

Ministère de l'Enseignement Supérieur et de la Recherche scientifique
وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des Sciences de l'Ingéniorat
Département d'Electronique

Année : 2015/2016

THESE

Présentée en vue de l'obtention du diplôme de *DOCTORAT*

**Etude et Application des transformées géométriques à la
Compression des images hautes résolutions et à la Biométrie
(Authentification/Vérification de l'empreinte palmaire)**

Option

Signaux et Images

Par

BOUCHEMHA Amel

Directeur de thèse : M. Nouredine DOGHMANE Professeur Université Badji Mokhtar, Annaba

Co-directeur : M. Amine NAIT ALI Professeur Université Paris-Est Créteil, France

Devant le Jury

| | | |
|-------------|-------------------------|------------------------|
| Président : | Pr BENNACER Layachi | Université de Annaba |
| Examineur : | Pr SERIR Amina | Université USTHB Alger |
| Examineur : | Pr BOUKROUCHE Abdelhane | Université de Guelma |
| Examineur : | Pr ROUAINIA Mounira | Université de Skikda |

Année Universitaire : 2015/2016

Dédicaces

Je dédie ce travail à :

La mémoire de mes parents

Toute personne que j'ai aimée et que j'ai perdue au cours du chemin de ma vie. Leur perte reste une douleur immense dans mon cœur (Mon père Ahmed le bien aimé, Ma mère Louiza, ma lumière, Ma sœur Hamida, ma douce et mon beau frère Hamid, l'amî perdu).

La personne qui est toujours à mes côtés, toi mon mari, mon ami et mon confident

Mes sœurs, mes frères.

Mes beaux parents

Mes neveux et nièces

Toute ma famille

Amel

Remerciements

Je remercie mon Dieu de m'avoir permis de réaliser et de faire aboutir ce travail de thèse.

A l'issue de la rédaction de ce travail, je suis convaincue que la thèse est loin d'être un travail solitaire. En effet, je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser.

Je tiens tout d'abord à remercier mes encadreurs, J'exprime toute ma gratitude à monsieur N. Doghmane, Professeur à l'Université Badji Mokhtar Annaba et à monsieur Amine Nait-Ali, Professeur à l'université Paris-Est Créteil (UPEC), de m'avoir dirigé et guidé tout le long de ce travail. Leurs intuitions et leur savoir, mais aussi leur exigence et leurs critiques m'ont permis de progresser dans mes démarches scientifiques. Les nombreuses discussions tout au long de la thèse ont été très riches et fructueuses. Qu'ils trouvent ici l'expression de ma considération et ma profonde admiration pour toutes leurs qualités scientifiques et humaines.

Je remercie chaleureusement les membres du jury d'avoir bien voulu accepter d'examiner et d'évaluer le présent travail, et pour l'honneur qu'ils me font d'être dans mon jury de thèse.

J'exprime toute ma reconnaissance et mes remerciements à Monsieur BENNACER Layachi, Professeur à l'université de Annaba, qui a bien voulu me faire l'honneur de présider le jury de ma thèse.

Je tiens à remercier vivement, Madame SERIR Amina, Professeur à l'université USTHB Alger et Madame ROUAINIA Mounira, Professeur à l'université de Skikda, qui m'ont fait l'honneur d'examiner mon travail et de faire partie de mon jury de thèse.

Je remercie Monsieur BOUKROUCHE Abdelhani, Professeur à l'université de Guelma, pour l'intérêt qu'il a bien voulu porter à ce travail en acceptant de le juger et de faire partie du jury.

Je tiens à remercier mon conjoint d'être à mes côtés et de m'avoir supporté durant ces années de recherche, j'espère qu'il trouve dans ces mots la gratitude et le respect que je porte envers lui.

Que toutes les personnes qui ont contribué de près ou de loin, directement ou indirectement à ce travail, trouvent ici le témoignage de ma profonde reconnaissance.

ملخص

ثراء الصور الطبيعية يجعل السعي من اجل التمثيل الأمثل في مجال معالجة الصور و الرؤية الحاسوبية تحديا كبيرا للباحثين الذين غالبا ما يطلب منهم حلا وسط بين الكفاءة و التعقيد. للأسباب من البساطة, كثيرا ما استخدمت المويجات بطريقة منفصلة على المحاور الأفقية و العمودية للصورة. وهذا ينتج عنه فك الارتباط جزئيا بين نقاط الصورة. مما يؤدي لظهور على حدود الصورة العديد من المعاملات عالية الطاقة. المويجات الكلاسيكية غير قادرة على تمثيل بشكل فعال هذه المزايا الخاصة بكفاءة بسبب قالب المربع للموجة علاوة عن ذلك فإنها تفتقر على خصائص اتجاهية. إذا تحويل المويجات يسمح بعزل نقاط التقاطع الموجودة في الصورة لكنه لا يسمح بتمييز المنحنيات المتجانسة في الصورة. التحويلات الموجهية المستمدة من التحويلات الموجتية توفر لهذه الإشكالية حلا و تسمح بالتمثيل الفعال لهذا التجانس الهندسي الموجود في الصورة مثل الحدود و النسخ. في الواقع هذا التجانس الهندسي هو نوع من أنواع التكرار الذي يجدر الاهتمام باستغلاله لتمثيل أجوف للصورة وهذا من اجل الحصول على تقريبية جديدة للصورة مع قليل من المعاملات.

في هذه الأطروحة تم دراسة و انجاز محوريين للبحث, المحور الأول يخص الضغط التدريجي للصورة العالية الدقة (صور الأقمار الصناعية و الصور الطبية) باستخدام تحويلات الشرائط (*bandelottes*), اما المحور الثاني يخص اقتراح نظام بيومتري الذي يسمح بتحديد هوية الشخص باستعمال صور راحة اليد المتعددة الأظياف و صور راحة اليد باستخدام التحويلات الموجهية الغير التكيفية (*non adaptative*). لهذا الغرض تم استعمال تحويلات *Ridgelets* و تحويلات *Radon* للاستخراج سمات تمييزية لوصف النماذج البيومترية المستخدمة. في أول المساهمة العلمية تم استخدام تحويلات *Radon* لوصف الخطوط الرئيسية لليد و تجرى عملية التحقق باستخدام مسافة *Hausdorff* مع تثليث ديلوناي. أما في المساهمة الثانية, نقترح استعمال تحويلات *Ridgelets* مع دمج المعلومات على مستوى *Pixel* و على مستوى السيمة لصور راحة اليد و صور عروق اليد و هذا لتحسين كفاءة النظام البيومتري المقترح.

الكلمات الرئيسية : التحويلات الاتجاهية, التجانس الهندسي, متعددة البيومترية, تصنيف, بصمة راحة اليد المتعددة الأظياف.

Résumé

La richesse des images naturelles rend la quête d'une représentation optimale dans le domaine de traitement de l'image et de vision par ordinateur un enjeu majeur pour les chercheurs qui sont souvent amenés à satisfaire un compromis entre l'efficacité et la complexité. Pour des raisons de simplicité et d'efficacité, les ondelettes ont souvent été utilisées de manière séparable, sur les axes horizontal et vertical, pour représenter l'image. Il en résulte une décorrélation partielle de l'image, qui se traduit par la présence de nombreux coefficients de forte énergie le long des contours. Donc, les ondelettes classiques ne sont pas capables de représenter la régularité des structures géométriques des images de façon efficace à cause de leur support carré. D'autre part, elles manquent des caractéristiques directionnelles pour l'image. Alors, les ondelettes permettent d'isoler des points de discontinuités mais elles ne permettent pas de distinguer les courbes régulières dans l'image. Les transformées directionnelles dérivées généralement de la transformée en ondelettes apportent une solution à cet enjeu, et permettent une représentation efficace de la régularité géométrique contenue réellement dans l'image. En effet, toute régularité géométrique est une forme de redondance qu'il s'avère intéressant d'exploiter pour une représentation creuse de l'image, ceci en vue de l'obtention de bonnes approximations avec peu de coefficients. Dans le cadre de cette thèse, deux applications ont été développées, la première concerne la compression progressive des images hautes résolutions (images satellitaires et images médicales) en utilisant les transformées Bandelettes et la seconde propose un système biométrique qui permet la reconnaissance d'images palmaires multispectrales et des images d'empreintes palmaire en utilisant les transformées non adaptatives. A cet effet, nous avons utilisé les transformées Ridgelets et la transformée de Radon pour l'extraction d'attributs discriminants afin de caractériser les modalités biométriques utilisées. Dans notre première contribution, nous explorons l'utilisation de la transformée de Radon pour la caractérisation des lignes principales de la main. Le processus de vérification est conduit en utilisant la distance de Hausdorff combinée à la triangulation de Delaunay. La deuxième contribution concerne l'identification des individus à partir de leurs empreintes palmaires multispectrales en se basant sur la transformée Ridgelets . A cet effet, nous avons proposé la fusion des informations au niveau pixel ou au niveau d'attributs des images palmaires et des images des veines de la paume, pour l'amélioration de l'efficacité et la robustesse du système biométrique proposé.

Mot-clé : *Transformées directionnelles, ondelettes, régularité géométrique, compression, biométrie, multibiométrie, classification, Authentification, Empreintes palmaires multispectrales.*

Abstract

The richness of natural images makes the quest for optimal representation in the field of image processing and computer vision a major challenge for researchers who are often required to meet a compromise between efficiency and complexity. For a sake of simplicity and efficiency, the wavelets have often been used separately on the horizontal and vertical axes to represent the image. This leads to a partial decorrelation of the image mirrored by the presence of numerous high energy coefficients along the contours. Hence, the classical wavelets failed to represent such singularities effectively because of their square support. The directional transforms commonly derived from the wavelet transform provide a solution to this limitation. Effectively, they allow an effective representation of the geometric regularity contained in the image. Indeed, any geometrical regularity is a form of redundancy that is worth exploiting for a sparse representation of the image to obtain good approximations with low coefficients. In this thesis, two applications have been developed, the former is the progressive compression of high resolution images (satellite images and medical images) using the bandelets transforms. The latter proposes a biometric system that allows recognition of palmprint and multispectral palmprint images using non-adaptative transforms. For this purpose, we have used ridgelets and Radon transforms to extract discriminating attributes that characterize the biometric modalities. In our first contribution, we have explored using the Radon transform for the characterization of the principal lines of the hand, then we have conducted a verification process using Hausdorff distance combined to Delaunay triangulation. The second contribution has do with identification of individuals using their multispectral palmprints. To this end we have proposed an approach based on ridgelets tranforms for palmprint and palm vein fusion in both pixel and feature levels to improve the efficiency and robustness of biometric systems.

Keyword: *directional transforms, wavelets, geometric regularity, compression, biometrics, multibiometrics, classification, Authentication, Multispectral palmprints.*

Glossaire

| | |
|----------|---|
| ACP | Analyse en Composantes Principales |
| ANN | Artificial Neural Network |
| CRT | Continuous Ridgelet Transform |
| DICOM | Digital Imaging and Communications in Medecine. |
| DWT | Discrete Wavelet Transform |
| EBCOT | Embedded Block Coder with Optimal Truncation |
| EER | Equal Error Rate |
| ENL | Equivalent Number of Looks |
| EZW | Embedded coding with Zerotree of Wavelet coefficients |
| FAR | False Accept Rate |
| FCM | Fuzzy C-Means |
| FRAT | Finite RAdon Transform |
| FRIT | Finite Ridgelet Transform |
| FRR | False Reject Rate |
| HH | High – High |
| HL | High – Low |
| JPEG | Joint Photographic Experts Group |
| JPEG2000 | Joint Photographic Experts Group 2000 |
| KNN | K Nearest Neighbors |
| LDA | Linear Discriminant Analysis |
| LH | Low – High |
| LL | Low – Low |
| LSB | Least Significant Bit |
| LMS | Least Mean Square |
| MSB | Most Significant Bit |
| MSE | Mean Squared Error |
| PMC | Perceptron MultiCouche |
| PSNR | Peak Signal-to-Noise Ratio |
| QMF | Quadrature Mirror Filters |
| RC | Rapport de Compression |
| RBF | Radial Basis Function |
| RLE | Run-Length-Encoding |
| RMSE | Root-Mean-Square Error |
| ROC | Receiver Operating Characteristic curve |
| ROI | Region Of Interest |
| SAR | Synthetic Aperature Radar |
| SNR | Signal-to-Noise Ratio |
| SPIHT | Set Partitioning In Hierarchical Trees |
| SRM | Structural Risk Minimisation |
| SVM | Support vector Machines |

Liste des tableaux

| | | |
|---------------------|---|-----|
| Tableau 1.1. | Etude comparative des transformées orientées | 30 |
| Tableau 2.1. | Résultats du PSNR (dB) pour la compression des images SAR avec différents bpp (nombre bits/pixel) | 55 |
| Tableau 2.2. | Mesure d'évaluation qualitative de la compression d'images SAR et Despeckling | 56 |
| Tableau 3.1. | Propriétés des diverses modalités biométriques | 59 |
| Tableau 3.2. | Quelques exemples de systèmes biométriques multimodaux avec leur technique de fusion | 70 |
| Tableau 4.1. | Matrice de contingence d'une classification binaire | 74 |
| Tableau 4.2. | Exemple de distances utilisées pour le classificateur K-NN | 82 |
| Tableau 4.3. | Fonction d'activation d'un modèle de neurone artificiel..... | 83 |
| Tableau 5.1. | Mesure de la similarité en utilisant la distance de Hausdorff pour différentes classes d'empreintes palmaires avec plusieurs acquisitions | 113 |
| Tableau 5.2 | Mesure de similarité des images bruitées avec différents SNR (dB) utilisant la distance de Hausdorff pour quelques images de la base | 114 |
| Tableau 5.3. | Comparaison et évaluation de plusieurs méthodes de fusion | 125 |
| Tableau 5.4 | Résultats expérimentaux pour KNN et SVM binaire sur la base CASIA | 127 |
| Tableau 5.5 | Résultats expérimentaux avec les SVMs binaires validés sur la base PolyU | 128 |
| Tableau 5.6 | Résultats expérimentaux avec la méthode OAO validés sur les deux bases | 128 |
| Tableau 5.7 | Confrontation de l'approche proposée à celles de l'état de l'art | 129 |

