow \infty} \inf_{x \in \mathbb{R}^d} \frac{\#(B_r(x) \cap \Lambda)}{(r)^d} \quad (2.30)$$

qui mesure le nombre moyen d'échantillons par unité de surface (la dimension $d = 2$ dans le cas des images). $D(\Lambda) \geq 1$ est bien évidemment une condition nécessaire pour pouvoir reconstruire u (à bande limitée dans $[-\pi, \pi]^2$) à partir de $g = \Delta_{\Lambda} \cdot u$, mais la condition suffisante ne peut dépasser l'inégalité stricte $D(\Lambda) > 1$, comme il a été montré par Beurling et Landau [Beu89, Lan67]. Tout algorithme général de restauration stable et exacte d'échantillonnage irrégulier à bande limitée nécessite donc un sur-échantillonné par rapport à la densité critique $D(\Lambda) = 1$.

L'un des algorithmes les plus performants de restauration irrégulier-régulier est l'algorithme ACT développé initialement par Feichtinger et ses collaborateurs [FGS95] et analysé, amélioré et généralisé ultérieurement par Gröchenig, Strohmer [GS00], et Rauth [Rau98]. L'algorithme combine très astucieusement des *poids Adaptatifs* afin de réduire le numéro de condition du problème et de produire des estimations explicites du taux de convergence, la méthode itérative des *gradients Conjugés* pour accélérer la convergence, et la formulation du problème comme un système linéaire dont la matrice à inverser a une structure de *Toeplitz* ce qui permet de calculer très rapidement chaque itération de la méthode des gradients conjugués (CG). En plus les étapes d'initialisation de l'algorithme sont aussi très rapides grâce à l'utilisation de l'algorithme proposé par Beylkin pour calculer des séries de Fourier non-harmoniques [Bey95, Bey98, Pot01].

Le taux de convergence de l'algorithme CG est déterminé par le conditionnement κ , de la matrice de Toeplitz. Plus précisément à chaque itération l'erreur d'approximation est réduit d'un facteur $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ [GVL96]. Étant donné la structure du problème, Gröchenig et Strohmer [Grö92, GS00] donnent une caractérisation très utile du conditionnement κ dans le cas bidimensionnel :

Proposition 1 (taux de convergence de l'algorithme ACT). *Si l'ensemble d'échantillonnage Λ est γ -dense avec*

$$\gamma < \frac{\log 2}{4\pi} \quad (2.31)$$

alors l'algorithme ACT converge à la solution exacte avec un taux de convergence $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$, dominé par le conditionnement κ qui est majoré par

$$\kappa \leq \frac{4}{(2 - e^{4\pi\gamma})^2} \quad (2.32)$$

Le concept de densité d'échantillonnage utilisé ici est plus strict (et plus local) que celui de Beurling. On dit qu'un ensemble d'échantillonnage Λ est γ -dense si

$$\bigcup_{\lambda_k \in \Lambda} D_\gamma(\lambda_k) = \Omega \quad (2.33)$$

i.e. si le domaine $\Omega \subseteq \mathbb{R}^2$ de l'image peut être recouvert par des disques $D_\gamma(\lambda_k)$ de rayon γ qui sont centrés sur les échantillons λ_k . On désigne aussi par *écart maximal* (*maximal gap* en anglais) de Λ la valeur $v = 2\gamma$, s'il γ est la valeur plus petite pour laquelle Λ reste γ -dense (l'écart maximal v peut être compris comme le diamètre maximal des cellules de Voronoï associées à l'ensemble d'échantillonnage Λ). Les deux concepts de densité sont liés car si Λ est γ -dense, alors $D(\Lambda) \geq \frac{1}{\pi\gamma^2}$, mais il n'y a pas de relation inverse en général.

Ce qui est important à retenir de ce résultat est que la condition de convergence (2.31) de l'algorithme ACT implique un échantillonnage largement plus dense que l'échantillonnage critique $D(\Lambda) \geq 1$, tandis que dans le cas de l'imagerie satellitaire (et dans la plupart des systèmes imageurs en pratique) nous sommes toujours en dessous de l'échantillonnage critique ($D(\Lambda) \leq 1$). Dans ces conditions nous ne pouvons espérer une convergence à la solution exacte, encore moins une solution unique. Mais on peut chercher parmi toutes les solutions possibles, des solutions approchées \tilde{u} ayant certaines caractéristiques raisonnables et telles que l'erreur d'approximation $\|\tilde{u} - u\|$ soit comparable à la norme du bruit de mesure $\|n\| = \sigma$. Afin d'assurer des telles solutions approximées, plus ou moins raisonnables Strohmer et Rauth [Rau98] ont proposé de munir l'algorithme ACT d'une technique de régularisation semblable à celle de Tikhonov [Tik63] (évoquée précédemment dans la section 2.2.2), afin de choisir parmi toutes les solutions possibles celle qui est la plus cohérente avec la vitesse de décroissance connue $(1 + |\omega|)^{-p}$ des coefficients de Fourier de u .

Comme j'ai constaté dans [P14,P26 chap 2] cette technique produit le résultat attendu, mais pénalise très fortement le taux de convergence. En revanche, un algorithme alternatif très simple proposé à l'origine par Bernard Rougé et enrichi avec des idées de Jean-Michel Morel et moi même (que nous avons appelé algorithme de la pseudo-inverse) produit dans des conditions typiques en imagerie satellitaire des solutions approximées \tilde{u} avec un erreur d'approximation $\|\tilde{u} - u\|$ comparable à celui de l'algorithme ACT, mais avec un effort de calcul entre 3 et 10 fois moins important. Voilà pourquoi, pendant ma thèse, j'ai dédié une grande partie de mon temps à comprendre les raisons pour lesquelles cet algorithme converge aussi vite dans ces conditions, même s'il diverge peu après avoir atteint le point d'approximation optimal. Des réponses partielles à cette question se trouvent dans la section 4 et l'appendice A de [P14,P26 chap 2].

2.3.2 Vers une restauration basée sur un modèle linéaire complet de formation d'image

Enfin, pour résoudre le problème complet de la restauration d'images dans une situation assez réaliste, telle que je l'ai énoncé dans le modèle de formation d'image (2.28), tenant compte aussi bien de l'échantillonnage irrégulier que des dégradations subies par l'image physique telles que le flou, le bruit, l'aliasage et la perte de résolution, j'ai proposé d'estimer le paysage f en

minimisant la fonctionnelle suivante

$$\tilde{f} = \min_f \frac{1}{2} \|w \cdot \Delta_\Lambda \cdot (f * h * p) - g\|^2 + \lambda J(f). \quad (2.34)$$

Décortiquons les composantes de cette équation :

- $J(u) = \int |\nabla f|$ est la variation totale de f agissant comme terme de régularisation et d'extrapolation spectrale.
- w est un vecteur de pondération spatiale faisant fonction de déconditionnement. Typiquement $w(\lambda_k)$ est la surface de la cellule de Voronoï centrée sur l'échantillon $\lambda_k \in \Lambda$ par rapport à l'ensemble d'échantillonnage Λ . Une telle technique était déjà utilisée dans [GS00] afin d'éviter que des régions trop densément échantillonnées l'emportent par rapport à d'autres régions moins densément échantillonnées.
- p est un projecteur spectral (e.g. $\hat{p} = \mathbb{1}_R$) sur une région spectrale R qui dépend aussi bien de la FTM \hat{h} que de la géométrie Λ de l'échantillonnage.

L'idée ici est de combiner les conclusions des chapitres 1 et 2 de ma thèse [P26] et les unifier en fournissant un algorithme qui puisse à la fois déconvoluer, débruiter, désaliasier et extrapoler le spectre d'une image irrégulièrement échantillonnée ou échantillonnée sur un réseau régulier perturbé.

Nous savons, en effet (voir section 2.3.1), que la convergence de l'algorithme ACT n'est assurée que si u (dans ce cas $u = f * h * p$) est à support spectral strictement plus petit qu'une cellule réciproque D^* associée à la densité d'échantillonnage (par exemple $D^* = [-\pi, \pi]^2$). Pour l'assurer dans un cas plus réaliste une régularisation de Tikhonov était proposée, qui pénalisée elle aussi très fortement la convergence. Ici nous nous débarrassons de cette pénalisation en remplaçant la régularisation de Tikhonov par la variation totale, ce qui a deux avantages additionnelles : (i) nous permettons l'algorithme ACT, sans régularisation, de faire sa restauration sur une cellule R dans laquelle la convergence rapide est assurée ; (ii) en dehors de cette cellule R le spectre est extrapolé par minimisation de la variation totale, ce qui permet d'augmenter visuellement la résolution, tout en évitant les phénomènes de *ringing* associés à une coupure spectrale brutale dans R .

Quant au choix optimal de R , une indication préliminaire est fournie par les résultats décrits dans la section 2.2.3. Dans le cas de l'échantillonnage perturbé, avec une perturbation de fréquence maximale π/T cette cellule optimale devra être érodée approximativement d'une largeur π/T afin d'améliorer la convergence de l'algorithme ACT.

Je travaille depuis 2003 en collaboration avec Alter et Caselles [P24], puis avec Caselles, Haro et Rougé [P3] sur le développement d'un algorithme permettant de résoudre ce problème d'optimisation d'une manière performante. Il combine les idées de l'algorithme ACT avec une généralisation proposée par Caselles d'un algorithme exact de minimisation de la variation totale proposé par Chambolle [Cha02], et la technique décrite dans [BCRS03] pour adapter localement la valeur du multiplicateur de Lagrange λ . Même s'il consiste de quatre boucles imbriquées, la convergence dans les boucles internes est très rapide ne nécessitant en règle général plus qu'une dizaine d'itérations.

Les résultats obtenus à ce jour sont très encourageants (voir figure 2.7). Il s'agit à ma connaissance du premier essai de régulariser un problème de restauration d'échantillonnage irrégulier par la variation totale au lieu d'utiliser des techniques linéaires plus courantes. L'expérience

s'est montré utile car les erreurs d'approximation mesurées sont en général plus petits, et nous avons en plus la possibilité d'extrapoler le spectre.

Les expériences faites jusqu'à présent considèrent toutes les distorsions (échantillonnage perturbé, flou, bruit) prévues dans le modèle complet de formation d'image (2.28) sauf l'aliasage fréquentiel. Ce cas est aussi pris en compte par l'algorithme que nous avons développé, et il ne nécessite que des réglages concernant la cellule R sur laquelle le projecteur spectral p doit s'appliquer. Ces réglages feront l'objet de nos futures recherches et conduiront (on l'espère) à une méthode complète de restauration d'images.

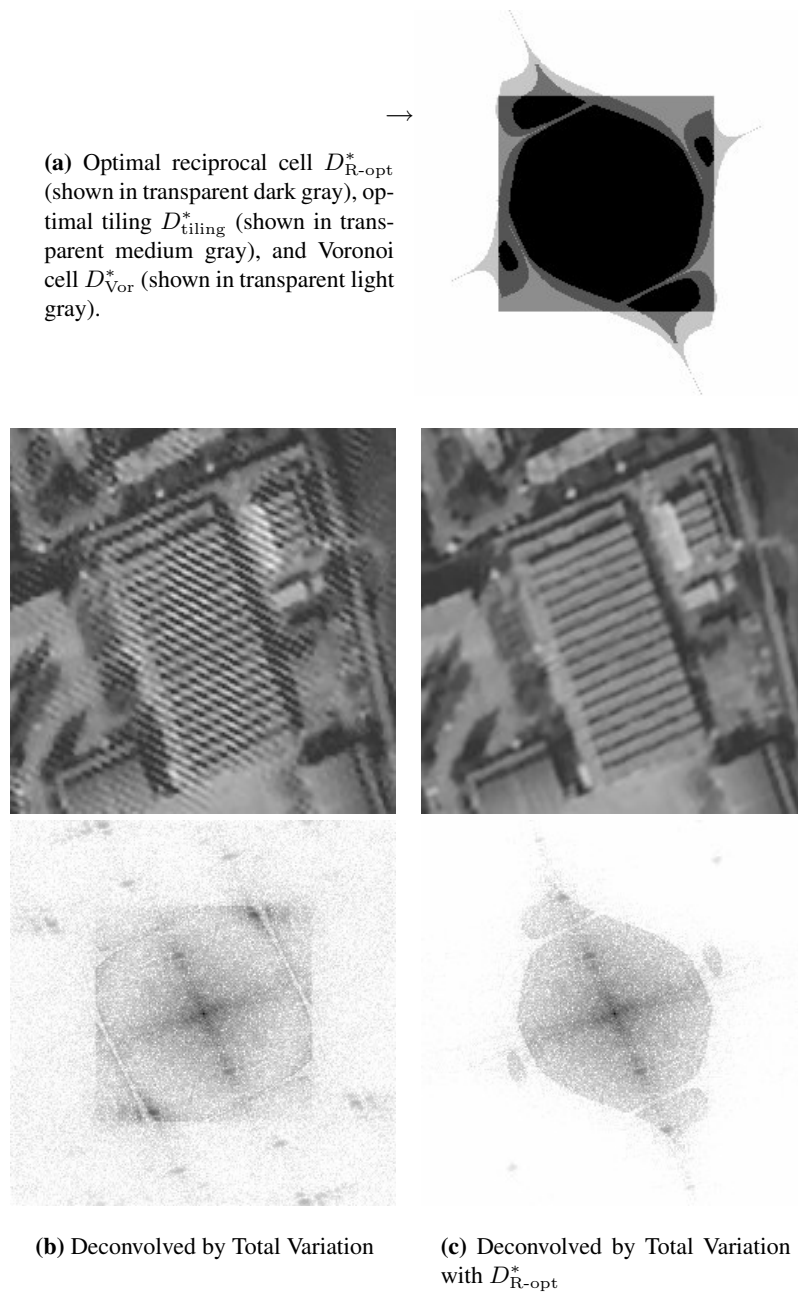


FIG. 2.6:

(a) paysage original f (b) image irrégulièrement échantillonnée g (c) restauration et zoom \tilde{f}

FIG. 2.7: Restauration et zoom d'une image floue, bruitée et irrégulièrement échantillonnée par minimisation exacte de la variation totale combinée avec l'algorithme ACT pour l'attache aux données irréguliers. L'image numérique g est simulée à partir du paysage f en utilisant le modèle (2.28). La perturbation ε du réseau régulier \mathbb{Z}^2 est un bruit coloré d'écart type $A = 1$ et fréquence maximale $\pi/20$. Le bruit blanc n a un écart type de $\sigma = 1$ niveau de gris (à comparer avec la plage de 256 niveaux de gris de l'image), et la FTM \hat{h} est celle d'Hipermode dans $R_{\mathbb{Z}^2}$ et nulle en dehors. Cette expérience ne simule donc pas l'aliasage fréquentiel.