Liste des figures

| | | |
|---------------------|---|----|
| Figure 1.1. | Exemple d'un triplet de fonctions ondelettes en 2D | 9 |
| Figure 1.2. | Représentation en ondelette discrète d'une image d'empreinte palmaire | 10 |
| Figure 1.3. | (a): Diagramme bloc de la transformée Ridgelets, (b): Diagramme bloc de la FRIT ... | 12 |
| Figure 1.4. | Représentation des contours lisses avec les ondelettes à support carrée et la Curvelet à support allongé | 13 |
| Figure 1.5. | Organigramme de la transformée en curvelets 1ère génération | 15 |
| Figure 1.6. | (a): Illustration de l'implantation des contourlets à l'aide d'un banc de filtres pyramidal directionnel (PDFB), (b): Exemple de partitionnement fréquentiel par la transformée en Contourlets | 17 |
| Figure 1.7. | La transformée en Contourlets de l'image Polygone (512x512 où l'image est décomposée en un seul niveau pyramidal et quatre sous-bandes directionnelles. Les petits coefficients sont représentés en noir alors que les grands coefficients sont représentés en blanc | 18 |
| Figure 1.8. | Le pavage de la transformée Wave atoms dans le plan fréquentiel où seul le premier quadrant est représenté | 19 |
| Figure 1.9. | Exemple d'un contour C et son flot τ définie par sa tangente | 20 |
| Figure 1.10. | Exemple de modèle d'horizon et déformation du domaine selon un flot géométrique | 22 |
| Figure 1.11. | Flots géométriques en utilisant différentes segmentations dyadiques..... | 23 |
| Figure 1.12. | Exemple segmentation en carrées dyadiques d'une image avec le quadtree correspondant | 23 |
| Figure 1.13. | Algorithme de la transformée en bandelettes seconde génération | 24 |
| Figure 1.14. | (a): Dictionnaire d'edgels trouvé pour l'image, (b) : Transformée wedgelets : zones uniformes mises en évidence | 25 |
| Figure 1.15. | Illustration de l'interaction directionnelle lorsque le filtrage est effectué le long des lignes numériques avec filtrage passe-haut le long $-1/2$ | 26 |
| Figure 1.16. | Directions de Prédiction dans la transformée en ondelettes orientées , (a) : Directions exploitées sur maillage séparé en pair-impair ; (b) : Directions exploitées sur treillis quinconces, la recherche de directions se fait pour les deux treillis et (c): Directions exploitées sur treillis quinconces, l'information de direction pour la prédiction est directement exploitée pour la mise à jour | 28 |
| Figure 2.1. | Schéma globale de la compression des images basée sur les transformations | 33 |
| Figure 2.2. | Schéma de compression par Prédiction : Algorithme DPCM | 34 |
| Figure 2.3. | Quantification scalaire, (a) : uniforme ; (b): uniforme avec zone morte | 36 |
| Figure 2.4. | Représentation des plans de bits allant du bits le plus significatif (MSB) vers le moins significatif (LSB) pour une transformée en ondelettes 2D | 40 |
| Figure 2.5. | Relation entre les coefficients d'ondelettes dans les différents sous-bandes; (a): Ordre de balayage des coefficients , (b): La dépendances parents-enfants entre des coefficients d'ondelettes | 41 |
| Figure 2.6. | Etude comparative entre EZW et SPHIT, (a): PSNR en fonction bpp et (b) : analyse visuel pour différents PSNR | 42 |
| Figure 2.7. | Structure du codec EBDCOT, (a): Subdivision en blocs 64×64 des coefficients d'une transformée en ondelettes; (b) : Déplacement de la fenêtre de contexte lors du codage d'un plan de bits | 44 |
| Figure 2.8. | Définition d'un voisinage dans le domaine ondelettes | 45 |
| Figure 2.9. | Exemple d'une image avec sa régularité géométrique, (a) : image originale, (b): Décomposition en ondelette $j=2$ et (c) : zoom sur les coefficients non nuls le long de la géométrie | 47 |
| Figure 2.10. | Réordonnancement discret des points d'échantillonnage | 49 |
| Figure 2.11. | Processus de la transformée en bandelettes d'un bloc f de coefficients d'ondelettes | 50 |

| | |
|---|-----|
| Figure 2.12. Courbe PSNR- débit pour différentes images testées, (a) : image médicale et (b) image aérienne | 51 |
| Figure 2.13. Reconstruction de l'image par bandelette 2G et Ondelettes pour un débit bpp= 0.73 (bit/pixel): Problème d'artefact | 52 |
| Figure 2.14. Résultats comparatifs entre SPHIT+Bandelettes 2G et SPHIT + Ondelettes | 52 |
| Figure 2.15. Comparaison de la compression de l'image SAR à Bpp=1.96 bit /pixel (DWT PSNR= 35.86 dB, FRIT PSNR = 34.71 dB et 2GBandelet PSNR = 37.18dB) et à bpp= 0.28 bit/ pixel (DWT PSNR=19.55 dB, FRIT PSNR= 18.85 dB et 2G Bandelet PSNR = 24.10dB) | 54 |
| Figure 2.16. Problème d'artefact, (a) : dans DWT, (b): dans bandelette 2G et (c) : dans FRIT domaines | 54 |
| Figure 2.17. Résultats pour la compression des images SAR utilisant les transformées directionnelles | 55 |
| Figure 2.18. Résultats de compression obtenus pour l'image SAR SABR Radar homogène | 56 |
| Figure 3.1. Quelques modalités biométriques | 58 |
| Figure 3.2. Schéma fonctionnel d'un système biométrique | 60 |
| Figure 3.3. Distributions des taux de vraisemblance des utilisateurs légitimes et des imposteurs d'un système biométrique | 62 |
| Figure 3.4. Les différentes Sources dans les systèmes multibiométriques | 65 |
| Figure 3.5. Schéma de fusion niveau capteurs | 65 |
| Figure 3.6. Concaténation des vecteurs caractéristiques | 66 |
| Figure 3.7. Les différents niveaux de fusion dans un système multibiométrique | 69 |
| Figure 4.1. Représentation d'une courbe ROC | 75 |
| Figure 4.2. Illustration de la discrimination linéaire (LDA) pour 2 classes (Projection sur l'axe optimal) | 80 |
| Figure 4.3. Modèle d'un neurone artificiel | 83 |
| Figure 4.4. Modèle du Perceptron multicouche | 85 |
| Figure 4.5 Synoptique de la méthode de Back propagation | 86 |
| Figure 4.6. Séparation linéaire des classes carré (étiquetée +1), et triangle (étiquetée -1), respectivement positive et négative par un hyperplan. w représente le vecteur normal à l'hyperplan, les γ_i représentent la marge séparant chaque objet de l'hyperplan | 89 |
| Figure 4.7. Les objets triangles et carrés représentés en gras, situés respectivement sur les hyperplans $H_1: wx + b = -1$ et $H_2: wx + b = 1$ sont appelés supports vecteurs les objets x_i , et x_t sont considérés comme non classés, les objets positifs situés dans la partie entre H_2 et H_0 et les objets négatifs situés entre H_1 et H_0 | 91 |
| Figure 4.8. (a): représente un ensemble d'objets non linéairement séparables, (b): montre qu'après application de la transformation non linéaire, les objets en question pouvaient être séparés par une droite dans \mathbb{R}^2 | 93 |
| Figure 4.9. Principe de la méthode One-Against-One appliquées à trois classes | 95 |
| Figure 5.1. Diagramme bloc de l'algorithme proposé pour la vérification des images palmaires basée sur la transformée de Radon et la triangulation de Delaunay | 104 |
| Figure 5.2. Extraction du ROI, (a): image originale ; (b): image en niveaux de gris ; (c) : histogramme de l' image ; (d) : Binarisation de l' image ; (e) : extraction du contour de l'image (d) et (f) : Extraction du ROI | 105 |
| Figure 5.3. Représentation d'une ligne rectiligne ; (a) dans le domaine spatial et (b) dans le domaine de Radon | 106 |
| Figure 5.4. Système d'acquisition des images de la main, (a):acquisition d'une image palmprint (b): la distance utilisée dans l'acquisition | 110 |
| Figure 5.5. Transformation de Radon des lignes principales dans le ROI pour 2 classes de palmprints avec plusieurs acquisitions; (a) : Palmprint images; (b) : les lignes principales extraites; (c): Transformation de Radon et (d) : Extraction des centroids utilisant le seuillage et la binarisation des coefficients de Radon..... | 111 |

| | |
|---|-----|
| Figure 5.6. Histogramme des coefficients de Radon pour différentes projections($\theta=4,45,80,90^\circ$) | 111 |
| Figure 5.7. Exemples de signatures obtenues par triangulation sur trois individus | 112 |
| Figure 5.8. Invariance de la transformation à la rotation, (a): Coefficients de Radon ; (b): Rotation du ROI et $\theta = 4^\circ$ et 8° et (c): Extraction des caractéristiques et triangulation de Delaunay | 112 |
| Figure 5.9. Triangulation de Delaunay des lignes principales d'images bruitées avec plusieurs SNRs | 113 |
| Figure 5.10. L'algorithme de l'approche proposée basée sur la fusion niveau pixel des images palmaires multispectrales en utilisant les classificateurs SVM /KNN | 115 |
| Figure 5.11. Représentation multi-échelle de l'empreint palmaire, (a): décomposition DWT à niveau égal à 2 et (b): diagramme bloc de transformée en Ridgelets | 117 |
| Figure 5.12. Exemple des ROIs extraits dans la base de données, (a):ROI de l'empreinte palmaire, (b): ROI de l'image palmaire des veines et (c): ROI des images fusionnées. | 121 |
| Figure 5.13. Histogramme et valeurs du kurtosis et skewness de trois classes: (a) kurtosis=8.82, skewness=2.62, (b):kurtosis=6.9, skewness =2.29 et (c): kurtosis=8.6, skewness=2.50... | 122 |
| Figure.5.14. Coefficients des énergies des blocs pour; (a) : intra-class (3 acquisitions de la même Multispectral ROI) et (b) : Inter-class (3 différentes multispectral ROI) | 122 |
| Figure 5.15. Courbe ROC de l'approche proposée | 123 |

Table des matières

| | |
|---------------------|----|
| Glossaire | IV |
| Listes des tableaux | V |
| Liste des figures | VI |

| | |
|----------------------------|---|
| Introduction générale..... | 1 |
|----------------------------|---|

Chapitre 1

Représentation multi-échelle des images par des ondelettes géométriques

| | |
|--|----|
| 1.1. Introduction..... | 5 |
| 1.2. Représentation par Ondelettes | 6 |
| 1.2.1. Transformée en Ondelette continue | 6 |
| 1.2.2. Analyse des singularités ponctuelles par la transformée en ondelettes | 7 |
| 1.2.3. Transformée en ondelettes discrètes | 8 |
| 1.2.4. Extension de la transformée en ondelette aux signaux 2D | 9 |
| 1.3. Transformées géométriques à base fixe (non adaptatives) | 11 |
| 1.3.1. Transformée en Ridgelets | 11 |
| 1.3.2. Transformée en Curvelets | 13 |
| 1.3.2.1. La transformée en Curvelets première génération (Curvelet 1G)..... | 14 |
| 1.3.2.2. La transformée en Curvelets seconde génération (Curvelet 2G) | 15 |
| 1.3.3. Transformée en Contourlets | 16 |
| 1.3.4. Transformée Wave atoms | 18 |
| 1.4. Transformées géométriques adaptatives..... | 19 |
| 1.4.1. La Transformée en Bandelettes | 19 |
| 1.4.1.1. Définition de la géométrie | 20 |
| 1.4.1.2. Les bandelettes de la première génération (Bandelet 1G) | 21 |
| 1.4.1.3. Les bandelettes de la seconde génération (Bandelet 2G) | 23 |
| 1.4.2. Les transformées Wedgelets | 24 |
| 1.4.3. Les transformées Directionlets | 25 |
| 1.4.4. Les ondelettes orientées | 26 |
| 1.5. Etude comparative des différentes transformées directionnelles | 28 |
| 1.6. Conclusion | 28 |

Chapitre 2

Transformées en Bandelettes pour la compression des images haute résolution

| | |
|--|----|
| 2.1. Introduction et état de l'art sur la compression..... | 31 |
| 2.1.1 La décorrélation | 33 |
| 2.1.1.1. Prédiction | 33 |
| 2.1.1.2. Transformation | 34 |
| 2.1.2. Quantification | 35 |
| 2.1.3. Codage entropique | 36 |
| 2.1.4. Mesures des performances en compression | 37 |
| 2.1.4.1. Taux de compression et débit binaire | 37 |
| 2.1.4.2. Le compromis débit-distorsion | 37 |
| 2.2. Compression progressive des images basée sur la transformée en ondelettes | 39 |
| 2.2.1. Introduction | 39 |
| 2.2.2. Codage Inter sous-bandes | 40 |
| 2.2.2.1. L'algorithme EZW | 40 |
| 2.2.2.2. L'algorithme SPIHT | 42 |

| | | |
|----------|--|----|
| 2.2.3. | Codage Intra sous-bande | 43 |
| 2.2.4. | Exploitation de la dépendance entre les coefficients d'ondelettes | 44 |
| 2.3. | Compression des images basé sur l'utilisation de la transformée en bandelettes | 46 |
| 2.3.1. | Introduction | 46 |
| 2.3.2. | Construction de bases de bandelettes par optimisation débit-distorsion | 47 |
| 2.3.2.1. | Transformée en Bandelettes d'ordre 0 | 47 |
| 2.3.2.2. | Transformée en Bandelettes par regroupement | 50 |
| 2.3.2.2. | Études expérimentales des performances de la transformée en bandelettes d'ordre 0 | 50 |
| 2.4. | Etude comparative de la compression des images satellitaires basée sur les transformations directionnelles | 53 |
| 2.5. | Conclusion et perspectives | 57 |

Chapitre 3

Introduction à la multibiométrie

| | | |
|----------|---|----|
| 3.1. | Généralité sur le système biométrie | 58 |
| 3.2. | Fonctionnement d'un système biométrique | 59 |
| 3.2.1. | L'enrôlement | 59 |
| 3.2.2. | La vérification | 59 |
| 3.2.3. | L'identification | 60 |
| 3.3. | Performance d'un système biométrique | 61 |
| 3.4. | La multibiométrie | 62 |
| 3.5. | Niveau de Fusion d'information dans le système multibiométrique | 64 |
| 3.5.1. | Fusion pré-classification | 64 |
| 3.5.1.1. | Fusion niveau capteur | 65 |
| 3.5.1.2. | Fusion niveau caractéristiques | 65 |
| 3.5.2. | Fusion post-classification | 66 |
| 3.5.2.1. | Fusion niveau score | 66 |
| 3.5.2.2. | Fusion niveau décision | 68 |
| 3.5.2.3. | Fusion niveau rang | 69 |
| 3.6. | Conclusion | 70 |

Chapitre 4

Technique de classification pour la biométrie

| | | |
|----------|--|----|
| 4.1. | Introduction..... | 71 |
| 4.2. | Aptitude de généralisation et mesures de performance | 72 |
| 4.3. | Approches paramétriques | 75 |
| 4.3.1. | Classification naïve Bayésienne | 75 |
| 4.3.2. | Analyse discriminante Linéaire (LDA) | 77 |
| 4.4. | Approches non paramétrique | 81 |
| 4.4.1. | Classificateur K-NN | 81 |
| 4.4.2. | La classification en utilisant les réseaux de neurones | 83 |
| 4.4.2.1. | Structure d'un neurone formel | 83 |
| 4.4.2.2. | Architecture d'un réseau de neurones | 84 |
| 4.4.2.3. | Algorithme de BACK-PROPAGATION | 85 |
| 4.4.3. | Les Machines à Vecteurs Supports (SVM)..... | 88 |
| 4.4.3.1. | Le concept | 88 |
| 4.4.3.2. | Extension des SVMs au cas multi-classes | 95 |
| 4.4.3.3. | Les SVM dans la pratique | 96 |
| 3.5. | Conclusion | 97 |

Chapitre 5

Les transformées géométriques pour l'identification/vérification des images palmaires

| | | |
|------|---|-----|
| 5.1. | Introduction | 99 |
| 5.2. | Etat de l'art de l'identification/vérification des images palmaires | 100 |

| | |
|---|------------|
| 5.3. Contributions | 103 |
| 5.3.1. La transformée de Radon pour la caractérisation des images palmaires | 103 |
| 5.3.1.1. Approche proposée | 103 |
| 5.3.1.2. Extraction des régions d'intérêt | 104 |
| 5.3.1.3. Extraction des lignes et Caractérisation dans le domaine de Radon | 105 |
| 5.3.1.4. Mesure de Similarité utilisant la distance de Hausdorff | 109 |
| 5.3.1.5. Expérimentation et résultats | 110 |
| 5.3.2. La transformée Ridgelets pour la caractérisation des images palmaires multispectrales | 114 |
| 5.3.2.1. Fusion des images palmaires par les ondelettes | 115 |
| 5.3.2.2. Extraction des caractéristiques des images palmaires multispectrales | 117 |
| 5.3.2.3. Résultats expérimentaux | 120 |
| 5.4. Conclusion | 130 |
| Conclusion générale et perspectives | 131 |
| Références | 134 |

Introduction Générale

La régularité géométrique des images naturelles est un élément important qui permet à notre cerveau de faire une représentation efficace et compacte des scènes que nous percevons [1-8]. En effet, les études de D.H. Hubel et T.N. Weisel [2] ont montré que toute région du cortex analyse le champ visuel en termes de directions "Lumière-obscurité" autour des contours, en particulier autour de leurs tangentes. A cet effet, ils ont établi que la transformation effectuée par les neurones peut être considérée comme une décomposition de l'image sur des ondelettes orientées. Un modèle de l'image naturelle fût proposé par Mallat [5-8], il présente la particularité de ne pas être seulement une simple fonction $L^2([0,1]^2)$, mais aussi de comporter en plus de la texture et du bruit, une composante géométrique qui contient l'essentiel des structures géométriques. Entre autres, leur étude suggère que toute régularité géométrique est une forme de redondance qu'il faut exploiter pour une représentation optimale de l'image. Par ailleurs, la régularité locale d'une fonction monodimensionnelle peut-être définie par l'exposant de Lipschitz qui mesure son degré et indique la présence ainsi que le type de la singularité [9-11]. Généralement, le calcul de la transformée en ondelettes permet d'évaluer le degré de Lipschitz et de détecter les points singuliers de la fonction, tout en considérant le choix du nombre de moments nuls. Ainsi, La transformée en ondelette est l'une des premières représentations temps-fréquence, multirésolution à support compact. En effet, elle demeure une alternative à la transformée de Fourier pour la représentation des signaux réguliers par morceaux monodimensionnels, permettant une localisation plus efficace des singularités ponctuelles en raison de son support isotrope [5-11]. Cependant, l'inconvénient majeur de la transformée en ondelette est qu'elle ne parvient pas à saisir la régularité géométrique le long des singularités curvilignes (bord ou contour) dans les signaux bidimensionnels qui définissent l'image. Pour cela, de nouvelles transformées ont été proposées par de nombreux chercheurs combinant l'aspect multi-échelles à l'aspect directionnel, il s'agit des ondelettes de seconde génération qui utilisent avantageusement les structures géométriques contenues dans les images. Par ailleurs, le but de ces transformées est d'améliorer la concentration de l'énergie dans un minimum de coefficients pour une représentation géométrique parcimonieuse et efficace à l'amélioration des algorithmes de traitement d'images tels que la compression, la caractérisation des structures contenues dans l'image, le débruitage, etc..

Parmi ces nouvelles transformées directionnelles, il existe dans la littérature deux approches différentes. La première consiste en les transformées géométriques adaptatives

généralement non redondantes, et dont les fonctions de base sont choisies pour mieux s'adapter à la géométrie d'une image donnée, telle que les transformées en bandelettes [6-8]. La deuxième approche concerne les transformées non-adaptatives généralement redondantes qui utilisent des bases de projection fixes dont le but est de corriger l'aspect isotrope des transformées en ondelettes classiques, telle que la transformée en Curvelets [12] et la transformée en Contourlets [13].

Les travaux développés dans le cadre de cette thèse reposent essentiellement sur l'utilisations de ces transformées géométriques, d'abord pour la compression progressive des images hautes résolution, puis dans la biométrie en considérant les empreintes palmaires (*Palmprint*) et les empreintes palmaires multispectrales (*Multispectral palmprint*) comme modalités.

Contributions

Au cours de cette thèse, nous avons réalisé les contributions suivantes :

Partie compression

- Notre contribution consiste à utiliser la transformée en bandelettes d'ordre 0 à la compression des images SAR (*Synthetic Aperature Radar*) et médicales, dans laquelle nous avons présenté une étude comparative entre plusieurs transformées directionnelles [14]. Par ailleurs, dans une optique d'évaluation de la robustesse de cette transformée, nous avons proposé dans cette thèse un schéma de compression progressif en combinant la transformée en bandelettes avec l'encodeur SPIHT.

Partie Biométrie palmaire

- Dans la première contribution nous avons utilisé la transformée de Radon dans le but de la caractérisation des lignes principales des empreintes palmaires, à cet effet nous nous sommes basés sur l'encodage de leurs positions dans le domaine de Radon [15]. Ceci a permis de construire une structure stable et robuste au bruit propre à chaque individu.
- Dans la deuxième contribution, nous avons utilisé la transformée en Ridgelets par bloc pour caractériser les images des empreintes palmaires multispectrales (empreintes palmaires et veines palmaires). Par ailleurs, nous avons exploré deux types de classificateurs à savoir: les k-plus proches voisins et les machines à supports vecteurs pour les phases d'identification et authentification [16,17].

- Dans la dernière contribution en vue d'améliorer les performances du système biométrique basé sur les empreintes palmaires multispectrales, nous avons proposé d'explorer deux niveaux de fusion : le niveau pixel et le niveau attribut. A cet effet, nous nous sommes basé sur une analyse statistique et une analyse par transformée en Ridgelets par bloc des images multispectrales [18].

Organisation du manuscrit

Le premier chapitre traitera des transformées directionnelles dérivées généralement de la transformée en ondelettes. Ces transformées permettent une représentation efficace de la régularité géométrique contenue réellement dans l'image. En effet, toute régularité géométrique est une forme de redondance qu'il s'avère intéressant d'exploiter pour une représentation creuse de l'image en vue de l'obtention de bonnes approximations avec peu de coefficients. Donc, dans ce chapitre, nous présenterons tout d'abord les transformées non adaptatives pour lesquelles la décomposition ne dépend pas du signal à analyser, telle que la transformée en Ridgelets, La transformée en Curvelets et la transformée en Contourlets. Ensuite, nous présentons les transformées géométriques adaptatives qui dépendent du signal analysé, pour lesquelles une étape préalable d'estimation de la géométrie du signal est nécessaire avant de procéder à la décomposition. Les approches les plus intéressantes développées dans ce chapitre sont : les transformées en bandelettes, les wedgelets et les directionlets.

Le deuxième chapitre portera sur nos travaux sur la compression dans lesquels nous exploitons les transformées en ondelettes directionnelles pour la compression des images. D'abord, nous commençons par la transformée en ondelettes utilisée dans les systèmes de compression les plus récents. Par ailleurs, nous détaillerons quelques encodeurs progressifs inter-bandes et intra-bandes. Puis, nous mettons en évidence les corrélations et l'information résiduelle entre les coefficients d'ondelettes pour expliquer les limites de cette transformée pour la compression, en plus de son incapacité de représenter efficacement les régularités géométriques. Plusieurs familles d'algorithmes de codage par plan de bit emploient différentes techniques pour exploiter ces dépendances. En particulier, nous analyserons les performances des codeurs SPIHT et EZW. Finalement, nous présenterons la transformée en bandelettes qui possède des propriétés intéressantes pour la compression vu qu'elle exploite les régularités géométriques contenues dans les images. La transformée en bandelette d'ordre 0 sera utilisée pour la compression des images SAR et les images médicales. Par ailleurs, une étude comparative des résultats obtenus avec les différentes techniques seront présentés dans ce chapitre.

Le troisième chapitre portera sur des généralités sur la biométrie, en se basant essentiellement sur la biométrie palmaire qui fera l'objet de notre étude. Ensuite, nous introduirons la multibiométrie qui permet de palier aux limitations d'un seul système biométrique en fusionnant des informations discriminantes issues de plusieurs modalités biométriques.

Le quatrième chapitre portera sur l'étude de quelques techniques de classification parmi les plus utilisées en biométrie, la liste de ces techniques n'est pas exhaustive car elle ne constitue pas en soit l'objet de notre étude. Par ailleurs, cela a comme priorité d'abord d'exposer avec plus de détails les notions utilisées dans le cinquième chapitre. Plus précisément, le principe de base de chaque méthode sera présenté et discuté d'une manière synthétique. Nous nous proposons d'aborder les méthodes de classifications dites paramétriques et les méthodes non paramétriques.

Le cinquième chapitre portera sur nos contributions relatives au domaine de la biométrie; dans lesquels nous soulignons l'efficacité et la robustesse des systèmes biométriques proposés à base d'images d'empreintes palmaires (*palmpoint*) et d'images d'empreintes palmaires multispectrales. A cet effet, nous avons utilisé les transformées géométriques pour l'extraction d'attributs discriminants pour la caractérisation des modalités biométriques utilisées. Dans notre première contribution, nous explorons l'utilisation de la transformée de Radon pour la caractérisation des lignes principales de la main. Le processus de vérification est conduit en utilisant la distance de Hausdorff combinée à la triangulation de Delaunay. La deuxième contribution concerne l'identification des individus à partir de leurs empreintes palmaires multispectrales en se basant sur la transformée en Ridgelets . A cet effet, nous avons proposé la fusion des informations au niveau pixel et au niveau des attributs des images palmaires et des images des veines de la paume, pour l'amélioration de l'efficacité et la robustesse du système biométrique proposé.

Nous terminons ce manuscrit par une conclusion dans laquelle nous présentons une synthèse sur les travaux réalisés au cours de cette thèse et les perspectives qui peuvent en découler.

Chapitre 1

Représentation multi-échelle des images par des ondelettes géométriques

1.1. Introduction.

La construction d'une représentation optimale d'images avec peu de coefficients significatives, nécessite l'exploitation des structures géométriques contenues dans ces dernières. Généralement, ces représentations sont réalisées via des transformations inversibles non redondantes et directionnelles, plus adaptées à l'extraction de structures géométriques lisses et continues, telles que les contours d'objets. La transformée de Fourier est l'une des premières représentations creuses du signal, elle consiste à décomposer le signal à partir d'une famille de sinusoides dont le support correspond au support du signal analysé. Bien que cette représentation est idéale pour les signaux réguliers, elle ne l'est pas aux signaux contenant des singularités [6-8,11,12]. En effet, ceci s'explique par le fait qu'elle n'arrive pas à modéliser le comportement de l'information sur la singularité, ce qui correspond à une dispersion d'énergie. Par ailleurs, la transformée de Fourier ne comporte aucune donnée sur l'information spatiale du signal. La transformée en ondelette (*Wavelet Transform*) est l'une des premières représentations temps-fréquence, multirésolution à support compact qui devient une alternative à la transformée de Fourier [5,19-21]. C'est une transformée idéale pour la représentation des signaux réguliers par morceaux monodimensionnels avec une localisation plus efficace des singularités ponctuelles en raison de son support isotrope. Cependant, l'inconvénient majeur de la transformée en ondelette est qu'elle ne parvient pas à saisir la régularité géométrique le long des singularités de bord ou contour dans les signaux bidimensionnels qui définissent l'image [5-13]. Les bases orthonormales sont particulièrement pratiques pour construire des approximations concises des images. Cependant, les représentations d'images dans des bases orthonormées séparables telles que Fourier, cosinus local ou ondelettes ne peuvent pas profiter de la régularité géométrique des structures de l'image. De nouvelles transformées ont été proposées par de nombreux chercheurs combinant des aspects multi-échelles à des aspects directionnels et fréquentiels, il s'agit des ondelettes de seconde génération qui utilisent avantageusement les structures géométriques contenues dans les images. Le but de ces transformées est d'améliorer la concentration de l'énergie dans un minimum de coefficients de la transformée pour aller dans le sens d'une

représentation creuse. Cette concentration correspond à une amélioration de la décroissance de l'erreur quadratique entre l'image initiale et l'image reconstruite.

Parmi ces nouvelles représentations directionnelles cherchant à remédier aux problèmes induits par les ondelettes 2D séparables, on peut trouver deux approches différentes dans la littérature : les transformées adaptatives (basées sur un dictionnaire de fonctions qui s'adaptent à la régularité géométrique) et les transformées à bases fixes ou non-adaptatives (basées sur un dictionnaire de fonctions fixes). Dans le reste de ce chapitre; on présentera un état de l'art des transformées directionnelles les plus utilisées dans les processus de traitement de l'image.

1.2. Représentation par Ondelettes

1.2.1. Transformée en ondelette continue

La transformée en ondelettes continues permet la représentation des signaux monodimensionnels réguliers par morceaux (de classe C^n avec $n \in \mathbb{N}$) et $\in L^2(\mathbb{R})$ dans une base bien localisée dans les domaines spatial et fréquentiel [510,19-23]. Cette base est construite à partir d'une unique fonction $\psi(x)$ appelée ondelette mère qui doit être localisée (c'est à dire à support compact ou à une décroissance rapide à l'infini) et oscillante. La fonction $\psi(x)$ dilatée par un facteur d'échelle a et translatée par des facteurs b , forme une base de $L^2(\mathbb{R})$, définie par :

$$\left(\frac{1}{\sqrt{a}} \psi \left(\frac{x-b}{a} \right) \right)_{a \in \mathbb{R}_+^*, b \in \mathbb{R}} \quad (1.1)$$

Ainsi, une transformée en ondelettes continue Wf de la fonction f , peut-être définie de l'espace $L^2(\mathbb{R})$ vers $L^2(\mathbb{R}_+^* \times \mathbb{R})$, par l'expression suivante:

$$Wf(a,b) = \langle f, \psi_{a,b} \rangle = \int_{\mathbb{R}} f(\tau) \tilde{\psi}_{a,b}(x) d\tau \quad (1.2)$$

avec $\tilde{\psi}_{a,b}$: l'ondelette duale de l'ondelette analysante, $Wf(a,b)$: le coefficient d'ondelettes de la fonction et $\langle \cdot, \cdot \rangle$: désigne le produit scalaire.