Chapitre 3

Interpolation non shannonienne : application aux modèles de terrain

Quand il s'agit de remplir un trou compact de grande taille Ω dans une image ou une fonction connue par ailleurs ($\mathbb{R}^2 \setminus \Omega$) avec une résolution très fine par rapport à la taille du trou, le modèle de Shannon basé sur l'analyse harmonique, et même celui basé sur l'analyse non harmonique est moins utile. Simplement parce que, en termes d'échantillonnage irrégulier, l'écart maximal ou *maximal gap* de l'ensemble d'échantillonnage (ici le diamètre du trou) est largement plus grand que le pas d'échantillonnage régulier sur $\mathbb{R}^2 \setminus \Omega$. Dans ces conditions le modèle shannonien décrit dans le chapitre précédent est contraint : (i) soit d'approximer la solution par une fonction à bande limitée, trop régulière pour pouvoir satisfaire les conditions d'interpolation sur $\mathbb{R}^2 \setminus \Omega$ et particulièrement sur le bord $\partial\Omega$; (ii) soit de permettre une bande plus large et risquer une solution très oscillante à l'intérieur de Ω ainsi qu'une grande sensibilité au bruit. Dans [FS94] Feichtinger et Strohmer proposent une application de l'algorithme ACT pour remplir des trous (des défauts) dans des images. Bien que très surprenants ces résultats ne peuvent être obtenus qu'avec des images très exactement à bande limitée et quasiment sans bruit, des conditions qui sont peu réalistes en pratique. En dehors de ces conditions nous devons espérer, en fonction des résultats du chapitre précédent, une grande instabilité de l'interpolateur et en particulier des solutions très oscillantes.

Dans des conditions plus réalistes il me paraît plus sensé de partir de la théorie axiomatique de l'interpolation proposée par Caselles, Morel et Sbert [CMS98], qui ont montré que dans un domaine plan Ω un opérateur d'interpolation monotone, stable, régulier et invariant par isométrie était forcément donné par la solution d'une équation différentielle

$$\begin{cases} G(u_{\xi\xi}, u_{\eta\eta}, u_{\eta\xi}) = 0, & \text{dans } \Omega \\ u = \varphi, & \text{sur } \partial\Omega \end{cases} \quad (3.1)$$

où G doit satisfaire des conditions d'ellipticité et nous avons utilisé la notation $u_{\xi\xi} = D^2u \left(\frac{Du}{|Du|}, \frac{Du}{|Du|} \right)$ pour la dérivée seconde de u dans la direction du gradient, et la notation $u_{\eta\eta} = D^2u \left(\frac{Du^\perp}{|Du^\perp|}, \frac{Du^\perp}{|Du^\perp|} \right)$ et $u_{\eta\xi} = D^2u \left(\frac{Du^\perp}{|Du^\perp|}, \frac{Du}{|Du|} \right)$ pour la dérivée seconde de u dans la direction des lignes de niveau,

et la dérivée croisée respectivement. On reconnaîtra en particulier le laplacien (pour $G(a, b, c) = a + b$). Deux autres équations duales l'une de l'autre ont suscité l'intérêt des communautés mathématique et de vision par ordinateur.

La première est l'équation $D^2u \left(\frac{Du^\perp}{|Du|}, \frac{Du^\perp}{|Du|} \right) = 0$ qui signifie que la courbure des lignes de niveau est nulle à l'intérieur de Ω , autrement dit les lignes de niveau de l'interpolateur sont des droites. Comme cette équation peut avoir plusieurs solutions, Masnou et Morel [Mas98] ont proposé d'en choisir la meilleure en introduisant un critère variationnel. La méthode résultante de désocclusion (*disocclusion* en anglais) fournit des résultats très convaincants en propageant les bords de l'image vers l'intérieur de l'occlusion Ω d'une façon assez convaincante à l'oeil. Ce travail a été le précurseur de nombreuses recherches qui –visant à permettre de propager les bords de l'image pas seulement par des lignes droites mais aussi par des courbes régulières (comme le font les restaurateurs de peintures dans les musées)– ont substitué l'EDP d'ordre deux de la méthode de Masnou et Morel par des EDPs d'ordre 3 [BSCB00] ou 4 [BBC⁺01].

Dans l'application à laquelle je me suis intéressé (à nouveau motivé par des problèmes posés par le CNES) les caractéristiques géométriques principales à préserver ne sont pas –comme dans le cas des images naturelles– les bords contrastés, c'est à dire des courbes plus ou moins régulières le long desquelles $|Du|$ est maximal (ces courbes coïncident localement avec les lignes de niveau de l'image). Le problème qui nous a été proposé par le CNES en 2000 est au contraire l'interpolation de modèles de terrain qui ont très rarement des bords contrastés. Dans ce cas les caractéristiques importantes qui nous permettent de reconnaître le terrain sont les sommets, et les arêtes, *i.e.* des courbes le long desquelles la courbure principale $|D^2u(p, p)|$ est maximale tandis que $|D^2u(q, q)|$ est minimale, p et q étant les vecteurs propres unitaires maximal et minimal respectivement (en valeur absolu) du Hessien D^2u . Dans ces conditions une extension naturelle de la méthode de désocclusion consisterait à choisir l'interpolateur u en minimisant la fonctionnelle

$$\int_{\Omega} |D^2u(q(u), q(u))(x)|^p dx \quad (3.2)$$

ou en résolvant l'équation différentielle

$$D^4u(q(u), q(u), q(u), q(u))(x) = 0, \quad \text{for } x \in \Omega \quad (3.3)$$

en respectant dans les deux cas au minimum les conditions de bord

$$\begin{aligned} u(x) &= \varphi(x), & \text{for } x \in \partial\Omega \\ \frac{\partial u}{\partial n}(x) &= \psi(x) & \text{for } x \in \partial\Omega. \end{aligned} \quad (3.4)$$

comme je le propose dans [P26 chap 3]. Je n'ai pas réussi, pendant le temps assez court dédié à ce projet, à trouver les outils mathématiques permettant d'étudier l'existence et unicité des solutions d'une telle équation différentielle de quatrième ordre, et les outils numériques permettant de calculer des solutions sur ordinateur.

Dans ma collaboration avec Cao, Gousseau et Rougé [P6] sur l'interpolation de modèles de terrain nous nous sommes contentés d'un objectif plus modeste consistant à évaluer dans quelle

mesure les interpolateurs bien connus s'adaptent bien à ce type de données. Pour rendre l'étude plus systématique nous avons proposé une extension de l'axiomatique dans [CMS98]. Cette extension s'est avérée nécessaire car, au contraire de ce qui arrive sur les images naturelles où l'oeil n'est sensible qu'aux discontinuités sur u , dans les modèles de terrain l'oeil est aussi assez sensible aux discontinuités du gradient Du . C'est normal, car une discontinuité du gradient peut représenter une arête ou un sommet, *i.e.* les informations géométriques primordiales dans un modèle de terrain. Dans une image naturelle la géométrie se trouve sur les discontinuités de u (les frontières entre deux objets) et les discontinuités de Du passent donc inaperçues. Il est donc important d'imposer à un interpolateur de modèles de terrain pas seulement la continuité C^0 mais aussi la continuité C^1 à travers la frontière de l'occlusion $\partial\Omega$. Ceci nous oblige à relaxer la condition de régularité (afin d'inclure des EDPs de quatrième ordre permettant d'interpoler u et Du sur le bord), et aussi à adapter la condition de monotonie (ou principe de comparaison), afin d'assurer que l'interpolateur respecte l'ordre des deux types de condition aux bords.

Nous n'avons pas trouvé dans la littérature des méthodes d'interpolation satisfaisant toutes les propriétés requises par la nouvelle axiomatique. Néanmoins l'interpolateur biharmonique¹ qui consiste à annuler le bilaplacien à l'intérieur de l'occlusion

$$\Delta^2 u(x) = 0, \quad \text{si } x \in \Omega \quad (3.5)$$

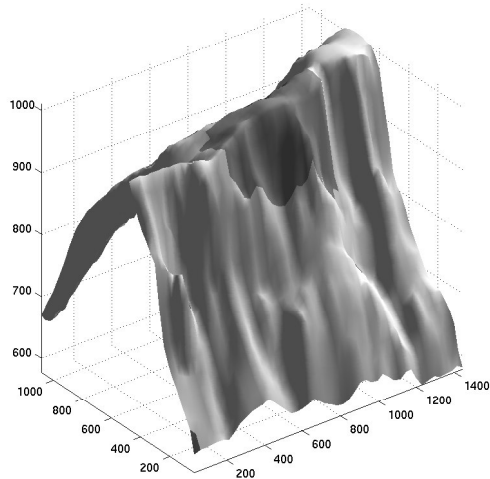
sous les mêmes conditions aux bords (3.4) remplit toutes les conditions sauf le principe de maximum, dont la conclusion reste plus subtile. En fait Boggio a démontré qu'il est vrai quand l'occlusion Ω est circulaire, ce qui a conduit Boggio et Hadamard [Bog01, Had08] à conjecturer qu'il était vrai en général. Néanmoins 40 ans plus tard un certain nombre de contre-exemples dus à Garabedian, Coffman et Duffin ont falsifié cette conjecture [Cof82, Gar51, 86]. Néanmoins il a été montré plus récemment que le principe de maximum reste vrai pour l'interpolateur biharmonique quand le domaine est un cercle légèrement perturbé [GS96]. La question du principe de maximum est d'importance pratique car un interpolateur qui ne le satisfait pas peut produire des solutions très oscillantes qui ne se déduisent pas des conditions aux bords. Tel est le cas du Krigeage [Mat71, Cre93, JH78] quand le paramètre de régularité α est plus grand que 2 (le cas $\alpha = 2$ correspond exactement à l'interpolateur biharmonique), comme nous avons constaté avec Gousseau dans [P6].

méthode	exemple a	exemple b
AMLE	52.88	25.21
biharmonique	16.72	24.84
AMLE + biharmonique	13.98	21.87

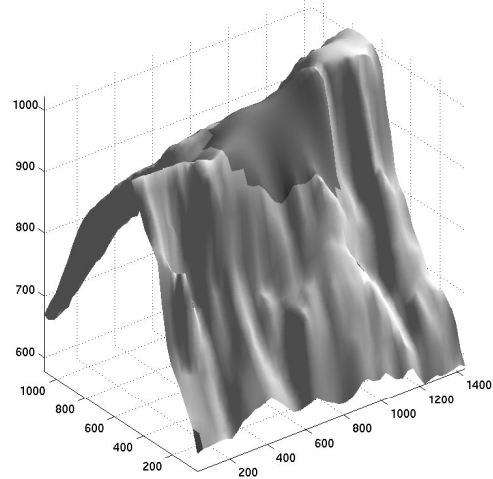
TAB. 3.1: Distances L^2 entre l'interpolateur et la vérité terrain sur l'occlusion pour les exemples a et b, (tous les valeurs sont exprimés en mètres)

Dans l'absence d'un interpolateur satisfaisant toutes les propriétés requises, nous avons retenu le biharmonique, car il est celui qui les satisfait au mieux, et nous avons comparé les

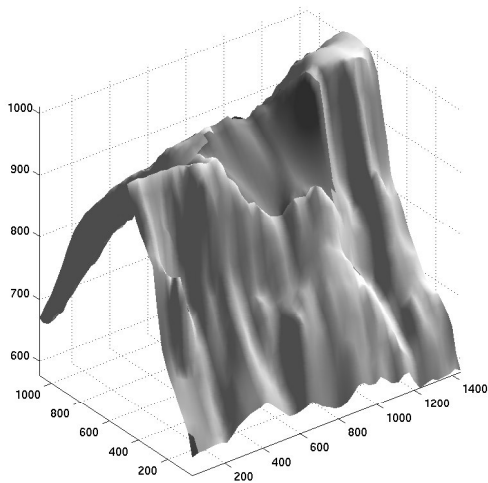
¹aussi appelé *thin-plate* (plaque-mince) car l'équation biharmonique est l'Euler-Lagrange associé à la minimisation de l'énergie de flexion d'une plaque $\int_{\Omega} |\Delta u|^2$.



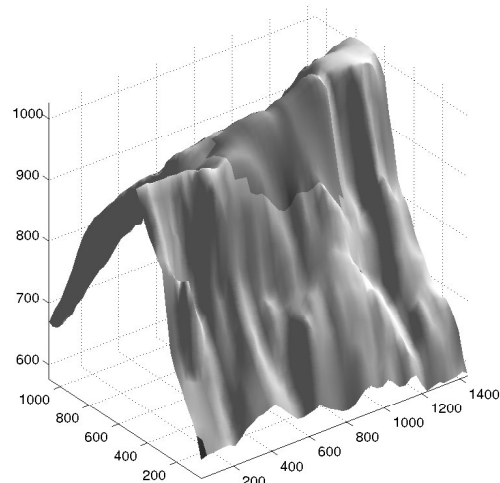
(a) MNE original (occlusion peinte en gris plus sombre)



(b) interpolation biharmonique (krigeage avec $\alpha = 2$)

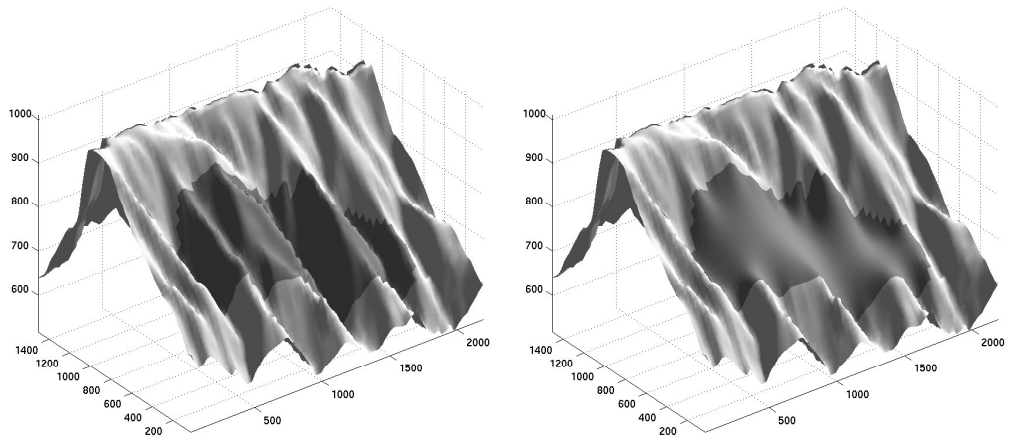


(c) interpolation AMLE



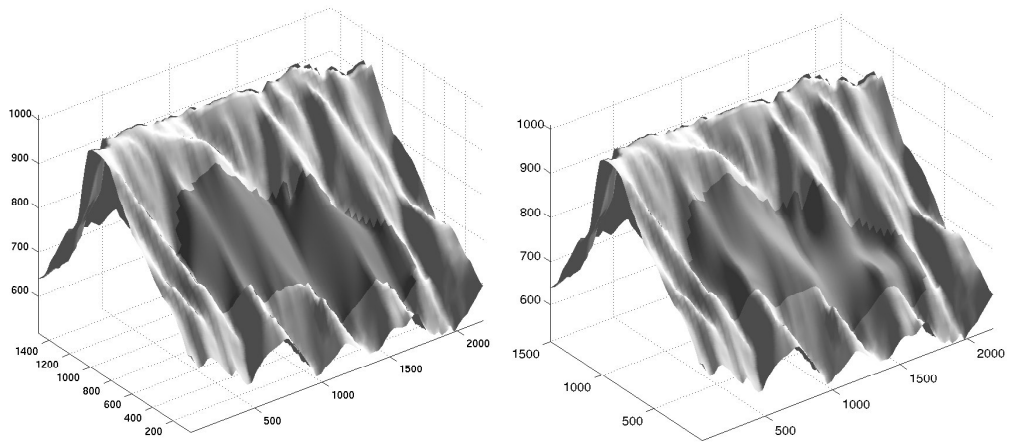
(d) interpolation AMLE + biharmonique

FIG. 3.1: Exemple a : Interpolation d'un modèle de terrain dans une occlusion contenant un sommet



(a) MNE original (occlusion peinte en gris plus sombre)

(b) interpolation biharmonique (krigeage avec $\alpha = 2$)



(c) interpolation AMLE

(d) interpolation AMLE + biharmonique

FIG. 3.2: Exemple b : Interpolation d'un modèle de terrain dans une occlusion contenant un sommet

résultats avec l'AMLE (voir figures 3.1, 3.2). Bien que l'AMLE n'arrive pas à reconstruire un sommet (car il n'interpole pas les pentes), il fait qualitativement mieux que le biharmonique sur les arêtes. Ceci est dû au fait qu'il est diffusif seulement dans la direction du gradient (qui coïncide souvent avec la direction des arêtes), tandis que la diffusion biharmonique est isotrope. Afin de combiner les avantages des deux j'ai proposé dans [P26 chap 3] une façon de les combiner, en sorte que l'interpolateur soit C^1 aux bords (grâce au biharmonique) et préserve mieux mes arêtes (grâce à l'AMLE). La combinaison est dans la plupart des cas étudiés plus satisfaisante aussi bien du point de vue qualitatif que quantitatif (voir le tableau 3.1). Malgré les résultats encourageants il est clair qu'un vaste champ de recherche reste ouvert si nous voulons trouver un bon interpolateur de MNE : aussi bien du point de vue théorique (donner une caractérisation des interpolateurs permis par la nouvelle axiomatique, et dans quelle mesure il existe des solutions uniques pour ces interpolateurs) que numérique (trouver des méthodes performants et stables pour leur calcul).

Enfin ma collaboration avec Cao a abouti à une meilleure compréhension des relations entre l'AMLE et des méthodes d'interpolation de MNE par des transformées distance géodésiques pondérées [Soi91, Soi94]. Une indication préliminaire (avec certaines erreurs) a été publiée dans [P6], mais corrigée et étendue après [P26 chap 3]. En particulier je propose un algorithme rapide (basé sur les transformées distance) pour le calcul de l'AMLE dans le cas où l'occlusion est une région délimité entre deux lignes de niveau.

Chapitre 4

Détection *a contrario* de structures géométriques

Les travaux présentés dans ce chapitre suivent de près le programme de recherche initié à la fin des années 90 par Desolneux, Moisan et Morel [[Des00](#), [DMM00](#), [DMM01](#), [DMM02b](#)] dans le but de développer une théorie quantitative ou computationnelle de la perception visuelle basée sur les lois de la Gestalt. Selon la thèse des gestaltistes [[Kan80](#), [Wer23](#)] la perception visuelle est guidée par un processus de groupement de structures géométriques ayant des caractéristiques similaires, tels que la forme, la couleur, etc. Malgré la pertinence de leurs observations qualitatives sur la perception, l'école gestaltique a évité de répondre à la question quantitative de déterminer précisément le seuil au delà duquel une structure géométrique est "noyée dans le bruit" et donc plus visible par notre perception.

C'est justement ce point crucial que Desolneux, Moisan et Morel ont attaqué en proposant d'utiliser le *principe de Helmholtz* pour déterminer de tels seuils. Il consiste à considérer que une structure géométrique déterminée G est perceptuellement significative quand son occurrence est très peu probable dans une image de bruit. Plus précisément G sera ε -significative si l'espérance du nombre d'occurrences de G dans une image de bruit est plus petite que ε . On note cette espérance par $NFA(G)$, qui signifie "nombre de fausses alarmes de G ". Nous voyons bien que cette approche se trouve à l'opposé de l'estimation bayésienne : elle n'utilise pas un modèle statistique sophistiqué d'image que la structure cherchée devra maximiser ; au contraire elle utilise un modèle (souvent très simple) de fond ou de bruit que la structure cherchée doit minimiser pour être significative. En partie pour cette raison cette approche a été appelée aussi détection *a contrario*.

Supposons que nous cherchions à détecter une certaine structure G et que nous mesurons sur l'image n caractéristiques (variables aléatoires) indépendantes sous le modèle de fond. Ces mesures peuvent être par exemple les niveaux de gris de l'image, mais elles doivent toutes être pertinentes pour la détection de G . Supposons ensuite que k parmi ces n mesures soient cohérentes avec l'existence de G et que la probabilité que chacune de ces mesures soit cohérente avec G soit égale à p . Alors nous pouvons montrer assez facilement que le nombre de fausses alarmes

de G est simplement calculé par le produit d'une binomiale

$$NFA(G) = N_T B(p, k, n) \quad (4.1)$$

par le nombre total N_T de structures G possibles dans l'image. Pour fixer les idées, dans le cas de la détection d'alignements, G est un segment de ligne droite, et les caractéristiques mesurées sont les orientations du gradient de l'image. Ces orientations sont cohérentes avec la présence du segment A lorsque elles sont orthogonales au segment (à une précision $p\pi$), et l mesures sont pertinentes pour la détection de A , l étant le nombre de pixels appartenant au segment. Le NFA est donc dans ce cas

$$NFA(A) = N_T B(p, \frac{k}{2}, \frac{l}{2}) \quad (4.2)$$

où le facteur 2 est une constante de normalisation utilisée pour tenir compte du fait que les mesures d'orientation du gradient de deux pixels voisins ne sont pas indépendantes.

Enfin si deux structures significatives G et G' sont très proches le plus probable est que l'une des deux soit masquée par l'autre, ce qui conduit à ne considérer que les gestalts *maximales* G , dans le sens qu'il ne contiennent ni sont contenues dans aucune autre gestalt G plus significative.

4.1 Détection de points de fuite

Afin de mesurer –dans une seule image d'une scène– la taille des objets qui y sont présents il est très utile de trouver les point de fuite, surtout si l'on connaît certains éléments de la géométrie de la scène. Tel est le but d'une partie du logiciel Investigator [SJ95] et du travail de Criminisi [CRZ00] qui a motivé mon travail de détection de points de fuite [P5]. Une version plus étendue existe dans ma thèse [P26, chap. 4], et une version préliminaire a été présentée à Oberwolfach [P27] ainsi que comme preprint du CMLA (CMLA2001-24).

Dans ce cas nous commençons par détecter tous les n segments maximaux significatifs de l'image en utilisant une distribution uniforme des orientations du gradient comme modèle de fond. Ensuite nous prolongeons ces segments par des droites, et dans une application typique de l'approche hiérarchique de l'école gestaltique, nous utilisons un nouveau modèle *a contrario* pour les droites ainsi obtenus. Le plus neutre est d'utiliser une distribution aléatoire uniforme dans le plan, avec pour seule contrainte qu'elles doivent passer par le domaine Ω de l'image. Alors nous nous demandons si la convergence de k parmi ces n droites dans une région convexe V du plan peut arriver par hasard. L'espérance du nombre d'occurrences d'un tel événement est

$$NFA(V) = N_T B(p, k, n) \quad (4.3)$$

où p est la probabilité qu'une droite aléatoire donnée passe par V (sachant qu'elle passe aussi par Ω). Le calcul de cette probabilité est une belle application des résultats de la géométrie intégrale [San76] :

$$p = \frac{\text{Per}(V)}{\text{Per}(\Omega)} \text{ si } V \subseteq \Omega. \quad (4.4)$$

(Le cas où $V \not\subseteq \Omega$ peut aussi se calculer en fonction des périmètres intérieurs et extérieurs aux deux convexes).

Afin de considérer toutes les régions de fuite possibles, il faut prendre un recouvrement du plan \mathbb{R}^2 par des régions de fuite V ayant toutes la même probabilité p de rencontrer une droite qui passe par l'image Ω . Le nombre N_T de régions nécessaires pour recouvrir le plan donne alors la valeur du nombre de tests dans l'équation (4.3).

Ce que je voudrais souligner de cette méthode est qu'elle es la première à donner les points de fuite significatifs d'une image *et rien d'autre*, et ceci sans aucune connaissance a priori sur les paramètres de calibration de la caméra, ou les relations (par exemple d'orthogonalité) entre les points de fuite d'une image.

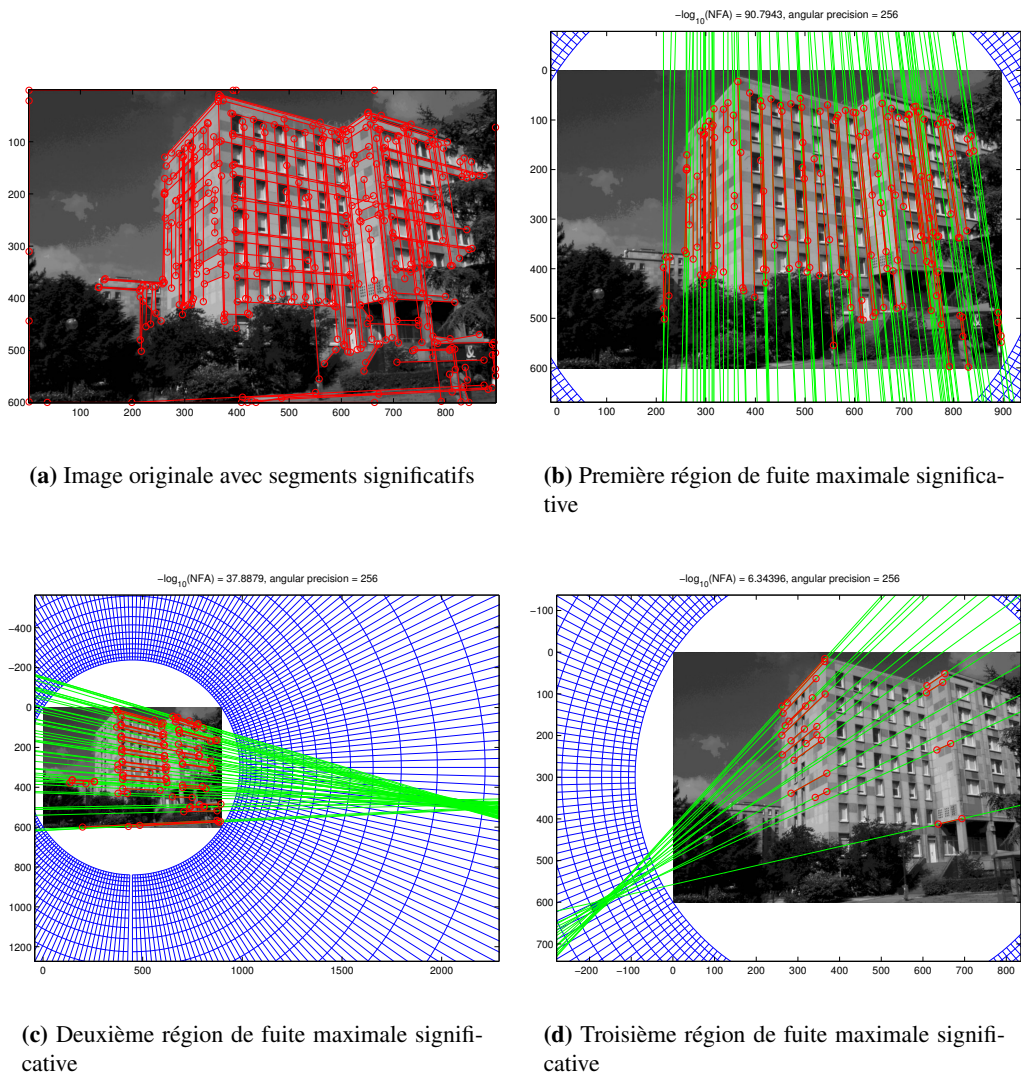


FIG. 4.1: Une image et les seuls trois points de fuite maximaux significatifs détectés.

4.2 Détection de rayures dans les films¹

Lors de mon séjour en 2003 à l'UPF (Barcelone) j'ai été amené à travailler avec Machiel Kolstein et Vicent Caselles sur le projet SpeedFX de postproduction numérique de films. Comme une partie de ce projet, on nous a demandé de détecter automatiquement les rayures verticales présentes dans beaucoup de vieux films, afin de les restaurer après. Nous avons vite réalisé que bien qu'il existe des méthodes (bayésiennes) bien performantes pour l'estimation fine des paramètres des rayures, la détection initiale de ces dernières reste un problème avec des seuils difficiles à régler, ce qui nous a amenés à éviter ce problème par des méthodes *a contrario*. La méthode proposée dans [Kok96], par exemple s'adapte bien à ce but. Un point x appartenant à un candidat R à être une rayure claire est caractérisé par un forte contraste de sa valeur de gris $u(x)$

$$C(x) = u(x) - \max(u_G(x), u_D(x)) \quad (4.5)$$

par rapport aux deux cotés de la rayure ($u_G(x)$ est une moyenne de u prise sur quelques points à gauche de la rayure, et même opération pour $u_D(x)$ à droite). Supposons que parmi les n pixels appartenant à la rayure ($x \in R$) il y a k d'entre eux qui ont un contraste $C(x) > \lambda$ au dessus d'un seuil fixé à l'avance. L'espérance du nombre d'occurrences d'un tel événement est

$$NFA(R) = N_TB(p, \frac{k}{2}, \frac{n}{2}) \quad (4.6)$$

où $p = P[C(x) > \lambda]$ est la probabilité sous le modèle *a contrario* d'un point assez contrasté. Ici le modèle *a contrario* consiste comme d'habitude à supposer les pixels de u indépendants et avec une loi égale à la loi empirique globale de u . Le facteur 2 joue le même rôle que dans le cas des alignements. Il donne une distance minimale entre deux pixels pour qu'ils soient indépendants. Deux points voisins ne le sont pas à cause du flou de l'image avant échantillonnage.