Pour une construction du signal sans perte d'information i.e. conservation d'énergie du signal, il faut que l'ondelette soit réversible, c'est à dire qu'elle doit vérifier la condition d'admissibilité

$C_\psi = \int_{\mathbb{R}} \frac{|\widehat{\psi}(\omega)|^2}{\omega} d\omega < +\infty$; avec $\widehat{\psi}(\omega)$: étant la transformée de Fourier de l'ondelette mère. Cette

condition implique que son moment d'ordre 1 (moyenne) soit nul i.e.:

$$\int_{\mathbb{R}} \psi(x) dx = 0 \quad (1.3)$$

ce qui justifie que l'ondelette soit oscillante et localisée en temps. D'une façon général, l'ondelette mère est caractérisée par son nombre de moments nuls M , défini par :

$$\int_{\mathbb{R}} x^m \psi(x) dx = 0, \quad \text{pour } m = 0 : M - 1 \quad (1.4)$$

ce qui constitue une caractéristique importante pour l'estimation de la régularité locale. Etant donné que la fonction ondelette est réversible, sa fonction originale $f \in L^2(\mathbb{R})$ peut être reconstruite à partir de ces coefficients d'ondelettes à l'aide de l'équation ci-après :

$$f(x) = \frac{1}{C_\psi} \iint_{\mathbb{R}^2} Wf(a,b) \psi_{a,b}(x) db \frac{da}{a^2} \quad (1.5)$$

1.2.2. Analyse des singularités ponctuelles par la transformée en ondelettes

En pratique, le calcul d'une transformation en ondelette sous la condition du choix du nombre de moments nuls permet de détecter les points singuliers d'une fonction f , et d'évaluer l'exposant de Lipschitz qui peut mesurer le degré de régularité locale, et éventuellement indiquer le type de la singularité [5-10]. Formellement, une fonction $f(x) \in L^2(\mathbb{R})$ admet une régularité Lipschitzienne localisée en x_0 , s'il existe une constante $K > 0$ telle que :

$$\forall x \rightarrow x_0 \text{ alors } |f(x) - f(x_0)| < K \|x - x_0\|^\alpha \quad (1.6)$$

avec $0 < \alpha < 1$, si $\alpha \neq 1$ alors, on dit alors que la fonction est singulière en x_0 .

Il est intéressant de remarquer que cette notion de régularité ne privilégie pas de direction particulière. Par conséquent, ceci justifie l'emploi dans les sections suivantes d'autres transformations dont la propriété est d'abord de chercher certaines directions particulières pour enfin estimer la régularité le long de cette direction. La transformée en ondelettes permet la détection, la localisation des singularités et l'accès à l'exposant de Lipschitz des points singuliers. Les singularités de la fonction analysée se traduisent par des points d'amplitude maximale dans sa transformée en ondelettes sur une large gamme d'échelles. En ne retenant que les points qui sont des maxima locaux de la transformée en ondelette, on obtient des chaînes de maxima pointant

vers les points singuliers de la fonction. Par conséquent, il existe une constante S telle que, lorsque l'échelle a tend vers 0, on a :

$$|Wf(x, a)| \leq S.a^\alpha \quad (1.7)$$

1.2.3. Transformée en ondelettes discrètes

L'ondelette continue est en théorie infiniment redondante du fait que l'ondelette est traduite de manière continue, cependant, il existe des méthodes pour diminuer cette redondance, parmi ces méthodes il en existe une qui consiste en l'emploi de la transformée en ondelettes discrète. J. Morlet [24] a proposé de construire des bases ou des frames de fonctions construits par une discrétisation dyadique des deux paramètres (a et b) sur le modèle suivant :

$$\psi_{j,k}(x) = a_0^{-j/2} \psi\left(\frac{x - k.b_0.a_0^{-j}}{a_0^j}\right)_{j=1,2,\dots,J-1 \text{ et } k=1,2,\dots,2^{j-1}} \quad (1.8)$$

où j et k sont des entiers. Généralement, $b_0 = 1$ et $a_0 = 2$, ces valeurs sont normalisées dans le but d'avoir des bases orthonormées d'ondelettes, et les coefficients seront alors décorrélés. Ceci permettra de réduire le volume d'information redondante. La base orthonormée sera reformulée par :

$$\left(\frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) = \frac{1}{2^{j/2}} \psi\left(\frac{x-2^j k}{2^j}\right) = \psi_{j,k}(x)\right)_{j \in \mathbb{Z}; k \in \mathbb{Z}} \quad (1.9)$$

La transformée en ondelette discrète (DWT: *Discret Wavelet Transform*) d'une fonction $f(x) \in L^2(\mathbb{R})$, engendre un ensemble de coefficients d'ondelettes dits détails; définis par :

$$d_{DWT}(j, k) = \langle f(x), \psi_{j,k}(x) \rangle = 2^{-j/2} \int_{\mathbb{R}} f(x) \psi\left(\frac{x-2^j k}{2^j}\right) dx \quad (1.10)$$

L'analyse multirésolution introduite par Mallat [5] et Meyer [20] a la capacité de reconstruire le signal. Quand le signal $f(x) \in L^2(\mathbb{R})$ est représenté dans différentes résolutions, le signal reconstruit contient différentes bandes de fréquences du signal original. La représentation basse résolution de $f(x)$ est construite par la fonction échelle $\phi_{j,k}(x)$ en projetant successivement le signal sur un sous-espace V_j (plus j augmente plus V_j contient des fonctions dont la composante spectrale est de fréquence élevée), ce qui est similaire à un filtrage passe-bas de $f(x)$. Par ailleurs, la différence entre deux approximations successives représente le détail perdu qui se

trouve dans un sous-espace W_j construit par la base d'ondelette et qui est complémentaire à V_j ; telle que $V_{j-1} = V_j \oplus W_j$. L'ensemble des coefficients d'approximation est définie formellement par :

$$c_{j,k} = \left\langle f(x), \phi_{j,k}(x) \right\rangle = \int_{\mathbb{R}} f(x) \cdot \phi_{j,k}(x) dx \quad (1.11)$$

A chaque échelle j fixée, les ondelettes $\psi_{j,k}(x)$ forment une base d'un espace détail W_j alors que l'espace V_j d'approximation est engendré par les fonctions échelles $\phi_{j,k}(x)$.

1.2.4. Extention de la transformée en ondelette aux signaux 2D

La représentation des signaux 2D (image), nécessite l'utilisation des ondelettes de manière séparable sur les axes horizontal et vertical pour des raisons de simplicité et d'efficacité. Les bases d'ondelettes de $L^2([0,1]^2)$ sont obtenues par translations et dilatations de trois ondelettes élémentaires $\{\psi_{j,k}^H(x, y), \psi_{j,k}^V(x, y), \psi_{j,k}^D(x, y)\}$ qui oscillent dans les directions horizontale, verticale et diagonale. Ces ondelettes bidimensionnelles séparables ont prouvé leurs capacités à détecter les singularités horizontales, verticales, ou ponctuelles (assimilées à des singularités diagonales). Elles s'obtiennent par un produit séparable d'ondelettes monodimensionnelles comme suit :

$$\begin{cases} \psi_{j,k}^1(x, y) = \psi_{j,k}^H(x, y) = \psi_{j,k}(x) \cdot \phi_{j,k}(y) \\ \psi_{j,k}^2(x, y) = \psi_{j,k}^V(x, y) = \phi_{j,k}(x) \cdot \psi_{j,k}(y) \\ \psi_{j,k}^3(x, y) = \psi_{j,k}^D(x, y) = \psi_{j,k}(x) \cdot \psi_{j,k}(y) \end{cases} \quad (1.12)$$

avec $\psi_{j,k}(\cdot)$ et $\phi_{j,k}(\cdot)$ représentant respectivement, l'ondelette et la fonction échelle monodimensionnelle associée. La Figure 1.1 illustre un exemple d'ondelettes 2D.

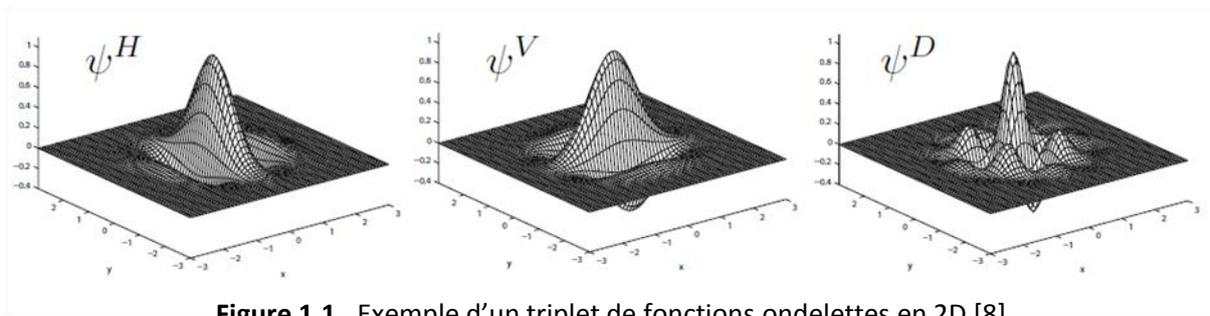


Figure 1.1. Exemple d'un triplet de fonctions ondelettes en 2D [8]

Par ailleurs, à un niveau de résolution J donné, toute fonction $f(x, y)$ peut être décomposée comme suit :

$$f(x, y) = \sum_k c_{J,k} \phi_{J,k}(x, y) + \sum_{j=1}^J \sum_n \sum_k d_{j,k} \psi_{j,k}^n(x, y) \quad (1.13)$$

où $\phi_{J,k}(x,y)$ est la fonction échelle définie par $\phi_{J,k}(x,y) = \phi_{J,k}(x) \cdot \phi_{J,k}(y)$; $n = 1, 2, 3$; $c_{J,k}$: les coefficients d'approximation et d_{jk} : les coefficients d'ondelettes.

La figure 1.2 ci-après illustre un exemple de décomposition à un niveau de résolution $j = 2$, d'une image palmaire de la base d'image d'empreintes palmaires CASIA.

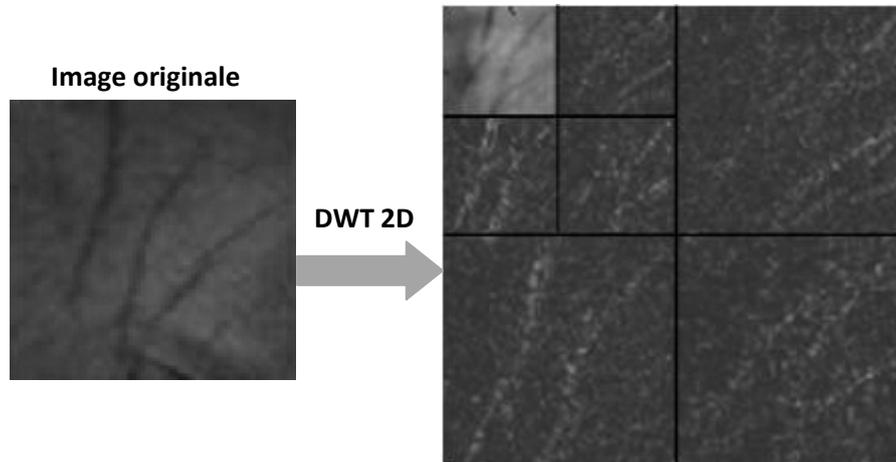


Figure 1.2. Représentation en ondelette discrète d'une image d'empreinte palmaire

En dépit de son succès, la transformée en ondelettes 2-D standard n'est pas complètement adaptée pour représenter des contours ou des textures orientées d'une image. La transformée en ondelettes bidimensionnelle séparable a prouvé son efficacité optimale pour la représentation des singularités horizontales, verticales ou ponctuelles. Cependant, elle fait montre de difficultés à capturer la régularité géométrique le long des singularités de bords ou les contours en raison de leur soutien isotrope. Ainsi, les coefficients d'ondelettes de tels contours vont se retrouver de façon redondante dans les différentes sous-bandes d'une même résolution. Cette information redondante contenue dans l'image, dont la régularité géométrique est un exemple, permet de caractériser les structures géométriques existantes dans les différentes images. Aussi, le nombre d'orientations qui est une caractéristique très importante pour les signaux multidimensionnels est limité et fixe.

La représentation de l'image dans des bases orthonormées séparables telles que Fourier, cosinus local ou ondelettes ne peuvent pas tirer profit de la régularité géométrique des structures d'image. Par conséquent, l'enjeu majeur est de trouver des représentations géométriques efficaces pour l'amélioration de l'efficacité des algorithmes de traitement d'images. Afin de représenter ces contours avec plus de parcimonie, de nouvelles transformées dérivées des ondelettes ont été proposées pour exploiter la régularité géométrique le long des bords. Dans ce contexte les chercheurs Do et Vetterli [25-26] ont listé les propriétés recherchées suivantes :

1. la multirésolution ;
2. la localisation spatiale et fréquentielle des éléments de base ;
3. l'échantillonnage critique pour garantir la compacité en un nombre minimal d'échantillons de représentation;
4. la multidirectionnalité : les éléments de base doivent être orientés dans de multiples directions; plus que les trois directions offertes par les ondelettes séparables ;
5. l'anisotropie : Pour capturer les contours lisses dans les images, la représentation doit inclure des fonctions de formes plus ou moins allongées qui sont pratiquement parallèles aux bords.

Parmi ces nouvelles transformées directionnelles, il existe dans la littérature : les transformées géométriques adaptatives, dont les fonctions de base sont choisies pour mieux s'adapter à la géométrie d'une image donnée, et les transformées à bases fixes ou non-adaptatives qui utilisent des bases de projection fixes, mais qui cherchent à corriger l'aspect isotrope des transformées en ondelettes classiques.

1.3. Transformées géométriques à base fixe (non adaptative)

1.3.1. Transformée en Ridgelets

La transformée en Ridgelets développée par Candes et Donoho en 1999 [27-29] compte parmi les premières transformées en ondelettes géométriques; où les éléments de base sont des ondelettes géométriques anisotropes. Cette transformée dotée d'une sélectivité directionnelle arbitraire, fournit une clé pour l'analyse des singularités à hautes dimensions. Elle correspond à la construction d'un dictionnaire de fonctions d'analyse s'appuyant sur la transformée de Radon (voir figure 1.3 (a)). La transformation en Ridgelets est une analyse en ondelettes dans le domaine de Radon, en effet, des discontinuités linéaires sont projetées comme des singularités ponctuelles par la transformée de Radon. L'application de la transformée en ondelettes sur les différentes projections de Radon, permet d'obtenir un codage optimal du contour de l'objet dans une image. Par ailleurs, l'intérêt de cette transformée est de permettre de concentrer efficacement l'énergie d'un signal régulier par morceaux sur quelques coefficients, ainsi le coefficient le plus significatif représente les directions des lignes d'image ayant la plus forte magnitude. Contrairement à la transformée en ondelettes, cette transformée permet de prendre en charge des ruptures 2D linéaires. Fondamentalement, la transformée en Ridgelet a été introduite pour la prise en charge des fonctions dans les espaces continus. La transformée en Ridgelet Continue (*CRT: Continuous Ridgelet Transform*) est définie par [27-30] :

$$CRT_f(a,b,\theta) = \int_{\mathbb{R}^2} f(x) \psi_{a,b,\theta}(x) dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2) \psi_{a,b,\theta}(x_1, x_2) dx_1 dx_2 \quad (1.14)$$

où $a = 2^j$ est le facteur échelle, b : le facteur de translation et $\psi_{a,b,\theta}(x)$: La fonction bidimensionnelle définie à partir de la fonction ondelette mère monodimensionnelle $\psi(x)$ orientée selon un angle de projection θ , et constante le long des droites d'équations $x_1 \cos \theta + x_2 \sin \theta = \text{constante}$. Elle est définie par :

$$\psi_{a,b,\theta}(x) = a^{-1/2} \psi((x_1 \cos \theta + x_2 \sin \theta - b) / a) \quad (1.15)$$

Formellement, la transformée de Radon est définie par l'équation :

$$R_f(\theta, \tau) = \int_{\mathbb{R}^2} f(x) \delta(x_1 \cos \theta + x_2 \sin \theta - \tau) dx \quad (1.16)$$

Par conséquent, la CRT est finalement formulée par :

$$CRT_f(a,b,\tau) = \int_{\mathbb{R}} \psi_{a,b}(\tau) R_f(\theta, \tau) d\tau \quad (1.17)$$

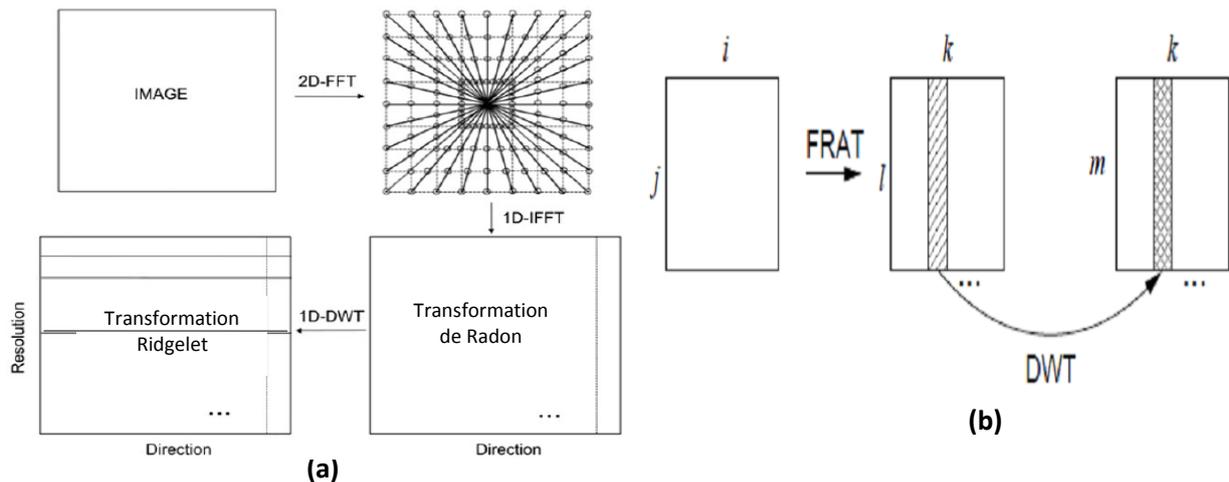


Figure 1.3. (a) : Diagramme bloc de la transformée Ridgelets, (b) : Diagramme bloc de la FRIT

Dans le domaine du traitement d'image, il est nécessaire d'avoir une version orthonormée et discrète de la transformée en Ridgelets pour les images numériques de taille finie [31]. La Transformée en Ridgelet Finie (FRIT) est accomplie via la Transformée de Radon Finie (FRAT : Finite Radon Transform), définie comme la somme des valeurs de pixels de l'image le long d'un ensemble de lignes. Considérons l'ensemble $Z_p = \{0, 1, \dots, p - 1\}$ où p est un nombre premier; la FRAT d'une fonction réelle $f(x, y)$ sur la treillis bornée Z_p^2 , est définie par :

$$r_k(m) = FRAT_f(k, m) = \frac{1}{\sqrt{p}} \sum_{(x,y) \in L_{k,m}} f(x, y) \quad (1.18)$$

où $L_{k,m}$ représente un ensemble de points qui forment une ligne sur la grille, k définit la pente de la ligne et m réfère au coefficient de la transformée en ridgelets dans chaque direction. La transformée en Ridgelet discrète FRIT est obtenue en considérant la transformée en ondelettes discrètes (DWT) de chaque de projection FRAT de l'ensemble $(r_k(0), r_k(1), \dots, r_k(p-1))$, elle est définie formellement par :

$$FRIT(k, n) = \langle FRAT(k, m), \phi_n^k(m) \rangle \quad (1.19)$$

où $\phi_n^k(m)$ est la base d'ondelette conçue pour la FRAT (voir figure 1.3(b))

Malheureusement, par construction la transformée en Ridgelets ne s'avère efficace que pour la caractérisation des régularités rectilignes. A cet effet, afin de représenter les singularités curvilignes ou localement linéaires, il suffit de subdiviser l'image en blocs, puis d'appliquer la transformation Ridgelet aux sous-images obtenues. Cette analyse multi-échelles qui applique une transformée en Ridgelet par blocs est appelé la transformée en Curvelets [12,30,32-35].

1.3.2. Transformée en Curvelets

Les curvelets ont été proposées par E. Candes et D. Donoho [12]; elles forment une nouvelle famille de frames d'ondelettes géométriques à base non adaptative très simple, et adapté à la régularité géométrique des images, elles permettent de représenter de façon parcimonieuse les contours. Comparées aux ondelettes, les curvelets peuvent représenter un contour lisse avec moins de coefficients pour la même précision grâce à leur forte anisotropie (voir figure 1.4). L'idée de base derrière les curvelets est de représenter une courbe comme une superposition de fonctions de différentes longueurs et largeurs, obéissant à la loi anisotropique d'échelle $Largueur \approx Longueur^2$.

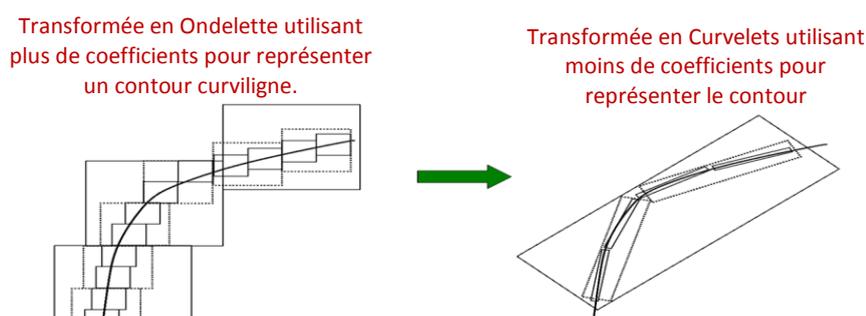


Figure 1.4. Représentation des contours lisses avec les ondelettes à support carrée et la Curvelet à support allongé [12]

1.3.2.1. La transformée en Curvelets première génération (Curvelet 1G)

La première génération de transformées est dérivée des ridgelets multi-échelles, elle nécessite pour sa construction les quatre étapes suivantes :

- La décomposition en sous bandes : cette décomposition est basée sur une décomposition en sous-bandes comme dans une transformée en ondelettes classiques. Chacune de ces sous-bandes est alors transformée à l'aide d'une transformée directionnelle qui nécessite l'application d'un ensemble de filtres passe-bandes. Etant donné une fonction $f \in L^2(\mathbb{R}^2)$, B_0 : un filtre passe-bas et $\{H_1, H_2, \dots\}$ un ensemble de filtres passe-bandes concentrés à proximité des fréquences $[2^j, 2^{j+2}]$, la décomposition de la fonction f en appliquant l'ensemble des filtres se résume à ce qui suit :

$$f \rightarrow \{B_0 f, H_1 f, H_2 f, \dots, H_j f\} \quad (1.20)$$

- Partitionnement dyadique : le support de l'image sera alors partitionné en carrés de tailles variables avec recouvrement afin d'éviter les effets de bords, puis ces carrés seront alors décomposés par une analyse en Ridgelets discrètes. La construction de la transformée en Curvelets nécessite de définir une fonction fenêtre w (*Window function*), qui satisfait la condition d'admissibilité $\int_{-1}^1 w^2(t) dt = 1$. Par exemple, si on considère la fenêtre de Meyer

[20] donnée par :

$$w(t) = \begin{cases} 1 & |t| \leq 1/3 \\ \cos \left[\frac{\pi}{2} v(3|t| - 1) \right] & 1/3 \leq |t| \leq 2/3 \\ 0 & \text{sin on} \end{cases} \quad (1.21)$$

où v représente la fonction de lissage, étant donné un carré dyadique Q ; chaque sous-bande est subdivisée en carrés dyadiques en multipliant la sous-bandes par la fonction fenêtre de lissage; comme suit :

$$H_j f \rightarrow h_Q = (w_Q H_j f)_{Q \in \mathbb{Q}_j} \quad (1.22)$$

où \mathbb{Q}_j représente l'ensemble de tous les carrés dyadiques à l'échelle j .

- Renormalisation : dans cette étape, on normalise le carrée obtenu en carrée unitaire conformément à l'équation :

$$g_Q = (T_Q)^{-1} (w_Q H_j f)_{Q \in \mathbb{Q}_j} \quad (1.23)$$

où T_Q n'est autre que l'opérateur de normalisation.

- Application de la transformée Ridgelete : chaque carré normalisé est analysé dans le domaine de la transformée en Ridgelets orthonormale.

La figure 1.5 ci-après donne un aperçu sur l'organisation de la transformée en Curvelets, la décomposition fréquentielle associée à la segmentation dyadique permet de conditionner les données pour la transformée en Ridgelets dans le but de décrire des singularités; dont la taille et la forme du motif varient.

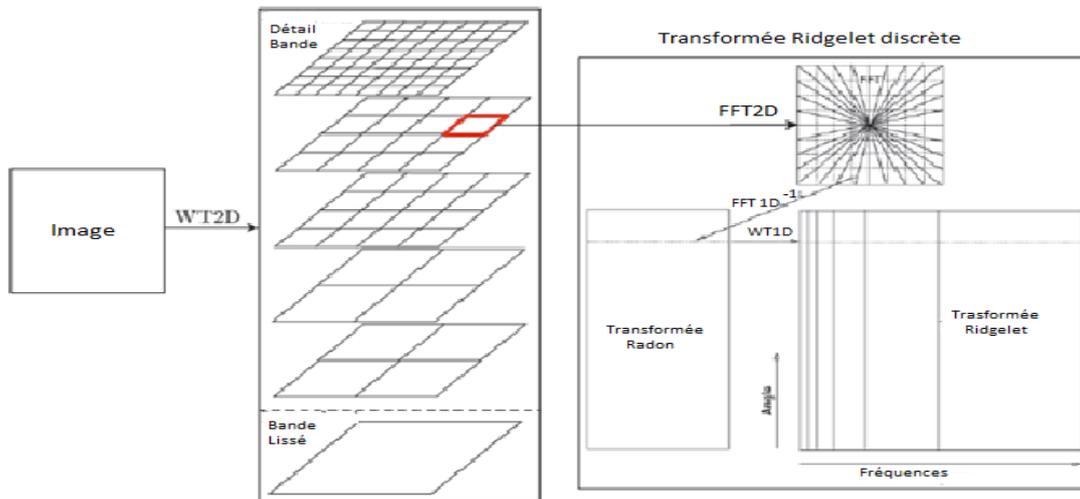


Figure 1.5. Organigramme de la transformée en Curvelets 1^{ère} génération

La transformée en Curvelets première génération améliore donc la prise en compte des singularités d'une image, par contre, elle est conçue de sorte que la transformée présente la particularité d'être très redondante. En effet, ceci est dû principalement à sa répartition spatiale qui utilise des fenêtres superposées pour éviter les effets de bord, ceci conduit à une augmentation de la redondance. Il existe des alternatives pour lesquelles cette redondance est limitée, telles que la transformée en Curvelets de deuxième génération et la transformée en Contourlets avec une redondance moins importante.