Les expériences faites sur des vrais films numérisés à très haute résolution de l'ordre de 2000x2000 pixels, nous ont conduit à des conclusions assez intéressantes : ²

1. Le facteur de normalisation 2 est crucial pour avoir un seuil de détection en correspondance avec la perception. Un facteur de 1 engendre une très forte sur-détection (trop de faux positifs), tandis qu'un facteur 4 a pour résultat une sous-détection. Je reviendrai sur ce point dans la section 4.4
2. La détection à la résolution originale engendre elle aussi un très grand nombre de "fausses détections", que ne sont plus détectées si l'on limite la longueur des rayures à environ 500 pixels par une réduction d'échelle préalable. Je pense que ce phénomène peut être dû au fait que l'oeil, composé d'une zone fovéale relativement petite, n'est pas capable de parcourir d'un seul coup toute une ligne de 2000x1 pixel à la résolution la plus fine possible. Il doit choisir entre regarder un détail plus petit à la résolution la plus fine, ou regarder l'ensemble à une résolution plus grossière. Il serait intéressant de confirmer cette hypothèse par des expériences psychophysiques.

¹Ce travail n'a pas conduit à une publication mais seulement au développement d'un logiciel faisant partie du projet SpeedFX, dont le produit "Baselight" qui en a découlé a reçu l'un des [prix européens IST 2004](#).

²Je n'ai malheureusement pas le droit de montrer ici ces grandes images pour des raisons de confidentialité et propriété intellectuelle des films. Pour une expérience sur une petite image, voir la figure 4.2



FIG. 4.2: *Détection de rayures* À gauche l'image d'entrée (de taille 256x256), et à droite la seule rayure détectée par la méthode a contrario. Seulement les rayures assez verticales et assez longues ont été cherchées. (Image obtenue du site web de Anil Kokaram).

3. On aurait pu utiliser la distribution empirique de la variable de contraste C comme modèle de fond, en suivant la même logique que pour les contours significatifs qui utilisent la distribution empirique de $|Du|$. Or il se trouve que dans ce cas, ce choix de modèle de fond conduit à un grand nombre de fausses détections. Pour l'éviter nous avons utilisé la distribution empirique de u comme modèle de fond et en déduit celle de C . Ce manque d'homogénéité est très inquiétante, mais elle pourrait être due au fait que C est trop éloignée de l'image (une sorte de dérivée seconde de u) pour que l'on puisse en déduire un modèle a contrario directement de ses observations. Cette hypothèse nécessite bien sûr des expériences dans des domaines plus variés pour être confirmée, et soulève la question des limites dans le choix d'un modèle a contrario pertinent.

4.3 Détection de taches

Une application biomédicale sur laquelle j'ai été amené à travailler récemment donne encore un exemple extrêmement simple de l'utilisation des modèles *a contrario*. Il s'agit de déterminer tout simplement s'il existe ou pas une tache ronde sur un fond blanc. Cette question relativement facile à répondre pour l'oeil expert et attentif, s'avérait très difficile à obtenir d'une manière automatique. De fait la société qui nous a contacté était insatisfaite par les réponses inconsistantes du logiciel de lecture automatique fourni avec les réactifs d'analyse biochimique. Un raffinement de la méthode que je décris ci-dessous donne par contre une réponse plus consistante et fiable.

Nous supposons *a contrario* que un sur R pixels sont indépendants ($R \approx 2$) et suivent la loi empirique estimée dans une région du fond blanc (ne contenant pas de tache). Étant donné un candidat T à être une tache nous fixons un seuil λ sur les niveaux de gris de l'image et comptons le nombre de pixels $k \in T$ tels que $u(x) < \lambda$ et le nombre total de pixels n de la tache. Alors la

significativité de la tache est donnée par

$$NFA(T) = N_{tests} B(p, \frac{k}{R}, \frac{n}{R}), \quad p = P[u(x) < \lambda] \quad (4.7)$$

et le nombre de tests tient compte de toutes les possibles positions et tailles de la tache, ainsi que de tous les seuils λ possibles. Une adaptation du principe de maximalité permet de n'obtenir que la réponse optimale.

Encore plus que dans le cas précédent un bon choix du facteur R s'avère crucial pour obtenir des résultats corrects.

4.4 Validation expérimentale du principe de Helmholtz dans des images dégradées

Le point de départ est le travail de Desolneux, Moisan and Morel [AD03], dans lequel il est montré par une série d'expérimentations psychophysiques que le modèle *a contrario* et le principe de Helmholtz correspondent bien à la perception visuelle, dans un cas très restreint d'images synthétiques binaires. Dans [FL04] deux élèves de mon cours de gestalt computationnelle en Uruguay ont essayé d'étendre cette vérification dans un cadre plus en accord avec le processus de formation d'images naturelles, en tenant compte du bruit, du flou et de l'aliasage en utilisant des modèles de formation et résolution d'images comme ceux qui ont été décrits dans le chapitre 2. Il est habituel dans la plupart des travaux basés sur l'application du principe de Helmholtz sur les pixels d'une image, d'utiliser une constante de normalisation $R = 2$ afin de tenir compte de la dépendance entre les pixels voisins due au flou. Le but de ce travail est de donner une explication plus systématique à cette constante en la remplaçant par la résolution effective du système d'acquisition d'image, et de vérifier que cette opération est toujours bien en accord avec la perception visuelle. Les expériences, bien que pas trop concluantes encore, n'ont pas encore falsifié la justesse de ce choix. Notons au passage que le choix canonique $R = 2$ correspond assez bien au nombre r_{nom}/r_{eff} de pixels nominaux par pixel effectif pour la plupart des systèmes d'acquisition courants, y compris un scanner.

4.5 Détection de logos

Pour conclure ce chapitre je présente une application plus élaborée du principe de Helmholtz ayant trois étapes hiérarchiques de groupement. Ce travail [P25] a été fait avec Frédéric Guichard et Thierry Cohignac lorsque je travaillais à Poseidon Technologies / VisionIQ dans le cadre de l'action RNRT "Indexation Sémantique des Images sur Internet (ISII)".

Le point de départ est la technique de mise en correspondance *a contrario* d'éléments de forme développée par Musé et Sur [MSCG03, MSCG04]. Ma contribution a été de montrer qu'en cherchant des regroupements assez denses (*clusters*) de ces correspondances dans l'espace des transformations (similitudes ou affines) il est faisable de répondre à la question si oui ou non une forme relativement complexe est présente dans une image. L'application visée, pour laquelle j'ai développé un prototype qui a contribué à montrer la faisabilité de cette technique

était la recherche d'un certain nombre fixe de logos dans une grande base de données d'images (voir figures 4.3 et 4.4). La méthode proposée dans [P25] a été ensuite beaucoup améliorée dans les thèses de Musé et Sur [Sur04, Mus04, CDD⁺04] en remplaçant le découpage dyadique de l'espace de transformations, par d'autres structures hiérarchiques mieux adaptées au problème, ainsi que par une étude plus systématique de la significativité et la maximalité des *clusters*.

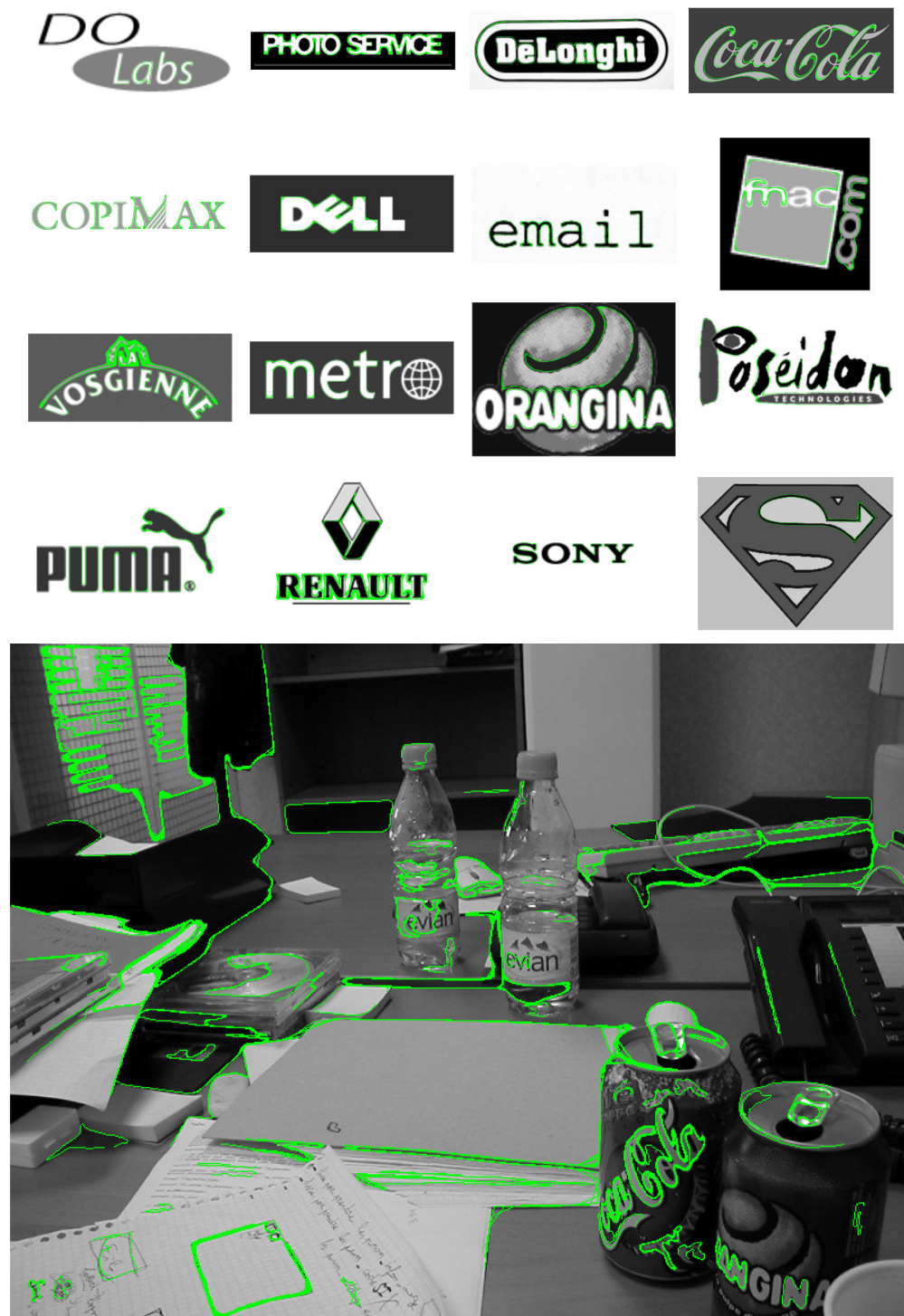


FIG. 4.3: Exemple de détection de logos. Les seize logos cherchés, et une image de requête. Les courbes vertes sont les morceaux de ligne de niveau qui sont codés d'une façon similitude-invariante pour leur mise en correspondance.



FIG. 4.4: *Exemple de détection de logos (cont.)* Le résultat de la mise en correspondance de tous les éléments de forme suivie d'un groupement par ressemblance de la transformation de similitude. Seul le logo de Coca-Cola donne un *cluster* de transformations significatif.

Chapitre 5

Photogramétrie à faible B/H en milieu urbain

Il est bien connu qu'une *disparité* ε dans la position à laquelle se trouve la projection (pin-hole ou push-broom) d'un objet sur deux images prises de deux points de vue différents donne une indication de la distance h à laquelle se trouve l'objet par rapport aux points focaux des caméras. C'est le principe de base de la vision stéréoscopique. Cette relation qui dans un cas général implique l'estimation d'une bonne dizaine de paramètres des transformations projectives correspondantes, se réduit dans le cas de l'imagerie satellitaire (ou aérienne d'une très grande hauteur) à la simple relation

$$\varepsilon = \frac{B}{H}h \quad (5.1)$$

où B est la distance (horizontale) entre les deux prises de vue, et H la hauteur du satellite. Une mesure de la disparité $\varepsilon(x)$ entre un point x de l'image de référence et son homologue $x' = x + \varepsilon(x)$ (correspondant à la projection d'un même point dans l'espace) dans l'image secondaire permet donc de retrouver la hauteur $h(x)$ du terrain pour tout point x .

Plus le rapport B/H est petit, plus de précision est requise sur la disparité estimée ε afin d'obtenir une même précision objectif sur la hauteur h , ce qui a conduit naturellement à l'utilisation des rapports $B/H \approx 1$ assez grands (correspondant à un angle entre les prises de vue de 45 degrés). Mais ce choix n'est pas sans inconvénient car en pratique ceci implique des images prises à des heures très différentes et avec beaucoup de régions visibles dans l'une qui ne le sont pas dans l'autre (en raison des occlusions), ce qui fait la recherche de points homologues d'autant plus difficile à modéliser mathématiquement et à résoudre numériquement.

Le groupe du CNES autour de Bernard Rougé a décidé, au contraire, de se débarrasser du problème des occlusions et de la variabilité des images, en utilisant des rapports B/H très faibles ce qui implique, d'une part, que le modèle de correspondance

$$\tilde{u}(x) = u(x + \varepsilon(x)) \quad (5.2)$$

entre l'image primaire u et secondaire \tilde{u} est assez juste (c'est la bonne nouvelle), et d'autre part, qu'une précision raisonnable d'un mètre sur la hauteur h implique des précisions de l'ordre du dixième, voir du centième de pixel dans l'estimation de ε , ce qui représente un vrai défi.

Pour atteindre ces précisions Julie Delon et Bernard Rougé [Del04, DR01] ont fait une analyse mathématique et numérique en profondeur de la méthode classique de corrélation de paires stéréo qui consiste à estimer la disparité au point x_0 comme la valeur du décalage $m(x_0)$ qui minimise la distance L^2

$$\int \varphi_{x_0}(v(\cdot + m) - \tilde{v})^2 \quad (5.3)$$

entre les images normalisées ($v = u/\sqrt{\int u^2 \varphi_{x_0}}$) et localement dans une fenêtre φ_{x_0} centrée en x_0 . En utilisant le modèle (5.2) ils ont trouvé que la mesure m qui maximise la corrélation est en lien avec la vraie disparité ε , au moins au premier ordre, par la relation suivante

$$(\varepsilon d_{x_0}) * \varphi = m\rho \quad (5.4)$$

qu'ils ont appelée *équation fondamentale de la corrélation*, impliquant une *densité de corrélation* d_{x_0} qui est très concentrée sur les points à gradient fort de l'image, et la courbure de corrélation au maximum $\rho = d_{x_0} * \varphi$.

La disparité $m(x_0)$ estimée par corrélation dans un point x_0 est donc une moyenne pondérée des vraies disparités $\varepsilon(x)$ dans un voisinage de x_0 de taille égale au $\text{supp}(\varphi)$. Si cette pondération est le plus concentrée sur un point x_1 de ce voisinage, alors $m(x_0)$ sera en fait probablement plus proche de la vraie disparité au point voisin ($\varepsilon(x_1)$) qu'au point central x_0 . Ce comportement a été appelé *phénomène d'adhérence*, car il a pour effet de dilater les valeurs des disparités proches des bords de l'image (où la pondération d_{x_0} est forte) vers des régions proches où l'image est plus régulière. La solution proposée par Rougé pour corriger ce phénomène d'adhérence consiste à faire une *correction barycentrique* qui consiste à considérer le maximum de corrélation $m(x_0)$ comme une estimation de la disparité $\varepsilon(x_1)$ au barycentre x_1 de la fonction de pondération. L'analyse de Delon et Rougé donne aussi une estimation de la précision à laquelle la hauteur est estimée par corrélation et cette estimation est utilisée pour adapter localement la taille de la fenêtre de corrélation φ à une précision objectif sur la disparité. Dans certains points la précision n'est pas suffisante, même avec une fenêtre très grande, et la disparité doit être interpolée.

La figure 5.2(b) montre le résultat de l'application d'un raffinement de cette approche à la paire d'images stéréo de la figure 5.1(b) et 5.1(b). On peut observer que dans un milieu urbain comme celui-ci la correction barycentrique n'arrive pas à corriger complètement le phénomène d'adhérence sur les bords des bâtiments, et la géométrie très particulière du paysage urbain, essentiellement affine par régions, n'est pas très visible.

Le projet PDT dont je suis responsable scientifique en Uruguay a pour but d'améliorer ces résultats dans le cas urbain, en utilisant trois approches différentes : une approche variationnelle (section 5.1), une approche d'estimation affine par régions (section 5.2) et une approche basée sur la restauration (voir [P3], et la fin du chapitre 2). Cette dernière est nécessaire afin de corriger les distorsions de l'image (flou, bruit, échantillonnage perturbé, aliasage), augmenter la résolution et obtenir ainsi une plus grande précision dans les mesures de disparité. Bien qu'il existe des méthodes de restauration bien performantes, j'ai décidé d'inclure ce dernier point dans le projet pour deux raisons : (i) je ne connais pas des méthodes de restauration permettant de corriger conjointement toutes les distorsions linéaires comme nous visons dans [P3], et (ii) pour faire une estimation précise de la disparité il faut connaître très précisément les ca-



(a) image de référence

(b) image de secondaire

FIG. 5.1: Paire stéréo à faible B/H. Le décalage entre l'image de référence (a) et l'image secondaire (b) n'est pas facilement visible sur les images, car trop petit (de l'ordre d'un tiers de pixel).

ractéristiques des image de départ, ce qui est contrôlé le mieux en pratique si nous faisons nous mêmes la restauration.

5.1 Méthode variationnelle

Mon travail en cours avec Gabriele Facciolo et Alvaro Pardo dont [P2] donne des résultats préliminaires vise à substituer la correction barycentrique et l'interpolation des zones inconnues du logiciel MARC développé au CNES par un critère variationnel plus en accord avec la géométrie particulière du paysage urbain.

L'équation fondamentale (5.4) peut être en fait écrite comme une relation linéaire $K\varepsilon = m\rho$ entre la vraie disparité ε et la mesure de corrélation m . L'opérateur linéaire K est très mal conditionné, car quasiment nul dans les points où la courbure de corrélation est très faible. Nous cherchons donc à estimer la disparité en ajoutant à l'équation fondamentale un terme régularisateur bien adapté

$$\min_{\varepsilon} D(\varepsilon) + \lambda R(\varepsilon) \quad (5.5)$$

$$D(\varepsilon) = \int |(\varepsilon d_{x_0}) * \varphi(x_0) - m(x_0)\rho(x_0)|^2 dx_0 \quad (5.6)$$

$$S(\varepsilon) = \int \sqrt{a^2 + |\nabla\varepsilon(x)|^2} dx \quad (5.7)$$

Cette régularisation par surface minimale a été beaucoup utilisée en stéréo et dans des situations plus générales impliquant plusieurs caméras dans une disposition quelconque (voir par exemple les travaux de Faugeras et Keriven [FK98]). Plus récemment des méthodes très performantes de *graph-cuts* ont été développés par Kolmogorov et Boykov [BK04] pour trouver une surface minimale sous contraintes, et appliqués au problème de la stéréo avec succès [SS02]. Mais la plupart de ces méthodes utilisent un terme d'attache aux données ne tenant pas compte du phénomène d'adhérence ou ne permettent de calculer la disparité qu'à des pas entiers (1 pixel), ce qui les rend pas applicables au cas d'un rapport B/H très faible. La nouveauté de la méthode proposée consiste donc à utiliser un terme d'attache aux données $D(\varepsilon)$ tenant compte du phénomène d'adhérence via l'équation fondamentale.

La régularisation par surface minimale, bien qu'utile pour interpoler des régions affines peut avoir une tendance à arrondir les coins et les bords des bâtiments qui n'ont pas d'attache aux données assez forte. Mais l'on sait aussi que les lignes de niveau le plus contrastées de l'image délimitent des régions pouvant avoir des hauteurs très différentes. Afin de favoriser les contours du terrain qui suivent les contours contrastés de l'image nous utilisons la variante suivante du terme de régularisation

$$R(\varepsilon) = \int \sqrt{a^2 + |\nabla \varepsilon|^2} - \int \sqrt{a^2 + |\langle \nabla \varepsilon, z \rangle|^2} \quad (5.8)$$

$$z(x) = \nabla u / \sqrt{b^2 + |\nabla u|^2} \quad (5.9)$$

Le paramètre b agit comme un seuil doux entre les points à gradient faible où l'orientation z du gradient de l'image n'est pas fiable (et donc $z \approx 0$) et les points plus contrastés où $z = \nabla u / |\nabla u|$. Une approche similaire a été utilisée précédemment par Caselles et Morel [BCRV] pour interpoler des images en couleur, sachant que la carte topographique des différents canaux est très semblable.

La figure 5.2(d) montre le résultat d'inverser la formule fondamentale régularisée le terme par surface minimale S , ainsi que par son adaptation R à la carte topographique de l'image. Il est aussi possible de garder le résultat ε_0 correction barycentrique de MARC et appliquer seulement une petite régularisation ($\min_{\varepsilon} \|\varepsilon - \varepsilon_0\| + \lambda R(\varepsilon)$) pour lisser le résultat en préservant les bords entre régions.

L'implémentation numérique de cette approche variationnelle présente certaines difficultés. Celle que nous essayons de surmonter à présent concerne le fait que le pas maximal d'une méthode explicite stable de descente de gradient est fortement limité par le terme D qui a des très différentes échelles, en raison de la forte variabilité de la courbure de corrélation ρ . Cette variabilité, qui est bien utile pour donner un poids plus important aux points fiables (ayant une forte courbure de corrélation) s'avère difficile à résoudre numériquement. Nous envisageons donc de remplacer la pondération par le seuil ρ_{\min} qui correspond à une précision donnée sur la disparité dans la théorie développé par Delon et Rougé :

$$EF(x_0) = \begin{cases} \frac{(\varepsilon d_{x_0}) * \varphi(x_0)}{\rho(x_0)} - m(x_0) & \text{si } \rho > \rho_{\min} \\ 0 & \text{sinon} \end{cases} \quad (5.10)$$

Le terme d'attache aux données s'écrirait dans ce cas $D(\varepsilon) = \int |EF(x)|^2 dx$.

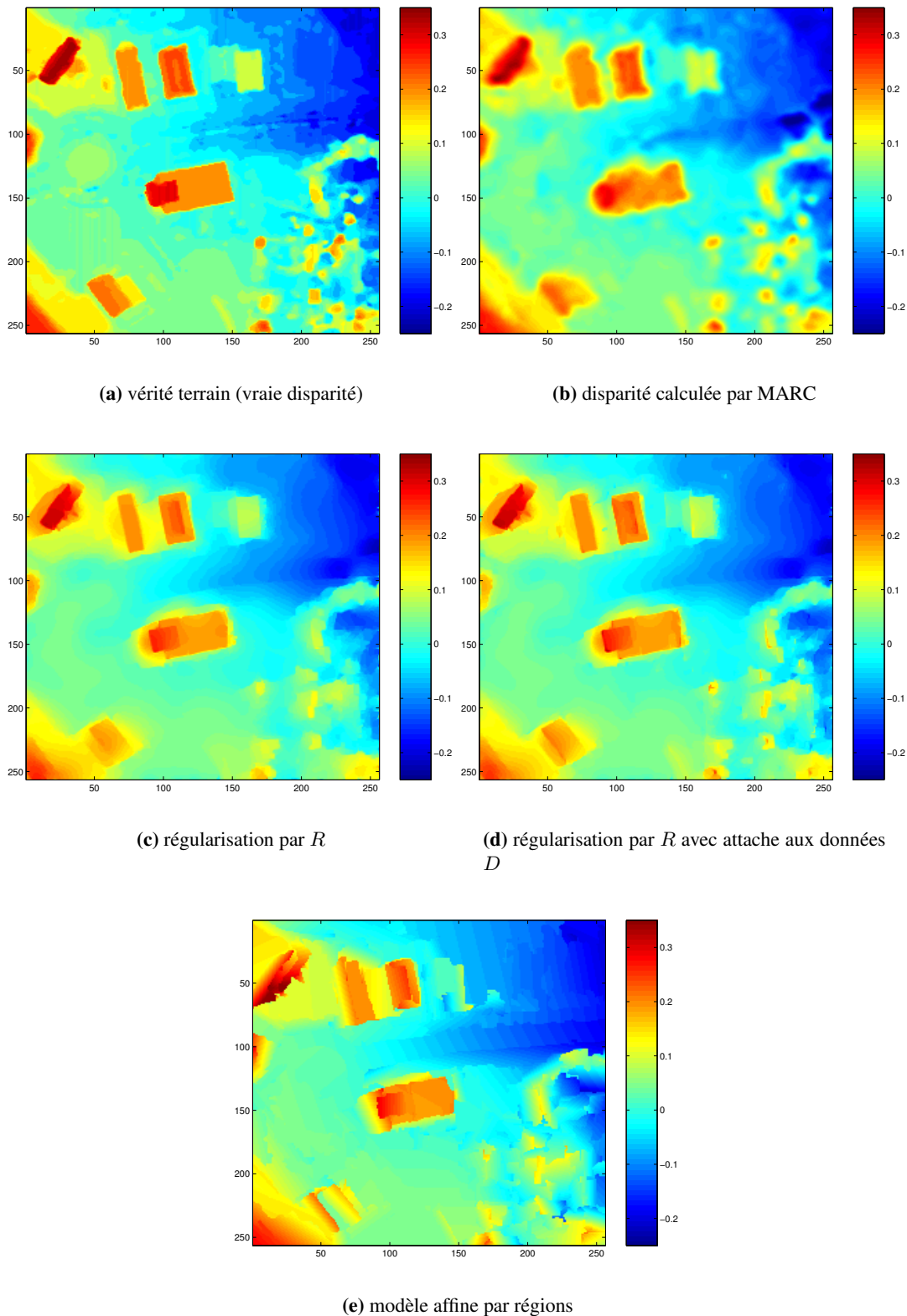


FIG. 5.2: Estimation de la carte de disparité par des méthodes variationnelles et affines (a) Vraie disparité. (b) Résultat ε_0 obtenu par le logiciel MARC qui calcule la corrélation fine avec correction barycentrique. (c) Régularisation simple par minimisation de $R(\varepsilon) + \lambda|\varepsilon - \varepsilon_0|^2$, (d) Régularisation sous la contrainte de l'équation fondamentale (minimisation de $R(\varepsilon) + \lambda D(\varepsilon)$). (e) Modèle affine par régions : la partition initiale en 500 régions obtenue par Mumford Shah, est ensuite modifiée par la méthode de *merging* de régions décrite dans la section 5.2

Encore une voie à explorer consiste à utiliser une méthode multigrille, afin d'accélérer la convergence et mieux gérer les différentes échelles présentes, ainsi que pour des raisons liées au fait que l'équation fondamentale n'est valable qu'au premier ordre pour des petites valeurs de ε .

5.2 Méthodes affines par régions

Une deuxième possibilité que nous sommes en train d'explorer avec Javier Preciozzi, Julie Delon et Bernard Rougé est de substituer le terme de régularisation R de la section précédente par un modèle affine par régions, ce qui est bien en accord avec les caractéristiques de la carte d'élévation urbaine composée principalement de toits de maisons, parfois inclinés. Même en supposant qu'une partition de l'image en régions R_1, \dots, R_n est connue, si nous partons directement des mesures de corrélation m et utilisons l'équation fondamentale pour ajuster un modèle affine par régions

$$\varepsilon(x, y) = \sum_i \mathbb{1}_{R_i} T_i(x, y), \quad \text{avec} \quad T_i(x, y) = a_i x + b_i y + c_i \quad (5.11)$$

à ces mesures, nous obtenons un système d'équations couplées faisant intervenir les paramètres a_i, b_i, c_i des affinités à chaque région (les paramètres de deux régions voisines sont couplés à cause de la convolution par φ). La résolution d'un tel système présente certaines difficultés.

Pour l'instant nous avons suivi une voie plus simple qui consiste à ne pas utiliser directement l'équation fondamentale pour corriger le phénomène d'adhérence. En revanche nous partons de la disparité $\varepsilon_0(x)$ après correction barycentrique sur les points valides (c'est à dire ceux qui ont une courbure de corrélation $\rho(x) > \rho_{\min}$ assez forte). Posée ainsi l'estimation du modèle affine de chaque région est découplée des autres régions. Elle peut être obtenue par une estimation robuste (voir par exemple [FLP01])

$$\min_{a_i, b_i, c_i} \sum_{x \in R_i, \rho(x) > \rho_{\min}} g(T_i(x) - \varepsilon_0(x)) \quad (5.12)$$

où g est une fonction positive avec son minimum en 0 permettant d'écarter des *outliers*. Nous avons préféré cette estimation (en dépit des moindres carrés habituels qui utilisent $g(r) = r^2$), parce que la correction barycentrique ne corrige pas totalement le phénomène d'adhérence, et il s'avère nécessaire de considérer certaines mesures $\varepsilon_0(x)$ comme *outliers*, car elles représentent plutôt la hauteur d'une région voisine.