1.3.2.2. La transformée en Curvelets deuxième génération (Curvelet 2G)

Pendant l'année 2003 les mêmes auteurs ont proposé une deuxième génération beaucoup plus simple que la première génération [32-35], basée sur une technique de pavage du plan fréquentiel. D'une part, cette deuxième génération de la transformée en Curvelets est rapide et d'autre part, elle est moins redondante car dans cette nouvelle version les transformées en ridgelets ne sont pas utilisées, ce qui réduit considérablement la redondance et augmente la vitesse. La deuxième génération de transformée en Curvelets est définie à la fois dans les

domaines continu et discret. Elle est définie à l'échelle 2^{-j} , à l'orientation θ_l et à la position $x_k^{j,l} = R_{\theta_l}^{-1}(2^{-j}x_1, 2^{-j/2}x_2)$ par la translation et la rotation de la curvelette mère φ_j (ou la forme d'onde : waveform), elle s'écrit formellement :

$$\varphi_{j,k,l}(x) = \varphi_j \left(R_{\theta_l} \left(x - x_k^{j,l} \right) \right) \quad (1.24)$$

où R_θ : la matrice de rotation, θ_l : séquence equi-espace de l'angle de rotation ($\theta_l = 2\pi 2^{-\lfloor j/2 \rfloor} l$ avec $0 < \theta_l < 2\pi$). Quand à la forme d'onde, elle est définie en utilisant les transformées de Fourier par :

$$\hat{\varphi}_j(r, \theta) = 2^{-3j/4} \hat{w}(2^{-j}r) \hat{v} \left(\frac{2 \lfloor j/2 \rfloor \theta}{2\pi} \right) \quad (1.25)$$

où \hat{w} et \hat{v} : représentent respectivement les fenêtres radiales et angulaires qui permettent de réaliser un pavage dans le domaine fréquentiel. Les coefficients des Curvelets 2G sont obtenus par le produit scalaire d'une fonction $f \in L^2(\mathbb{R}^2)$ et la base Curvelet 2G, ils sont donnés par l'équation suivante :

$$c_{j,k,l} = \langle f, \varphi_{j,k,l} \rangle = \int_{\mathbb{R}^2} f(x) \cdot \overline{\varphi_{j,k,l}} dx \quad (1.26)$$

Ces mêmes coefficients peuvent être aussi exprimés dans le plan fréquentiel, ceci est possible en utilisant le théorème de Plancherel, ce qui donne l'expression :

$$c_{j,k,l} = \langle f, \varphi_{j,k,l} \rangle = \int_{\mathbb{R}^2} \hat{f}(\omega) \cdot \hat{\varphi}_j(R_{\theta_l} \omega) e^{-ix_k^{j,l} \omega} d\omega \quad (1.27)$$

Par ailleurs, l'implémentation numérique est rendue possible par les deux approches proposées par E. Candès, L. Demanet et al [34], à savoir :

- L'approche basée sur l'algorithme *Unequi-spaced FFT*
- L'approche basée sur l'algorithme *Warping*

1.3.3. Transformée en Contourlets

Conçue directement dans le domaine discret, La transformée en Contourlets a été introduite par Minh Do et Martin Vetterli [13,25,26,36,37], c'est un nouveau schéma de décomposition d'images qui fournit une représentation éparsée des données aussi bien aux résolutions spatiales que fréquentielles. La décomposition d'images en Contourlets est construite en combinant successivement deux étapes de décomposition distinctes : une décomposition multi-échelle suivie d'une décomposition directionnelle. La première étape utilise une pyramide

Laplacienne pour décomposer de manière itérative une image 2-D en une suite de sous-images passe-bandes LP et une image de niveau passe-bas (approximation en basse fréquence de l'image). La deuxième étape applique de manière appropriée des filtres bidimensionnels de type quinconce et un échantillonnage critique pour décomposer chaque niveau LP passe-bande en un nombre de bandes directionnelles [38,39], capturant ainsi des informations directionnelles ou orientées. Finalement, l'image se trouve représentée par un ensemble de sous bandes multiéchelles et orientées (voir figure 1.6).

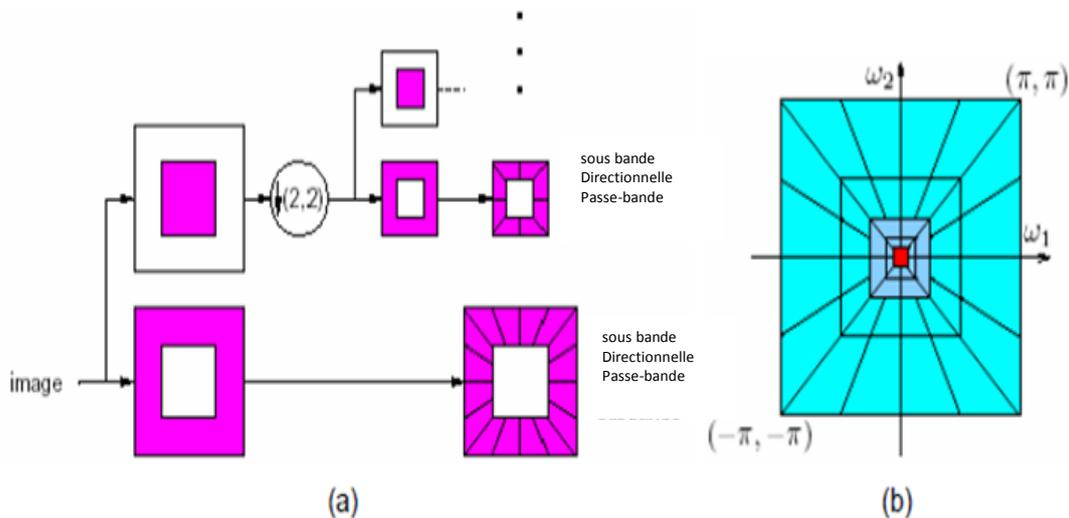


Figure 1.6. (a): Illustration de l'implantation des contourlets à l'aide d'un banc de filtres pyramidal directionnel (PDFB), (b): Exemple de partitionnement fréquentiel par la transformée en Contourlets [13].

La Figure 1.7 ci-après illustre un exemple de décomposition en transformée en Contourlets de l'image Polygone (512x512). En effet, L'image est décomposée en une image d'approximation passe-bas et en plusieurs sous-bandes passe-bande directionnelles. Seuls les contourlets qui coïncident à la fois avec la position et la direction des contours de l'image produisent des coefficients significatifs (en blanc). Par ailleurs, la transformée en contourlets garantit une reconstruction parfaite de l'image, son degré de redondance est relativement faible puisque le nombre total de ses coefficients est obtenu à l'issue d'une décomposition approchant les 4/3 du nombre d'échantillons de l'image originale (pour J échelles associés à la transformée, la redondance est $1 + \sum_{i=0}^J \frac{1}{4^i} < \frac{4}{3}$) [13,26,36], en raison des sous-échantillonnage et sur-échantillonnage présents à la fois dans LP et DFB.

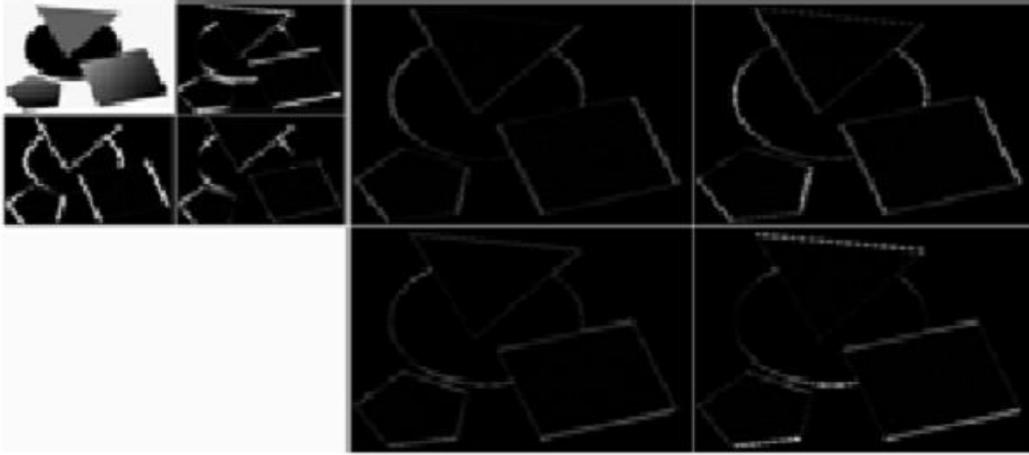


Figure 1.7. La transformée en Contourlets de l'image Polygone (512x512) où l'image est décomposée en un seul niveau pyramidal et quatre sous-bandes directionnelles. Les petits coefficients sont représentés en noir alors que les grands coefficients sont représentés en blanc.

1.3.4. Transformée Wave atoms

Les Auteurs L. Demanet et L. Ying proposent une nouvelle famille de transformation multi échelles et multidirectionnelles qui offre une représentation parcimonieuse de l'image [33,40]. Elles sont particulièrement adaptées à la bonne représentation des images contenant des modèles oscillatoires. Elle peut être considérée comme une variante de paquets d'ondelettes 2D et possède deux propriétés très importantes : La première est sa capacité d'adaptation aux directions locales arbitraires d'un motif; alors que la seconde est la capacité de représenter les modèles faiblement anisotropes alignées avec les axes.

La transformée wave atoms est mise en œuvre par l'intermédiaire des produits tensoriels de paquets d'ondelette 1D choisies adéquatement. Soit la base wave atoms $\phi_u(x)$ paramétrée avec les indices $u(j, m, n) = (j, m_1, m_2, n_1, n_2)$ choisis appartenant à \mathbb{N} (entiers naturels), associés au point (x_u, ω_u) dans l'espace de phase tels que :

$$x_u = 2^{-j} n, \omega_u = \pi 2^j n \quad \text{et} \quad C_1 2^j \leq \max_{i=1,2} |m_i| \leq C_2 2^j \quad (1.28)$$

où C_1 et C_2 sont deux constantes positives. La base orthonormale 2D de la transformée wave atoms $\phi_u^+(x)$ est engendrée par le produit de paquet d'ondelettes 1D ψ_m^j conformément à l'équation :

$$\phi_u^+(x) = \psi_{m_1}^j(x_1 - 2^{-j} n_1) \cdot \psi_{m_2}^j(x_2 - 2^{-j} n_2) \quad (1.29)$$

sa transformée de Fourier est aussi séparable et exprimée par :

$$\widehat{\phi}_u^+(\omega_1, \omega_2) = \widehat{\psi}_{m_1}(\omega_1) e^{-i2^j n_1 \omega_1} \cdot \widehat{\psi}_{m_2}(\omega_2) e^{-i2^j n_2 \omega_2} \quad (1.30)$$

Par ailleurs, La base orthonormale duale peut-être définie en utilisant la transformée de Hilbert par :

$$\phi_u^-(x) \equiv H\psi_{m_1}^j(x_1 - 2^{-j} n_1) \cdot H\psi_{m_2}^j(x_2 - 2^{-j} n_2) \quad (1.30)$$

Finalement, on peut définir la trame de la base wave atoms (*the wave atoms tight frame*) formellement par :

$$\phi_u^{(1)}(x) = \frac{\phi_u^+(x) + \phi_u^-(x)}{2}; \quad \phi_u^{(2)}(x) = \frac{\phi_u^+(x) - \phi_u^-(x)}{2} \quad (1.31)$$

La figure 1.8 représente le pavage de la transformée Wave atoms dans le plan fréquence-spatiale. Il est à noter que lorsque l'échelle j augmente de 1, la taille du carrée double. À une échelle j donné, les carrés sont indexés par m_1 et m_2 à partir de zéro à proximité de l'axe.

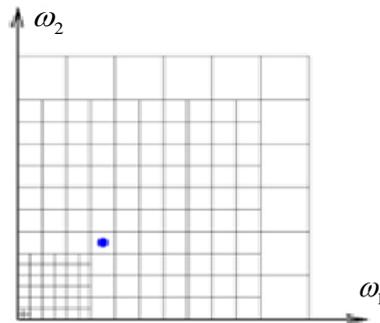


Figure 1.8. Le pavage de la transformée Wave atoms dans le plan fréquentiel où seul le premier quadrant est représenté [40]

1.4. Transformées géométriques adaptatives

De nombreuses représentations géométriques adaptatives ont été étudiées ces dernières années pour les applications de traitement d'images. Dans le cas de ces transformées, on cherche à décomposer une image sur une base dont les fonctions sont choisies pour mieux s'adapter à la régularité géométrique de l'image. Par conséquent, il faut déterminer préalablement la régularité géométrique de l'image par un modèle géométrique qui détermine explicitement la direction d'analyse locale avant de faire la décomposition. Dans ce qui suit, on présentera un aperçu détaillé sur les transformations géométriques adaptatives les plus utilisées.

1.4.1. Transformée en Bandelettes

La transformée en bandelettes exploite avantageusement les structures géométriques contenues dans les images en construisant des vecteurs orthogonaux qui sont allongés dans la

direction où la fonction a une régularité maximum. Ces bandelettes sont obtenues à partir de déformation d'ondelettes anisotropes. La base bandelettes décompose l'image le long de vecteurs multi-échelles qui sont allongés dans la direction d'un flux géométrique. Ce flux géométrique indique les directions où les niveaux de gris des images ont des variations régulières

1.4.1.1. Définition de la géométrie

La géométrie de l'image est représentée par ses contours et ces courbures. Sachant que la tangente à un contour est une direction de la régularité maximale, il en résulte que la régularité peut être mesurée tandis que la position ne peut être déterminée que dans le cas de contours nets [6,7]. On peut définir chaque courbe (contour) C paramétrée par son abscisse curviligne s par :

$$C = \{c(s) = (c_1(s), c_2(s)), s \in \mathbb{R}, c_1 : \mathbb{R} \mapsto [0,1] \text{ et } c_2 : \mathbb{R} \mapsto [0,1]\} \quad (1.32)$$

Le flot τ le long du contour (voir figure 1.9), désigne le vecteur unitaire de cette direction de régularité maximale le long du contour [6]. Il est défini par :

$$\tau(c(s)) = c'(s) = \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} \quad (1.33)$$

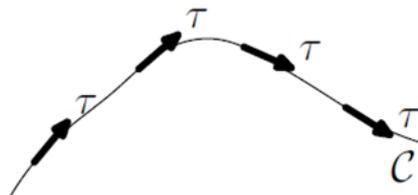


Figure 1.9. Exemple d'un contour C et son flot τ définie par sa tangente [6]

La direction de la régulation maximale d'une fonction dérivable et régulière, est égale à la direction orthogonale au gradient; qui est définie par :

$$\tau(x) = \frac{\nabla f(x)^\perp}{|\nabla f(x)|} \quad (1.34)$$

L'idée centrale dans la construction des bandelettes est de définir la géométrie comme un champ vectoriel ou un flux géométrique pour effectuer une analyse orientée le long des contours. Ce flux indique la direction du déplacement des valeurs de niveau de gris, non pas dans le temps, mais dans l'espace. Il existe deux versions de construction de la transformée en bandelettes à savoir : les bandelettes de la première et la deuxième génération.

1.4.1.2. Les bandelettes de la première génération (Bandelet 1G)

La première génération de transformées en bandelettes [6] permet de représenter la géométrie des images en utilisant des flots géométriques au lieu des contours. Ces bandelettes sont construites à partir d'ondelettes bidimensionnelles en effectuant un partitionnement du support de l'image en des régions, à l'intérieur des régions ces flots géométriques restent parallèles ou non définis. Pour chaque région, les bases orthonormées de bandelettes sont définies en déformant la base d'ondelettes localement sur des carrés dyadiques pour aligner la singularité avec la direction horizontale ou verticale. Il existe trois types de régions :

➤ Régions où l'image est uniformément régulière, et où il n'y a donc pas de flot, la représentation se fait alors avec une base d'ondelettes bidimensionnelles classiques.

➤ Régions où le flot géométrique est parallèle verticalement $\vec{\tau}(x_1)$

$$\vec{\tau}_i(x_1, x_2) = \vec{\tau}_i(x_1) = (1, c_i(x_1))$$

➤ Régions où le flot géométrique est parallèle horizontalement

$$\vec{\tau}_i(x_1, x_2) = \vec{\tau}_i(x_2) = (c_i(x_2), 1)$$

La première étape de la représentation en bandelettes consiste à isoler ces régions de flots géométriques parallèles, cette segmentation est réalisée à l'aide d'une décomposition en arbre quaternaire en carrés dyadiques.