Le problème principal reste de trouver une bonne partition de l'image en régions R_1, \dots, R_n , dans lesquelles un modèle affine de la disparité soit raisonnable. Dans la section précédente nous avons observé que les bords de l'image donnent une bonne indication, quoique pas suffisante, de l'endroit où peuvent se trouver les discontinuités du terrain. Il est donc censé d'utiliser comme partition initiale une sur-segmentation de l'image (obtenue par exemple par la méthode proposée par Mumford et Shah [MJS89, KLM94]), pour ensuite chercher à joindre des régions dont le modèle affine se ressemble. Une telle approche de *merging* de régions (voir par exemple [Jai95, GS98, KLM94]) nécessite un critère pour décider si une fusion (*merge*) entre deux régions voisines est raisonnable, et pour choisir le meilleur s'il y a plusieurs possibilités. Le

critère utilisé ici est guidé par la minimisation de l'erreur globale moyenne d'approximation

$$\sigma^2(R_1, \dots, R_n) = \sum_i |R_i| \sigma_i^2(R_i) / \sum_i |R_i| \quad (5.13)$$

L'erreur (quadratique) moyen d'approximation sur une région $\sigma_i^2(R_i)$ est calculée par une estimation robuste de la variance de $T_i - \varepsilon_0$, ce que devient infini si le nombre de points valides de l'équation (5.11) n'est pas supérieur au nombre 3 de paramètres. La fusion de ces petites régions est donc forcée au début. Ensuite à chaque itération nous fusionnons le paire de régions (R_i, R_j) faisant décroître le plus l'erreur moyen d'approximation, jusqu'à ce que l'erreur ne diminue plus. Cette approche très simple, malgré certains erreurs dans le choix des fusions, permet une assez bonne reconstruction du terrain (voir figure 5.2(e)).

Cette méthode est très inspirée de celle proposée précédemment par Caselles, Garrido et Igual [CGI05] pour l'estimation de mouvement entre deux *frames* consécutifs d'une séquence vidéo. Ils proposent aussi d'ajuster un modèle afin par régions aux données avec la même partition initiale de Mumford-Shah. En revanche, leur méthode n'utilise pas la corrélation pour trouver des correspondances, mais la similitude des orientations du gradient, donnant, en conséquence, une estimation invariante par changements monotones de contraste g . Leur méthode est donc consistente avec un modèle plus général que celui de l'équation (5.2) à savoir

$$\tilde{u}(x) = g(u(x + \varepsilon(x))) \quad (5.14)$$

Leurs estimations sont ensuite validées par un critère *a contrario* visant à réduire le nombre de fausses alarmes, et les régions non valides ou trop peu significatives fusionnées avec une de ses voisines par un critère de *merging* visant à maximiser la significativité.

Enfin Musé et Sur [MSCG04, CDD⁺04] ont développé une théorie de comparaison et recherche de formes qui pourrait s'avérer très utile pour la stéréo en milieu urbain. Leur approche consiste à trouver des correspondances sûres (validées *a contrario*) entre morceaux de ligne de niveaux (éléments de forme) codés d'une façon invariante par transformations affines. Puis ces correspondances sont groupées en *clusters* maximaux du point de vue de la ressemblance des transformations affines associées, ce que permet de rendre compte de la ressemblance entre formes plus complexes. Nos premiers essais avec cette théorie ont échoué à cause de la précision trop grossière des transformations affines qui relient deux éléments de forme mis en correspondance (de l'ordre de un ou deux pixels) ce qui les rend inutilisables dans le cadre d'un très faible rapport B/H. Néanmoins cette difficulté n'est pas forcément insurmontable.

5.3 Segmentation de textures

Le travail en cours avec Pablo Cancela sur la segmentation de textures dont je suis codirecteur de mastère avec Gregory Randall est très lié à la section précédente. Le problème dans ce cas est celui de la segmentation de textures, mais l'approche utilisée pour le résoudre très semblable à celle de [CGI05]. Dans ce cas au lieu d'utiliser des disparités ou des orientations de gradient, nous utilisons un vecteur de descripteurs de texture à chaque point dont nous cherchons une segmentation constante par régions. Une significativité (*NFA*) est calculée pour chaque paire

(R, T) composée par une région R et hyperrectangle T dans l'espace de descripteurs de texture, en utilisant comme modèle *a contrario* la distribution empirique des descripteurs de texture sur toute l'image. Une sur-segmentation initiale est ensuite modifiée par un critère de *merging* de régions qui vise à maximiser la significativité.

5.4 Conclusions partielles

Les expériences faites avec les différentes méthodes soulignent quelques éléments communs inhérents au problème de la stéréo :

- la pertinence du modèle affine par régions pour la stéréo en milieu urbain
- l'utilité des contours de l'image pour donner des indices sur la localisation des frontières entre régions
- la nécessité d'estimer conjointement le modèle affine de chaque région, en même temps que la partition en régions, une segmentation naïve des niveaux de gris de l'image de référence sans information de disparité n'étant pas suffisante.

Nos futures recherches pourront donc s'orienter vers une approche en deux étapes qui s'inspire de tous les travaux précédents : La première étape chercherait des correspondances affines et sûres T_i pour tous les points x_i de l'image où ceci est possible. Dans ce but il faudrait utiliser des caractéristiques plus locales que les classiques fenêtres de corrélation ou les codes invariants dans [MSCG04] afin d'éviter les phénomènes d'adhérence, mais plus riches que les orientations du gradient dans [CGI05] afin de permettre une meilleure précision. De préférence des caractéristiques morphologiques telles que des petits morceaux de lignes de niveau qui ont le double avantage d'être invariants par changement de contraste et de traverser très rarement la frontière entre deux régions affines différentes. Dans une deuxième étape ces correspondances locales doivent être groupées en paires de régions-transformations (R, T) telles que

- toutes les affinités locales T_i soient proches de T si $x_i \in R$
- l'image de référence u aie un bord bien contrastée le long de la frontière ∂R lorsque l'image secondaire \tilde{u} est bien contrastée le long de la frontière correspondante $T(\partial R)$.

Dans ce but une méthode de *merging* de régions pourrait être utilisée (ce qui a montré ces limites), ou alors un *clustering* du type proposé dans [CDD+04] peut être utilisé. Dans le deuxième cas il faudra au moins des adaptations pour tenir compte du fait que les régions doivent être connexes.

Chapitre 6

Conclusions et Perspectives

J'ai eu pendant une dizaine d'années la joie et la chance d'être confronté aussi bien à des applications réelles du traitement d'images et la vision par ordinateur qu'à des nouvelles théories dans ces deux domaines. J'ai déjà évoqué dans l'introduction l'effet fécond que cette double exposition peu avoir pour le mathématicien appliqué.

Il y a une deuxième synergie entre deux aspects de notre discipline qui a été inspiratrice de mes recherches, à savoir : les relations entre le traitement d'image de bas niveau (restauration d'images par exemple) et la vision par ordinateur d'un plus haut niveau d'abstraction. Ces deux niveaux d'abstraction qu'en termes physiologiques on pourrait appeler niveau "rétinien" et niveau "cortical" (V1, ... V4) sont trop souvent considérés comme deux mondes séparés, ce qui conduit à rater l'opportunité d'une optimisation globale, telle qu'elle est probablement faite par notre système nerveux. Dans mes recherches précédentes j'ai eu l'occasion d'observer certaines relations entre ces deux mondes qui méritent, à mon sens, d'être encore plus développées...

La détection rejoint la restauration : Bien que la réintroduction du principe de Helmholtz et le grand succès de la panoplie de méthodes de détection *a contrario* qui en ont découlé montrent bien l'utilité pratique de ces techniques et leur pertinence comme modèle (partiel) de la perception visuelle, leur généralisation s'est montrée cependant bloquée parfois par au moins deux difficultés : (i) s'assurer que les mesures prises sur l'image sont bien indépendantes, et aussi denses que possible ; (ii) estimer la précision de ces mesures. Nos travaux sur la détection de rayures, et des taches, ainsi qu'un premier essai de valider le principe de Helmholtz dans un cadre d'acquisition d'image plus réaliste ont donné des indices sur la pertinence de ces deux points. Or, aucun de deux ne peut être complètement résolu sans connaître à fond les caractéristiques d'acquisition de l'image, ou alors, sans une procédure préalable de restauration qui rende l'image sous une forme canonique du point de vue de sa résolution et des caractéristiques statistiques du bruit, comme il a été proposé dans le cas de la déquantification des orientations du gradient dans [DLMM00].

La restauration rejoint la détection : Nous avons proposé un modèle de formation d'image qui tient compte de quasiment toutes les dégradations linéaires et stationnaires (souvent pas inversibles et très mal posées) qui subit l'image idéale pendant le processus de capture (bruit, flou de différentes sources, aliasage, et perturbation de l'échantillonnage) et nous avons développé des algorithmes performants permettant de corriger de façon jointe toutes ses dégradations, ob-

tenant une reconstruction quasiment optimale parmi les algorithmes qui ne tiennent compte que d'une connaissance détaillée des caractéristiques de l'instrument de capture et d'une connaissance très faible des caractéristiques statistiques de l'image (vitesse de décroissance des coefficients de Fourier et variation totale faible par rapport au bruit). Ce chemin qui arrive presque à sa fin ouvre deux voies possibles de continuation :

- La substitution d'une connaissance détaillée de l'instrument de capture par des algorithmes d'estimation automatique de ces paramètres comme il a été proposé par Ladjal [Lad05].
- L'utilisation de modèles statistiques d'image plus élaborés, tels que des statistiques d'ordre élevé sur les coefficients d'ondelettes de l'image (voir les travaux de Baraniuk [RCB01] sur le débruitage), ou le fait qu'une image est souvent très autosimilaire comme il a été largement exploité dans les travaux qui ont découlé d'Efros [EL99] comme le débruitage proposé par Buades et Morel [BCCM05].

Dans les deux cas nous sommes ramenés à exploiter la structure géométrique de l'image et à utiliser des méthodes de détection de telles structures -souvent inspirés par notre perception visuelle- afin d'améliorer les performances d'algorithmes de restauration envisagés pour une classe plus large de donnés.

Des méthodes de détection a contrario plus performantes et précises peuvent à mon sens être développées.

Pour améliorer la performance nous pouvons orienter la recherche par la minimisation du NFN (nombre de faux négatifs) comme le proposent Amit et Geman [AG99], en coupant les branches ayant une trop faible probabilité de détection. Dans cette approche (dont ils montrent aussi des liens avec la perception visuelle humaine) le seuil du NFA n'arriverait que dans les feuilles de l'arbre de recherche (le dernier test de validité). Une telle approche peut être très utile dans les applications civiles de reconnaissance automatique d'empreintes digitales, où il est très important de montrer qu'un candidat à obtenir une nouvelle carte d'identité n'est pas dans la base de données de sujets connus (il faut assurer un NFN très bas !), afin d'éviter qu'une même personne puisse avoir deux identités différentes. Dans ce cas la minimisation du NFA (Nombre de Fausses Alarmes) est elle aussi très importante, mais seulement pour des raisons économiques : les fonctionnaires ne devraient pas passer leur temps à vérifier des longues listes de sujets déjà identifiés et ayant peut-être la même identité que le nouveau candidat à obtenir une nouvelle carte.

Pour améliorer la précision, c'est à dire, la localisation des structures détectées nous pourrions envisager une descente de gradient du NFA(G) par rapport aux paramètres définissant la gestalt cherchée G . Pour faire ceci il faut d'abord rendre cette expression différentiable. La majoration de Hoeffding ainsi que la relation entre le NFA et la distance de Kullback-Leibler donne une première indication dans ce sens. Une version différentiable du NFA aurait aussi d'autres avantages. On pourrait par exemple substituer la notion de maximalité par inclusion en faveur de la minimalité locale du NFA dans l'espace de paramètres, ce qui peut donner à son tour une indication plus systématique pour le calcul du nombre de tests : afin de ne pas rater aucune gestalt nous devons faire un échantillonnage de l'espace de paramètres d'un pas qui n'excède pas la distance minimale entre deux minima locaux de NFA.

Enfin mes recherches sur la photogramétrie et l'échantillonnage irrégulier trouvent une continuation très naturelle dans l'estimation de mouvement en vidéo, et son utilisation pour la

restauration de défauts dans les films. Pas seulement l'estimation de disparités en milieu urbain peut s'utiliser (probablement avec modifications) pour l'estimation de mouvement, mais aussi une fois estimé le mouvement nous pouvons utiliser conjointement les échantillons (recalés) de quelques frames précédents et suivants pour améliorer la résolution ou le rapport signal bruit du frame courant. Posé ainsi il s'agit d'un problème de restauration d'échantillonnage irrégulier qui pourrait donner une nouvelle application à la méthode que nous avons proposée avec Caselles, Haro et Rougé dans le cadre de l'imagerie satellitaire. Dans ce but nous avons soumis un projet ECOS avec Cao, Gousseau, Masnou, Randall, Preciozzi et Grompone, et un projet PDT avec Pardo, Musé, Memoli, Bartesaghi et Preciozzi. Des sujets autour de la photogramétrie, l'estimation de mouvement et la restauration de films feront donc très probablement l'objet de la thèse de Preciozzi en co-tutelle avec Morel, et de Facciolo en co-tutelle avec Caselles, à partir de 2006.

Annexe A

Publications et encadrement

Articles et Livres en preparation :

- [P1] Jean-Michel Morel, Bernard Rougé, and Andrés Almansa. *A First Course on Fourier Analysis for Image Processing*. En preparation pour la série LNM, Springer Verlag.
- [P2] Gabriele Facciolo, Andrés Almansa, and Alvaro Pardo. Variational approach to interpolate and correct biases in correlation of stereo pairs. In *GRETSI*, 2005. accepté. [PDF](#).
- [P3] Gloria Haro, Vicent Caselles, Andrés Almansa, and Bernard Rougé. Restoration and zoom of irregularly sampled, blurred and noisy images, by total variation minimization with local constraints. En preparation pour *Multiscale Modeling and Simulation*. [PDF](#).

Articles publiés dans des revues à comité de lecture

- [P4] Andrés Almansa, Sylvain Durand, and Bernard Rougé. Measuring and improving image resolution by adaptation of the reciprocal cell. *Journal of Mathematical Imaging and Vision*, 21(3) :235–279, November 2004. [PDF](#). Quasi-final version in [P26, chapter 1].
- [P5] Andrés Almansa, Agnès Desolneux, and Sébastien Vamech. Vanishing points detection without any a priori information. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(4) :502–507, April 2003. Special issue “Perceptual Organisation in Computer Vision”. Submitted Dec 2001 ; revised Jul 2002 ; final submission of accepted manuscript Dec 2002. [PDF](#). Extended version in [P26, chapter 4].
- [P6] Andrés Almansa, Frédéric Cao, Yann Gousseau, and Bernard Rougé. Interpolation of digital elevation models using AMLE and related methods. *IEEE Trans. on Geoscience and Remote Sensing*, 40(2) :314–325, February 2002. [PDF](#). Extended and corrected version in [P26, chapter 3].
- [P7] Andrés Almansa and Tony Lindeberg. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. *IEEE Trans. on Image Processing*, 9(12) :2027–2042, December 2000. [PDF](#)
- [P8] Carlo Graziani and Andrés Almansa. Ein Algorithmus zur Simulation linearer dynamischer Modelle in kontinuierlicher Zeitformulierung unter vollkommener Voraussicht. *Wirtschaftswissenschaftliches Studium*, 27(6) :319–324, June 1998.

- [P9] Carlo Graziani and Andrés Almansa. Un algoritmo para la simulación de modelos lineales en tiempo continuo bajo previsión perfecta. *Estudios de Economía, Universidad de Chile*, 24(1) :185–196, June 1997.
- [P10] Carlo Graziani and Andrés Almansa. Un procedimiento para la simulación de modelos lineales en tiempo continuo bajo previsión perfecta e histéresis. *Estudios de Economía del Colegio de México*, 13(1), January-June 1998. [PDF](#).

Chapitre de livre

- [P11] Andrés Almansa and Tony Lindeberg. Enhancement of fingerprint images using shape-adaptation of scale-space operators. In Jon Sporring, Mads Nielsen, Luc Florack, and Peter Johansen, editors, *Gaussian Scale-Space Theory*, volume 8 of *Computational Imaging and Vision*, chapter 2, pages 21–29. Kluwer Academic Publishers, 1997.

Conférences avec comité de lecture

- [P13] Andrés Almansa, Sylvain Durand, and Bernard Rougé. Measuring and improving image resolution by adaptation of the reciprocal cell. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Toulouse, France, June 2003. (extended abstract and poster).
- [P15] Andrés Almansa, Stéphane Jaffard, and Bernard Rougé. Perturbed sampling in satellite images and reconstruction algorithms. In *18th Symposium GRETSI*, Toulouse, France, September 2001.
- [P16] Andrés Almansa and Laurent Cohen. Fingerprint image matching by minimization of thin-plate energy using a two-step algorithm with auxiliary variables. In *Proc. Fifth IEEE Workshop on Applications of Computer Vision (WACV 2000)*, pages 35–40, Palm Springs, CA, December 2000. IEEE, Computer Society Press. [PDF](#).
- [P18] Andrés Almansa and Tony Lindeberg. Enhancement of fingerprint images using shape adaptation of scale-space operators. In *PhD School on Classical Scale-Space Theory*, Copenhagen, Denmark, May 1996. Computer Science Dept., University of Copenhagen.
- [P20] Gustavo Drets, Rosario Curbelo, Olaf Bergengruen, and Andrés Almansa. Métodos de impresión fiel para huellas dactilares. In *Memorias de la XXII Conferencia Latinoamericana de Informática (CLEI Panel 96)*, volume 2, pages 659–670, Santafé de Bogotá, Colombia, June 1996. [PDF](#).
- [P21] Rosario Curbelo, Andrés Almansa, Gustavo Drets, and Olaf Bergengruen. Transmisión remota de huellas dactilares para la justicia en el uruguay. In *Jornadas Chilenas de la Computación. Actas del IV Encuentro Chileno de Computación*, pages 93–98, Valdivia, Chile, November 1996. SCCC, Universidad Austral.
- [P22] Andrés Almansa, Olaf Bergengruen, Rosario Curbelo, and Gustavo Drets (autores en orden alfabético). Sistema de transmisión remota de huellas dactilares. In *Memorias del 5to. Congreso Internacional de Nuevas Tecnologías de La Habana, INFORMATICA'96*, La Habana, Cuba, 1996.

Conférences invitées ou sans comité de lecture

- [P14] Andrés Almansa, Stéphane Jaffard, and Bernard Rougé. Irregular sampling in satellite images and reconstruction algorithms. In *Congrès d'Analyse Numérique (CANUM)*, La Grande Motte, France, July 2003. [PDF](#). Quasi-final version in [P26, chapter 2].
- [P17] Andrés Almansa, Frédéric Cao, and Bernard Rougé. Image interpolation and application to digital elevation models. In *European Workshop Frontiers of Mathematical Morphology*. Université de Strasbourg, April 2000.
- [P23] Carlo Graziani and Andrés Almansa. Dos algoritmos para la simulación de modelos lineales con previsión perfecta en tiempo continuo. In *Anales de las X Jornadas Anuales de Economía*, Montevideo, Uruguay, December 1995. Banco Central del Uruguay.

Rapports, mémoires, logiciels, autres

- [P24] Andrés Almansa, François Alter, and Vicent Caselles. Restauration d'images régularisée part la variation totale : le cas de l'échantillonnage perturbé. Technical report, Universitat Pompeu Fabra, December 2003. Rapport Technique de l'UPF pour le CNES.
- [P25] Frédéric Guichard, Thierry Cohignac, and Andrés Almansa. Indexation sémantique d'images sur internet (ISII) : partie 2. Technical report, Vision IQ, December 2002. Rapport final de VisionIQ pour l'action RNRT du Ministère de la Recherche.
- [P26] Andrés Almansa. *Echantillonnage, Interpolation et Détection. Applications en Imagerie Satellitaire*. PhD thesis, Ecole Normale Supérieure de Cachan, 94235 Cachan cedex, France, December 2002. [PDF](#). Contains extended versions of [P4,P5,P6,P14].
- [P27] Andrés Almansa, Agnès Desolneux, Lionel Moisan, and Jean-Michel Morel. Vanishing point detection using Helmholtz principle. In *The Mathematical, Computational and Biological Study of Vision (URL)*, pages 2–3. Mathematisches Forschungsinstitut Oberwolfach ([MFO](#)), November 2001. [PostScript](#).
- [P28] Andrés Almansa, Sylvain Durand, and Jean-Michel Morel. Analyse de la résolution et de la stabilité des différents modes d'acquisition spot. Technical report, Centre de Mathématiques et Leurs Applications (CMLA), École Normale Supérieure de Cachan, 94235 Cachan cedex, France, February 2000. Internal Technical Report for CNES.
- [P29] Andrés Almansa, Stéphane Jaffard, and Jean-Michel Morel. Perturbed image sampling : Optimal sampling theorems and correction algorithms. Technical report, Centre de Mathématiques et Leurs Applications (CMLA), École Normale Supérieure de Cachan, 94235 Cachan cedex, France, September 2000. Internal Technical Report for CNES.
- [P30] Andrés Almansa. Hexagonal sampling in image processing. Technical report, Centre de Mathématiques et Leurs Applications (CMLA), École Normale Supérieure de Cachan, 94235 Cachan cedex, France, August 1999. Mémoire de stage de DEA.
- [P31] Andrés Almansa and Tony Lindeberg. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. Technical Report ISRN KTH/NA/P-99/01-SE, Dept. of Numerical Analysis and Computing Science, KTH (Royal Institute of Technology), 100 44 Stockholm, Sweden, 1999. [PDF](#). Preprint of [P7]. Earlier and more extended version in ISRN KTH/NA/P-98/03-SE [P32]

- [P32] Andrés Almansa and Tony Lindeberg. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. Technical Report ISRN KTH/NA/P-98/03-SE, Dept. of Numerical Analysis and Computing Science, KTH (Royal Institute of Technology), 100 44 Stockholm, Sweden, 1998. [PDF](#). Earlier version in J. Sporring, et al. (eds.) (1997) *Gaussian Scale-Space Theory*, chapter 2 [P11].
- [P33] Andrés Almansa. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. Master's thesis, Facultad de Ingeniería, Montevideo, Uruguay, April 1998. [PostScript](#).
- [P34] R. Curbello, G. Drets, A. Almansa, Proyecto CONICYT-BID 96/94 : Desarrollo de un sistema de identificación automática de huellas dactilares. Informe final. CeCal, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay, 1999.
- [P35] Andrés Almansa. Ridge enhancement of fingerprint images. Technical Report 3 :2, CeCal, Fac. de Ingeniería, Univ. de la República, 11300 Montevideo, Uruguay, 1994.
- [P36] Andrés Almansa and Vladimir Hinich (mentor). Widespread subsets in a cube. In *Scientific Works submitted by the participants of the 21st. Dr. Bessie F. Lawrence International Summer Science Institute*, volume 1, pages M13-M20, Rehovot, Israel, 1989. Weizmann Institute of Science.

Encadrement En 2003, pendant ma période de transition entre Paris, Barcelone et Montevideo, et avec le souci de trouver les moyens pour garder à plus long terme les collaborations académiques entre différents groupes de traitement d'images avec lesquels j'avais eu une expérience fructueuse auparavant, j'ai mis en place en collaboration avec Gregory Randall et Jean-Michel Morel un projet de mise en réseau de 7 universités européennes et latino-américaines pour la formation en co-tutelle de 14 docteurs sur une période de trois ans.

Ce projet Alfa appelé "Computer Vision Foundations and Applications" a été retenu par la Commission Européenne en 2004 et vient de démarrer sous la coordination de Morel, Randall et moi même. Dans ce cadre je me suis engagé à encadrer la thèse de Javier Preciozzi en co-tutelle avec Jean-Michel Morel à l'ENS Cachan, et celle de Gabriele Facciolo en co-tutelle avec Vicent Caselles à l'UPF, Barcelone.

Activités d'encadrement en cours ou passées :

- 2004-2005** Directeur de thèse de mastère de *Gabriele Facciolo* sur les "Approches variationnelles pour la correction du phénomène d'adhérence en corrélation de paires stéréo satellitaires en milieu urbain" [P2] en cours (PEDCIBA, UdelaR, Uruguay)
- 2004-2005** Directeur de mémoire de DEA et thèse de mastère de *Javier Preciozzi* sur les "Méthodes géométriques, morphologiques et a contrario pour l'estimation de la carte de disparité par photogrammétrie en milieu urbain" (ENS Cachan, France et PEDCIBA, UdelaR, Uruguay).
- 2004-2005** Codirecteur de thèse de mastère de Pablo Cancela sur la "Segmentation de textures" avec Gregory Randall (UdelaR, Uruguay).
- 2004-2005** Collaboration avec Vicent Caselles dans l'encadrement de la dernière partie de la thèse de doctorat de Gloria Haro sur la "Restauration d'images irrégulièrement échantillonnées"

par minimisation de la variation totale avec des contraintes locales” (Universitat Pompeu Fabra, Barcelone).

- 2005** Codirecteur de mémoire de fin d’études de Adrian Marques sur la conception d’”Algorithmes pour la conversion de vidéo entrelacé en progressif”, avec Marcelo Bertalmío, UPF, Barcelona. [Projet Racine](#).
- 2004** Codirecteur de mémoire de fin d’études de Gastón Fernandez, Natalia Tourn, Claudia Stocco sur ”Vision en football de robots” avec Gonzalo Tejera et Nelson Calero (UdelaR, Uruguay).
- 1996 - 1999** Directeur ou co-directeur de 5 mémoires de fin d’études en Informatique (UdelaR, Uruguay) sur le traitement d’images, d’empreintes digitales et de documents.

Bibliographie

- [86] *Partial Differential Equations*. Chelsea, New York, 1986. see page 275.
- [AD03] Jean-Michel Morel Agnès Desolneux, Lionel Moisan. Computational gestalts and perception thresholds. *Journal of Physiology - Paris*, 97 :311–324, 2003.
- [AG99] Yali Amit and Donald Geman. A computational model for visual selection. *Neural Computation*, 1999.
- [AGLM93] L. Alvarez, F. Guichard, P.L. Lions, and J.M. Morel. Axioms and fundamental equations of image processing. *Arch. Ration. Mech. Anal.*, 123(3) :199–257, 1993.
- [BBC⁺01] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. Vancouver, Canada, July 2001.
- [BCCM05] Antoni Buades-Capo, Bartomeu Coll, and Jean-Michel Morel. On image denoising methods. Technical Report 2004-15, CMLA, ENS Cachan, 2005.
- [BCRS03] M. Bertalmío, V. Caselles, B. Rougé, and A. Solé. TV based image restoration with local constraints. *Journal of Scientific Computing*, 19(1-3) :95–122, 2003.
- [BCRV] C. Ballester, V. Caselles, B. Rougé, and J. Verdera. A variational model for p+xs image fusion. Submitted. [PDF](#).
- [Beu89] A. Beurling. . In L. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, editors, *The collected works of Arne Beurling*, volume 2, pages 341–365. Birkhauser, Boston, MA, 1989.
- [Bey95] G. Beylkin. On the fast Fourier transform of functions with singularities. *ACHA*, 2 :363–381, 1995.
- [Bey98] G. Beylkin. On applications of unequally spaced fast Fourier transform. In *Mathematical Geophysics Summer School*, Stanford, August 1998.
- [BK04] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9) :1124–1137, September 2004.
- [Bog01] T. Boggio. Sull’equilibrio delle piastre elastiche incastrate. *Rend. Acc. Lincei*, 10 :197–205, 1901.

- [Boo89] Fred L. Bookstein. Principal warps : Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(6) :567–585, June 1989.
- [BSCB00] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *SIGGRAPH 2000*, pages 417–424, New Orleans, USA, July 2000.
- [CDD⁺04] F. Cao, J. Delon, A. Desolneux, Y. Gousseau, P. Musé, and F. Sur. An a contrario approach to hierarchical clustering validity assessment. Technical report, CMLA, ENS Cachan, 61 av. du Président Wilson, 94235 Cachan, France, 2004.
- [CDRJ05] Yi Chen, Sarat Dass, Arun Ross, and Anil Jain. Fingerprint deformation models using minutiae locations and orientations. In *Proc. of IEEE Workshop on Applications of Computer Vision (WACV)*, pages 150–156, Breckenridge, Colorado, 2005.
- [CGI05] Vicent Caselles, Luis Garrido, and Laura Igual. A contrast invariant approach to motion estimation and validation. In *IEEE International Conference on Image Processing (ICIP)*, Genova, Italy, 2005.
- [Cha02] A. Chambolle. An algorithm for total variation minimization and applications. *Preprint Ceremade 02040*, 2002.
- [CMS98] V. Caselles, J.M. Morel, and C. Sbert. An axiomatic approach to image interpolation. *IEEE Trans. Image Processing*, 7(3), March 1998.
- [Cof82] C.V. Coffman. On the structure of solutions of $\Delta^2 u = \lambda u$ which satisfy the clamped plate condition on a right angle. *SIAM Journal of Mathematical Analysis*, 13 :746–757, 1982.
- [Coh96] Laurent D. Cohen. Auxiliary variables and twostep iterative algorithms in computer vision problems. *Journal of Mathematical Imaging and Vision*, 6 :59–83, 1996.
- [Cre93] N. Cressie. *Statistics for Spatial Data*. Wiley, New York, revised edition, 1993.
- [CRZ00] A. Criminisi, I. Reid, and A. Zisserman. Single view metrology. *International Journal of Computer Vision*, 40(2) :123–148, November 2000.
- [CS93] A. Cohen and J.M. Schlenker. Compactly supported bidimensional wavelet bases with hexagonal symmetry. *Constructive Approximation*, (9) :209–236, 1993.
- [CS96] C. Chui and X. Shi. On stability bounds of perturbed multivariate trigonometric systems. *Applied and Computational Harmonic Analysis*, 3 :283–287, 1996.
- [Del04] J. Delon. *Fine comparison of images and other problems*. PhD thesis, Ecole Normale Supérieure de Cachan, 2004.
- [Des00] Agnès Desolneux. *Événements significatifs et applications à l’analyse d’images*. PhD thesis, École Normale Supérieure de Cachan, December 2000.
- [DGSS99] I. Daubechies, I. Guskov, P. Schröder, and W. Sweldens. Wavelets on irregular point sets. *Phil. Trans. Royal Society of London A*, 1999. Submitted.
- [DLMM00] A. Desolneux, S. Ladjal, L. Moisan, and J.M. Morel. Dequantizing image orientation. Technical Report 2000-23, CMLA, ENS Cachan, 2000.