Par ailleurs, la construction d'une base de bandelettes s'effectue sur un domaine Ω_i où la fonction $f \in L^2(\mathbb{R}^2)$. On suppose que cette région contienne une seule singularité, soit horizontale ou verticale, que l'on peut approcher par une courbe paramétrée (par exemple horizontalement voir figure 1.10), l'opérateur de déformation est définie comme suit :

$$\forall x \in \Omega_i, w(x) = (x_1, x_2 - c(x_1)) \tag{1.35}$$

Cet opérateur permet de définir le domaine déformé $W\Omega_i = w(\Omega_i)$ ainsi que la fonction warpée

$$Wf(x) = f(w^{-1}(x))$$

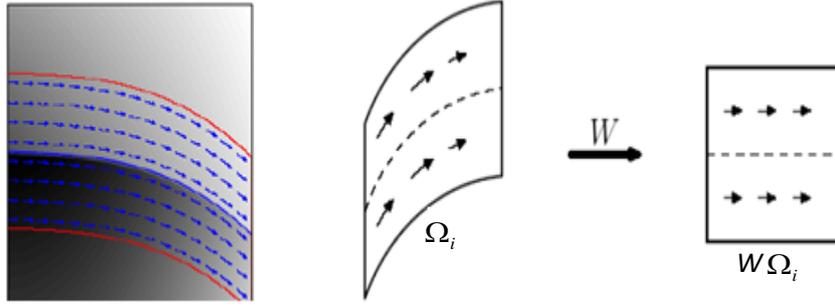


Figure 1.10. Exemple de modèle d'horizon et déformation du domaine selon un flot géométrique [6].

Par conséquent, pour construire la première génération de bandelettes, il faut tout d'abord diviser l'image f en régions Ω_i où le flot géométrique donne le sens de la variation régulière de l'image, ensuite effectuer les tâches énumérées ci-dessous:

- a- un ré-échantillonnage qui calcule les valeurs de l'image le long du flot géométrique.
- b- une transformation en ondelettes effectuée par un filtrage en sous-bandes le long du flot, si la région est régulière (sans flot géométrique) alors, la décomposition est effectuée en utilisant une base d'ondelettes bidimensionnelles classique à l'aide d'un banc de filtres de Daubechies biorthogonales 9/7. Une transformation en ondelettes discrètes séparables est définie par son ondelette mère ψ et sa fonction d'échelle associée φ , exprimée par:

$$\begin{cases} \varphi_{j,m_1}(x_1)\psi_{j,m_2}(x_2) \\ \varphi_{j,m_1}(x_1)\psi_{j,m_2}(x_2) \\ \psi_{j,m_1}(x_1)\psi_{j,m_2}(x_2) \end{cases} ; j \in \mathbb{Z}, (m_1, m_2) \in \mathbb{Z}^2 \quad (1.36)$$

Autrement, si la région a un flot géométrique, la décomposition est effectuée sur une base d'ondelettes déformée par un opérateur de déformation $W(\vec{\tau})$. La base orthonormée déformée pour un flot horizontal parallèle est formellement donnée par :

$$\begin{cases} \psi_{j,m_1}(x_1 - c_i(x_2))\varphi_{j,m_2}(x_2) \\ \varphi_{j,m_1}(x_1 - c_i(x_2))\psi_{j,m_2}(x_2) \\ \psi_{j,m_1}(x_1 - c_i(x_2))\psi_{j,m_2}(x_2) \end{cases} ; (j, m_1, m_2) \in I_{W\Omega_i} \quad (1.37)$$

Par conséquent, pour un flux horizontal l'image déformée est alors :

$$Wf(x_1, x_2) = f(x_1, x_2 + c(x_1)) \quad (1.38)$$

- c- une "bandeletisation" qui transforme les coefficients des ondelettes déformées en coefficients bandelettes en exploitant la régularité de la fonction le long du flot.

Il est important de préciser que la segmentation en carrées dyadiques est obtenue par une subdivision successive du support de l'image. La dimension maximale et minimale du carrée affecte les directions des flots géométriques. La figure 1.11 ci-après illustre les flots géométriques dans différentes partitions de l'image.

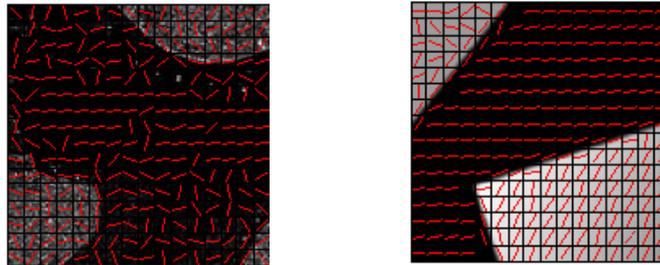


Figure 1.11. Flots géométriques en utilisant différente segmentations dyadiques

Les bandelettes de la première génération utilisent avantageusement les structures géométriques des images. Cependant, elles ne sont pas directement définies dans le cas discret et elles n'offrent pas une représentation multi-résolution de la géométrie. A cet effet, Peyré et al. [8,41] ont défini la seconde génération de bandelettes.

1.4.1.3. Les bandelettes de la seconde génération (Bandelet 2G)

Cette nouvelle version est basée sur la transformée en ondelettes, elle permet une représentation multirésolution de la géométrie. L'objectif de ces bandelettes est d'exploiter les régularités géométriques le long des contours dans les sous-bandes de la transformée en ondelettes. Sa construction consiste à appliquer les étapes ci-après :

- une transformation en ondelettes orthogonales ou biorthogonales discrètes à l'image f ,
- une segmentation hiérarchique en carrées dyadiques est réalisée dans chaque sous-bande dans le sens de la meilleure représentation de la géométrie, la figure suivante présente la construction d'une segmentation en *quadtree* [8,41-43],

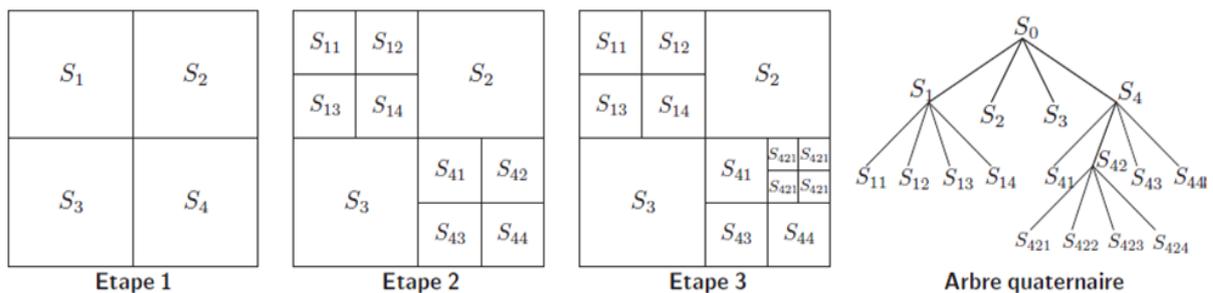


Figure 1.12. Exemple de segmentation en carrées dyadiques d'une image avec le quadtree correspondant [8].

- dans chaque carré dyadique, déterminer la meilleure géométrie qui définit la direction de la régularité géométrique. Une base de bandelettes $B(\Gamma)$ dépend de la géométrie des flots géométriques locaux définis par une segmentation dyadique des coefficients d'ondelettes, et par le choix d'un flot polynômial à l'intérieur de chaque carré. Le but étant d'optimiser l'approximation f_M de f , la meilleure géométrie est celle qui optimise l'erreur d'approximation $\|f - f_M\|^2$ pour un nombre M de paramètres décrivant f_M . Soit $T > 0$ et soit f une image de $N \times N$ pixels discrétisée, ayant une régularité géométrique C^α . Soit $B^*(\Gamma)$ la base de bandelette discrète définie par : $B^*(\Gamma) = \arg \min_{B^*(\Gamma) \in D_T} (\mathcal{L}(f, B(\Gamma), T))$; avec \mathcal{L} : le lagrangien, D_T : le dictionnaire des bases bandelettes. La meilleure géométrie est celle qui minimise le lagrangien pour un seuil T .
- une projection orthogonale 1D est effectuée à la géométrie spécifiée pour définir un signal discret 1D f_d ,
- une transformée en ondelettes discrète 1D est appliquée au signal 1D f_d donnant les coefficients de bandelettes b_k .

L'algorithme de la représentation en bandelettes de seconde génération d'une image est illustré par la figure 1.13.

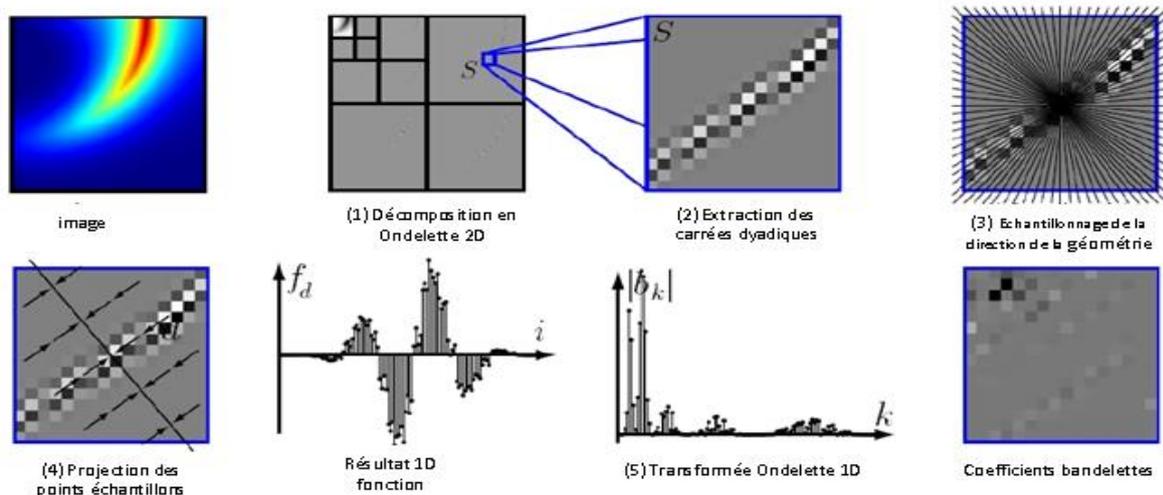


Figure 1.13. Algorithme de la transformée en bandelettes seconde génération [8]

1.4.2. Les transformées Wedgelets

Les wedgelets introduites par Donoho [44] est une représentation des images constantes par morceaux ("Horizon class function"), par une segmentation en quadrees du support de

l'image dans lequel chaque bloc dyadique $S(j, k_1, k_2)$ est défini par : $\{(x_1, x_2) : [k_1 / 2^j, k_1 + 1 / 2^j] \times [k_2 / 2^j, k_2 + 1 / 2^j]\}$ avec $0 \leq k_1, k_2 < 2^j$. Chaque bloc est séparé en deux régions d'intensités constantes réparties des deux côtés d'une frontière $L = [v_1 v_2]$ appelée *Edgel* (*Edge element*), qui traverse le carré S . L'idée principale de cette transformée est de chercher une rupture edgel coupant le carré en deux morceaux de fonctions constantes. S'il y'a plus de morceaux, la subdivision est réitérée sinon, si le bloc ne contient aucun morceau, l'absence d'arrêtes est indiquée et le processus de segmentation est arrêté.

Ce comportement de la transformée s'obtient de manière plus globale en cherchant à minimiser la différence entre l'ensemble des données traitées et sa projection sur l'ensemble des edgels. Nous obtenons donc l'ensemble quadtree plus les edgels formant le dictionnaire sur lequel s'appuie les wedgelets pour construire la transformée. Après la constitution de ce dictionnaire, les coefficients de wedgelets permettent de définir la valeur de chaque morceau constant pour obtenir la transformée wedgelet. Chaque wedgelet peut simplement représenter un contour rectiligne dans un bloc de l'image (voir figure 1.14).

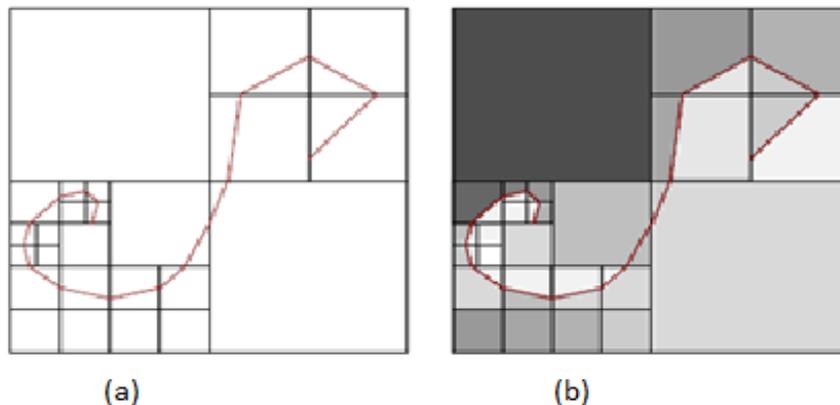


Figure 1.14. (a): Dictionnaire d'edgels déterminé pour cette image , (b): Transformée wedgelets : zones uniformes mises en évidence [44].

1.4.3. Les transformées Directionlets

Velisavljević propose dans sa thèse [45, 46] une transformée discrète à échantillonnage critique appelée transformée en directionlets. Cette technique fournit une représentation anisotrope de l'image, ce qui est très adapté aux images à forte structure géométrique. La transformée directionlets se base sur le concept des transformées en ondelettes classiques, mais en utilisant un filtre passe-haut ne possédant pas uniquement des moments nuls verticaux et horizontaux, mais aussi dans d'autres directions appelées moment nuls directionnels. Les

directions de ces moments sont définies par des lignes numériques qui représentent toutes les directions définies sur un treillis entier. Ces directions servent en réalité d'angle pour effectuer une rotation appropriée avant d'appliquer un filtrage horizontal. Techniquement, il s'agit d'une généralisation des ondelettes séparables qui peuvent être vues comme des directionlets sur un treillis carré.

Considérons une image avec deux directions dominantes et deux pentes $d_1 = b_1 / a_1$ et $d_2 = b_2 / a_2$, où a_1, a_2, b_1, b_2 sont des entiers. Soit la matrice conjointe des pentes des deux lignes de direction :

$$M(r_1, r_2) = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \quad (1.39)$$

Afin d'obtenir une représentation parcimonieuse ("*sparse representations*") de cette image, une transformation en ondelette 1D est appliquée le long des lignes numériques $L(d, n) : n \in \mathbb{Z}$ (l'équation de la ligne numérique est donnée par: $y = \lceil d \cdot x \rceil + n, \forall x \in \mathbb{Z} \text{ et } |d| \leq 1$). Ceci produit des coefficients non nuls qui correspondent aux discontinuités le long des lignes avec les deux pentes d_1 et d_2 . En appliquant un filtre passe-haut (HP) à moments nuls le long de la ligne numérique de pente d_1 , les coefficients obtenus sont placés dans la sous-bande HP. Alors que, les coefficients le long de la seconde direction sont retenus dans les deux sous-bandes. La figure 1.15 ci-après illustre un exemple d'application de la Transformée en ondelettes 1D le long de la ligne

numérique de pente $-1/2$; avec $M(r_1, r_2) = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$.