- [DMM00] Agnès Desolneux, Lionel Moisan, and Jean-Michel Morel. Maximal meaningful events and applications to image analysis. Technical report, preprint CMLA N° 2000-22, 2000.
- [DMM01] Agnès Desolneux, Lionel Moisan, and Jean-Michel Morel. Edge detection by Helmholtz principle. *Journal of Mathematical Imaging and Vision*, 14(3) :271–284, 2001.
- [DMM02a] Agnès Desolneux, Lionel Moisan, and Jean-Michel Morel. Gestalt theory and computer vision. Technical report, preprint CMLA N° 2002-06, 2002.
- [DMM02b] Agnès Desolneux, Lionel Moisan, and Jean-Michel Morel. A grouping principle and four applications. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2002. special issue on "Perceptual Organization in Computer Vision", to appear.
- [DR01] J. Delon and B. Rougé. Le phénomène d'adhérence en stéréoscopie dépend du critère de corrélation. In *GRETSI*, Toulouse, France, September 2001.
- [EL99] A.A. Efros and T.K. Leung. Texture synthesis by non-parametric sampling. In *ICCV99*, pages 1033–1038, 1999.
- [FGS95] H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numerische Mathematik*, 69 :423–440, 1995.
- [FK98] O. Faugeras and R. Keriven. Variational principles, surface evolution, pde's, level set methods, and the stereo problem. *T-IP*, 7 :336–344, 1998.
- [FL04] Gabriele Facciolo and Tomas Laurenzo. Towards a more general psychophysical validation of the Helmholtz principle. In *Symposium on Reality Representations by brain and machines : crossed views from neurosciences and computer vision*, Montevideo, Uruguay, November 2004.
- [FLP01] Olivier Faugeras, Quang-Tuan Luong, and T Papadopoulos. *The Geometry of Multiple Images*. The MIT Press, 2001.
- [FS75] R.W. Floyd and L. Steinberg. An adaptive algorithm for spatial gray scale. In Society for Information Displays, editor, *International Symposium Digest of Technical Papers*, volume 17, pages 75–77, 1975.
- [FS94] H.G. Feichtinger and T. Strohmer. Recovery of missing segments and lines in images. *Optical Engineering*, 33(10) :3283–3289, 1994. special issue on Digital Image Recovery and Synthesis.
- [FZ95] S.J. Favier and R.A. Zalik. On stability of frames and Riesz bases. *Applied and Computational Harmonic Analysis*, 2 :160–173, 1995.
- [Gar51] P.R. Garabedian. A partial differential equation arising in conformal mapping. *Pacific J. Math.*, 1 :485–524, 1951.
- [GM01] Yann Gousseau and Jean-Michel Morel. Are natural images of bounded variation ? *SIAM Journal of Mathematical Analysis*, 33(3) :634–648, 2001.
- [Grö92] K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comp.*, 59(181–192), 1992.

- [GS96] Christoph Grunau and Guido Sweers. Positivity for perturbations of polyharmonic operators with Dirichlet boundary conditions in two dimensions. *Math. Nachr.*, 179 :89–102, 1996.
- [GS98] L. Garrido and P. Salembier. Region based analysis of video sequences with a general merging algorithm. In *European Signal Processing Conference (EU-SIPCO)*, Rhodes, Greece, September 1998.
- [GS00] K. Gröchenig and T. Strohmer. Numerical and theoretical aspects of non-uniform sampling of band-limited images. In F. Marvasti, editor, *Theory and Practice of Nonuniform Sampling*. Kluwer/Plenum, 2000.
- [GVL96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition edition, 1996. isbn 0801854148.
- [Had08] J. Hadamard. Mémoire sur le problème d’analyse relatif à l’équilibre des plaques élastiques incastées. *Mémoires Présentées par Divers Savants a l’Académie des Sciences*, 33 :1–128, 1908.
- [Jaf89] Stéphane Jaffard. *Construction et propriétés des bases d’ondelettes. Remarques sur la contrôlabilité exacte*. PhD thesis, École Polytechnique, May 1989.
- [Jaf00] Stéphane Jaffard. Résultats généraux sur l’échantillonnage irrégulier. Personal communication, 2000.
- [Jai95] Anil K. Jain. *Fundamentals of Image Processing*. Prentice Hall, 1995.
- [JH78] A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, London, 1978.
- [JHB97] Anil Jain, Lin Hong, and Ruud Bolle. On-line fingerprint verification. *IEEE Trans. Pattern Analysis and Machine Intell.*, 19(4) :302–3013, April 1997.
- [Kad64] M. I. Kadec. The exact value of the Payley-Wiener constant. *Soviet Math. Doklady*, 5(559–561), 1964.
- [Kan80] Gaetano Kanizsa. *Grammatica del Vedere*. Società editrice Il Mulino, Bologna, 1980.
- [KLM94] G. Koepfler, C. Lopez, and J.M. Morel. A multiscale algorithm for image segmentation by variational method. *SIAM J. NUMERICAL ANALYSIS*, 31 :282–299, 1994.
- [Koe84] J. J. Koenderink. The structure of images. In *Biol. Cyb.*, volume 50, pages 362–370, 1984.
- [Kok96] A. Kokaram. Detection and removal of line scratches in degraded motion picture sequences. In *Signal Processing VIII* *Signal Processing VIII*, pages 5–8, September 1996.
- [KS98] J. Kovacevic and W. Sweldens. Wavelet families of increasing order in arbitrary dimensions. *IEEE Trans. Image Processing*, 1998. Submitted.
- [Lad05] Saïd Ladjal. *Quantification et flou dans les images numériques*. PhD thesis, ENS Cachan, 2005.

- [Lan67] H.J. Landau. Necessary density conditions for sampling and interpolation of entire functions. *Acta Math*, 117 :37–52, 1967.
- [Lin94] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.
- [Lin96] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. In *Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, pages 465–470, San Francisco, USA, June 1996. IEEE Computer Society Press.
- [LRS01] Christophe Latty, Bernard Rougé, and Baillarin S. La chaîne image SPOT5 THR : un exemple d’optimisation globale. In *GRETSI*, Toulouse, September 2001.
- [LS94] A. F. Laine and S. Schuler. Hexagonal wavelet representations for recognizing complex annotations. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 740–745, Los Alamitos, CA, USA, June 1994. IEEE Computer Society Press.
- [Mal99] François Malgouyres. *Augmentation de la résolution d’images satellitaires : Théorie variationnelle et applications*. PhD thesis, Université de Paris IX – Dauphine, 1999.
- [Mar82] David Marr. *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, 1982. ISBN 0716712849.
- [Mas98] S. Masnou. *Filtrage et Desocclusion d’Images par Methodes d’Ensembles de Niveau*. PhD thesis, Université Paris-IX Dauphine, 1998.
- [Mat71] G. Matheron. La théorie des variables régionalisées, et ses applications. *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau*, 5, 1971.
- [Mer79] R.M. Mersereau. The processing of hexagonally sampled two-dimensional signals. *IEEE Proceedings*, 67 :930–949, 1979.
- [Mey90] Yves Meyer. *Wavelets and operators*, volume 2. Hermann, 1990.
- [Mey02] Yves Meyer. Oscillating patterns in image processing and nonlinear evolution equations,. In *University Lecture Series*, volume 22. AMS, 2002.
- [MG01] François Malgouyres and Frederic Guichard. Edge direction preserving image zooming : a mathematical and numerical analysis. *Journal on Numerical Analysis*, 39(1) :1–37, 2001.
- [MJS89] D. Mumford and J. J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42 :577–684, 1989.
- [ML98] J.M. Morel and S. Ladjal. Analyse de Fourier et théorie de Shannon. Analyse de Fourier et traitement d’images. Journées X-UPS 1998. Centre de Mathématiques – Ecole Polytechnique. France., 1998.
- [MM01] Lionel Moisan and Jean-Michel Morel. Les défauts du DVD sous l’oeil de l’équation. *La Recherche*, (342) :56–57, May 2001.

- [MS83] R.M. Mersereau and T.C. Speake. The processing of periodically sampled multidimensional signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(1) :188–194, February 1983.
- [MSCG03] P. Musé, F. Sur, F. Cao, and Y. Gousseau. Unsupervised thresholds for shape matching. In *IEEE International Conference on Image Processing (ICIP03)*, Barcelona, Spain, September 2003.
- [MSCG04] P. Musé, F. Sur, F. Cao, and Y. Gousseau. An a contrario decision method for shape element recognition. Technical report, CMLA, ENS Cachan, 61 av. du Président Wilson, 94235 Cachan, France, 2004.
- [Mus04] Pablo Musé. *Sur la définition et la reconnaissance de formes planes dans les images numériques (On the definition and recognition of planar shapes in digital images)*. PhD thesis, ENS Cachan, 61 av. du Président Wilson, 94235 Cachan, France, October 2004. [PDF](#).
- [PM90] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. 12(7) :629–639, 1990.
- [PN82] Murphy P.K. and Gallagher N.C. Hexagonal sampling techniques applied to Fourier and Fresnel digital holograms. *Journal of the Optical Society of America*, 72(7) :929–37, July 1982.
- [Pot01] D. Potts. NFFT. <http://www.math.mu-luebeck.de/potts/nfft/>, 2001.
- [Pra78] William K. Pratt. *Digital Image Processing*. John Wiley and Sons, 1978.
- [Rau98] M. Rauth. *Gridding geophysical potential fields from noisy scattered data*. PhD thesis, University of Vienna, May 1998.
- [RCB01] J.K. Romberg, Hyeokho Choi, and Richard Baraniuk. Bayesian tree-structured image modeling using wavelet-domain hidden Markov models. *IEEE Trans. Image Processing*, 10(7) :1056–1068, July 2001.
- [RLO03] L. Rudin, P.L. Lions, and S. Osher. Multiplicative denoising and deblurring images with constrained non-linear pde's ii. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer Verlag, 2003.
- [ROF92] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60 :259–268, 1992.
- [Rou97] B. Rougé. *Théorie de la chaîne image optique et restauration à bruit final fixé. Mémoire en vue de l'obtention de l'habilitation à diriger des recherches. Option : Mathématiques appliquées*, May 1997.
- [SA90a] Eero P. Simoncelli and Edward H. Adelson. Subband image coding with hexagonal quadrature mirror filters. In *Picture Coding Symposium*, Cambridge, MA., March 1990.
- [SA90b] E.P Simoncelli and E.H Adelson. Subband transforms. In John W. Woods, editor, *Subband Coding*, chapter 4, pages 143–192. Kluwer Academic Publishers, 1990. Also available from MIT Media Laboratory as Technical Report #137, <ftp://whitechapel.media.mit.edu/pub/tech-reports/TR-137.ps.Z>.

- [San76] Luis A. Santaló. Integral geometry and geometric probability. In Gian-Carlo Rota, editor, *Encyclopedia of Mathematics and its Applications*, volume 1. Addison-Wesley, 1976.
- [SJ95] L. Rudin S. Jensen. Measure : an interactive tool for accurate forensic photo/video grammetry. In *Investigative & Trial Image Processing Conference*, volume 2567 of *Proceedings of the SPIE - The International Society for Optical Engineering*, San Diego, CA, 1995.
- [SL97] S. Schuler and A. Laine. Hexagonal QMF banks and wavelets. In M. Akai, editor, *Time-Frequency and Wavelet Transforms in Biomedical Engineering*, chapter ? IEEE Press, 1997.
- [Soi91] Pierre Soille. Spatial distributions from contour lines : An efficient methodology based on distance transformations. *Journal of Visual Communication and Image Representation*, 2(2) :138–150, June 1991.
- [Soi94] Pierre Soille. Generalized geodesic distances applied to interpolation and shape description. In Jean Serra and Pierre Soille, editors, *Mathematical Morphology and its Applications to Image Processing*. Kluwer Academic Publishers, 1994.
- [SS02] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47 :7–42, 2002.
- [ST93] G. Sapiro and A. Tannenbaum. Affine invariant scale space. *International Journal of Computer Vision*, 11(1) :22–44, 1993.
- [Sur04] Frédéric Sur. *Décision a contrario pour la reconnaissance de formes (A contrario decision for shape recognition)*. PhD thesis, ENS Cachan, 61 av. du Président Wilson, 94235 Cachan, France, October 2004. [PDF](#).
- [Tik63] A.N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Mathematics Doklady*, 4 :1624–1627, 1963.
- [Wei95] J. Weickert. Multiscale texture enhancement. In *Proc 6th Int Conf on Computer Analysis of Images and Patterns (Prague Czech)*, volume 970 of *Lecture Notes in Computer Science*, pages 230–237. Springer Verlag, Berlin, Sep 1995.
- [Wei98] Joachim Weickert. *Anisotropic Diffusion in Image Processing*. ECMI Series. Teubner-Verlag, Stuttgart, Germany, 1998. ISBN 3-519-02606-6.
- [Wer23] Max Wertheimer. Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forshung*, 4 :301–350, 1923.
- [Wit83] A. P. Witkin. Scale-space filtering. In *8th International Conference on Artificial Intelligence*, pages 1019–1022, Karlsruhe, Germany, August 1983.
- [WtHRV98] J. Weickert, B.M. ter Haar Romeny, and M.A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Processing*, 7 :398–410, 1998.
- [WZtHRN97] J. Weickert, K.J. Zuiderveld, B.M. ter Haar Romeny, and W.J. Niessen. Parallel implementations of aos schemes : A fast way of nonlinear diffusion filtering. In *IEEE International Conference on Image Processing (ICIP-97)*, volume 3, pages 396–399, Santa Barbara, October 1997.

- [You80] Robert M. Young. *Introduction to Nonharmonic Fourier Series*. Academic Press, 1980.