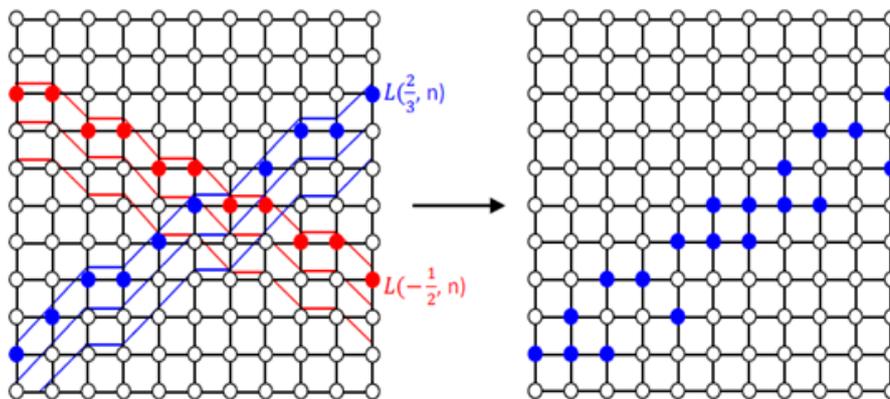


Figure 1.15. Illustration de l'interaction directionnelle lorsque le filtrage est effectué le long des lignes numériques avec filtrage passe-haut le long de la pente $-1/2$.

1.4.4. Les ondelettes orientées

La transformée en ondelettes orientées est basée sur la décomposition en ondelettes classiques et le schéma adaptatif et directionnel du lifting en exploitant les orientations locales identifiées sur l'image. W. Ding et al [47] ont déterminé les orientations par zone dyadique, une fois la direction déterminée, le calcul de la prédiction dans le lifting s'adapte à la direction en choisissant les échantillons voisins (voir figure 1.16 (a)). Soit le signal 2D $f(m,n)_{m,n \in \mathbb{Z}}$; le signal est d'abord décomposé en utilisant la transformée en ondelette 1D dans la direction verticale, ensuite dans la direction horizontale. Sachant que chaque phase de lifting est composée en trois étapes (subdivision, prédiction et mise à jour). W. Ding et al proposent d'abord de diviser tous les échantillons en échantillons pairs et impairs conformément à ce qui suit :

$$\begin{cases} f_e(m,n) = f(m,2n) \\ f_o(m,n) = f(m,2n+1) \end{cases} \quad (1.40)$$

ensuite, dans la phase de prédiction les échantillons impairs sont prédits par une combinaison linéaire des échantillons pairs de forte corrélation, comme suit :

$$h(m,n) = f_o(m,n) - P_e(m,n) \quad (1.41)$$

avec $P_e(m,n) = \sum_i \alpha_i f_e(m + \text{sign}(i-1) \tan g\theta_v, n+1)$, où $\text{sig}(x) = 1 \text{ si } x \geq 0; -1 \text{ ailleurs}$; α_i : des poids donnés par le filtre et θ_v : l'angle correspondant à une forte corrélation.

Finalement, vient la phase de mise à jour où chaque échantillon pair est remplacé par $l(m,n) = f_e(m,n) + u_h(m,n)$ avec $u_h(m,n) = \sum_j \beta_j h(m + \text{sign}(j) \tan g\theta_v, n+j)$

où β_j représentent les poids donnés par le filtre.

Une autre variante d'ondelette orientée de Chang et al. [48] introduit deux améliorations. En premier lieu, afin d'adapter au mieux la taille des champs de directions à la taille des motifs dans le signal, la segmentation dyadique est remplacée par un quadtree. Ensuite, afin d'éviter de privilégier les directions horizontale et verticale, la décimation associée à la transformée correspond à deux treillis quinconces entrelacés (voir figure 1.16(b)). Finalement, une version de transformée orientée toujours basée sur un treillis est introduite par V. Chapellier et al. [49-50]. Elle diffère des deux transformées précédentes par le nombre de directions prises en charge, à savoir deux directions : horizontale ou verticale. Cette limitation s'accompagne d'une plus grande précision au niveau spatial, car ils ne définissent pas seulement une direction par zone dyadique mais une direction par élément (voir 1.16 (c))

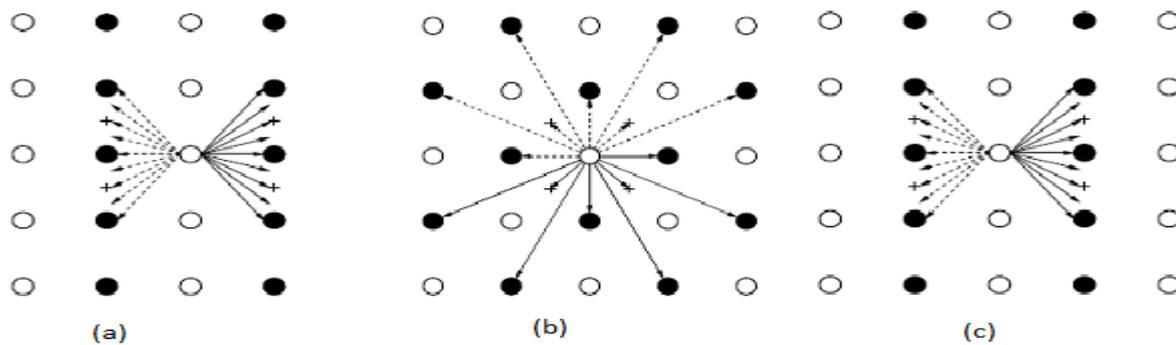


Figure 1.16. Directions de Prédiction dans la transformée en ondelettes orientées , (a) : Directions exploitées sur maillage séparé en pair-impair ; (b) : Directions exploitées sur treillis quinconces, la recherche de directions se fait pour les deux treillis et (c): Directions exploitées sur treillis quinconces, l'information de direction pour la prédiction est directement exploitée pour la mise à jour

1.5. Etude comparative des différentes transformées directionnelles

Il a été démontré que les images naturelles ne sont pas de simples fonctions de $L^2([0,1]^2)$, mais qu'elles possèdent une composante géométrique. Afin d'obtenir une représentation creuse et efficace de l'image, il est important de prendre en considération l'aspect géométrique. Par ailleurs, toute régulation géométrique implique une forme de redondance qu'il faut exploiter pour des processus de traitement d'image, tel que le débruitage et la caractérisation. L'exploitation de la redondance est une idée centrale pour la construction de ces représentations par ondelettes directionnelles de seconde génération qui permettent d'obtenir des approximations optimales avec peu de coefficients, et qui sont plus adaptées aux structures géométriques lisses et continues [6,8,9,12,13,27-52].

Une étude comparative de toutes les transformées introduites dans ce chapitre est présentée par le tableau 1.1, afin de donner une idée globale sur les caractéristiques de ces dernières.

1.6. Conclusion

La richesse des images naturelles rend la quête d'une représentation optimale dans le domaine de traitement de l'image, et de vision par ordinateur un enjeu majeur pour les chercheurs, qui doit satisfaire un compromis entre l'efficacité et la complexité.

Différents outils et modèles ont été exposés pour la représentation des structures géométriques contenues dans une image. Chaque transformée propose une approche originale pour capturer et décrire la géométrie et s'avère efficace dans le cadre d'un modèle restrictif de la

régularité géométrique modélisée (courbe, contour, segment, flot géométrique, ..etc.). Le choix d'une transformée correspond à des compromis que l'on doit fixer en général en fonction de l'application visée. Les transformées non adaptatives sont généralement redondantes; la transformée en Ridgelet et la transformée en Curvelet sont des exemples qui se prêtent à des applications de traitement d'images telles que le débruitage et l'extraction de caractéristiques discriminantes dans une image. En revanche, elles sont moins sollicitées pour la compression. A l'inverse, la plupart des transformées en ondelettes adaptatives sont idéales pour les applications de compression telles que les transformées en Bandelettes. Ces représentations directionnelles permettent une meilleure concentration de l'énergie avec peu de coefficients conservés, et la synthèse présente de meilleurs résultats visuels.

Dans le cadre de cette thèse, les applications développées concernent la compression des images hautes-résolutions (image SAR et images médicales) en utilisant les transformées adaptatives, telles que les transformées Bandelettes et la reconnaissance d'images palmaires multispectrales, et des images d'empreintes palmaires en utilisant les transformées non adaptatives.

| Type | Multiéchelle | Redondance | Nature de singularités | Nombre d'orientation | Approximation | Nombre d'opération |
|-------------------------------------|--------------|----------------|------------------------|-----------------------------------|--|---|
| Ondelettes séparable | Oui | Oui | Ponctuelle | 3 | $\ f - f_M\ _{C^2}^2 \leq AM^{-2\alpha}$ | $N \cdot \log(N)$ |
| Transformées non adaptatives | | | | | | |
| Ridgelet | Non | | Droites | ∞ ($p + 1$) pour FRIT | $\ f - f_M\ _{C^2} \leq C \cdot m^{-r}; r > 0$ | - |
| Curvelet | Oui | Redondante | Segments | 2^k | $\ f - f_M\ _{C^2}^2 \leq A \cdot \log(M)^3 M^{-2}$ | $N^2 \log N$ |
| Contourlet | oui | Peu redondante | directions privilégiés | 2^k | $\ f - f_M\ _{C^2}^2 \leq A \cdot \log(M)^3 M^{-2}$ | $L_1 \cdot L_2 \cdot \log(N)$ ($L_1 \times L_2$ avec taille du filtre) |
| Wave atoms | oui | Redondante | Direction | k | $\ f - f_M\ _{C^2}^2 \leq A \cdot (M)^{1-2/p} A _{L^p}^2$ | $N^3 \cdot \log N$ |
| Transformées Adaptatives | | | | | | |
| Wedgelet | oui | Peu redondante | Edgels | ∞ | | $N^2 \cdot \log N$ |
| Bandelet 1G | oui | Non redondante | Contours | - | $\ f - f_M\ _{L^2}^2 \leq AM^{-\alpha}$ | $N^2 \cdot (\log N)^2$ |
| Bandelet 2G | Oui | Non redondante | Contours dans quadtree | - | $\ f - f_R\ _{L^2}^2 \leq A \log(R)^\alpha R^\alpha$ | $N^2 \cdot T^{2(p-1)}$ avec T le seuil et p : nombre de moments nuls |

Tableau 1.1. Etude comparative des transformées orientées

Chapitre 2

Transformées en Bandelettes pour la compression des images haute résolution

2.1. Introduction et état de l'art sur la compression

La compression a émergé comme un domaine majeur de la recherche en raison de la croissance phénoménale des applications qui génèrent un traitement et une transmission des images volumineuses, comme pour l'imagerie satellitaires, l'imagerie médicale, la vidéoconférence, le multimédia, etc.. L'objectif principal de la compression d'image est de réduire la redondance des données afin d'éliminer les informations inutiles et redondantes dans l'image, et de représenter les informations pertinentes restantes avec le plus petit nombre de bits. En effet, ceci permet de stocker, traiter ou de transmettre les données d'une façon efficace et avec une tolérance faible pour la dégradation de la qualité de l'information source. Selon la qualité de l'information reconstruite par les systèmes de compression "Codec", il existe deux grandes familles d'algorithmes de compression : Les algorithmes sans perte d'information (Lossless compression) qui peuvent reconstituer l'information exactement équivalente à l'originale, et les algorithmes avec perte d'information (Lossy compression) qui tolèrent une élimination d'information pour augmenter le taux de compression, sous la contrainte de trouver un meilleur compromis entre le débit binaire et la distorsion acceptée par l'application.

Formellement, une image est une collection d'intensités situées à des positions spatiales distinctes, chaque intensité est notée $f_k = f(x_k, y_k)$ pouvant prendre une valeur dans l'ensemble des intensités $\{q_1, q_2, \dots, q_L\}$, souvent considéré comme l'ensemble des seuils de quantification admissibles [53]. L'information moyenne contenue dans une image est appelée entropie, elle est exprimée en (bit/pixel) et peut être calculée par l'équation: $H = -\sum_{i=1}^L p_i \log_2 p_i$ avec $p_i = \text{Probabilité}\{f(x, y) = q_i\}$. Etant donné un nombre de bit par pixel R , l'entropie d'ordre zéro est définie par : $H_0 = -\sum_{i=1}^{2^R-1} p_i \log_2 p_i$ [53,54]. L'entropie H_0 d'une image originale fournit le débit minimal qu'il est possible d'atteindre par compression pour un taux de compression sans perte maximale. Généralement, il existe trois types de redondances dans l'image, le premier type est directement liée à l'entropie, appelé le codage de la redondance (*redundancy coding*) [53], et défini par :

$$R_c = \frac{H_{\max} - H}{H_{\max}} \quad (2.1)$$

avec $H_{\max} = \log_2 L$. Une valeur nulle du codage de la redondance ($R_c = 0$) signifie que la distribution des valeurs des intensités dans une image est uniforme, par contre $R_c = 1$ si toutes les valeurs des intensités sont les mêmes [53,54]. Les méthodes dont l'objectif est d'éliminer la redondance de codage sont communément appelées méthodes de codage entropique; il en existe un bon nombre dont le codage Huffman [55] et le codage arithmétique [56] qui comptent parmi les plus populaires. Les codeurs entropiques font partie des algorithmes de compression sans perte, car ils éliminent la redondance sans aucune perte d'information.

Le deuxième type de redondance est directement lié à la corrélation spatiale entre les pixels. Afin d'éliminer ce type de redondance, il est préférable d'utiliser des codeurs à base de transformation (voir figure 2.1) qui modifient la représentation spatiale de l'image afin de réduire la corrélation entre les pixels voisins. Finalement, le troisième type de redondance est psycho-visuel ou perceptive, il est directement lié aux propriétés du système visuel humain HVS (Human Visual System) [53]. La figure 2.1 illustre le schéma d'un système de compression des images classique avec perte d'information, celui-ci comporte trois étapes de traitements. La première phase de traitement consiste en la décorrélation dont l'objectif est de réduire les redondances statistiques contenues dans l'image, et de ce fait réduire les entropies d'ordres faibles afin de séparer les composantes importantes de celles contenant peu d'information. Cette étape réversible sans pertes ou quasiment sans pertes fait intervenir une transformée, telle que la DCT (Discrete Cosine Transform) qui est utilisée dans le très populaire standard de compression d'images JPEG, ou encore la transformée en ondelettes qui est à la base du standard JPEG2000 [57]. La deuxième étape concerne la quantification, c'est la seule étape non réversible du schéma de compression, elle permet de réduire le nombre de bits nécessaires à la représentation de l'information. Finalement, la dernière étape est un codage entropique (sans perte) qui permet la compression des coefficients quantifiés en générant un flux binaire.

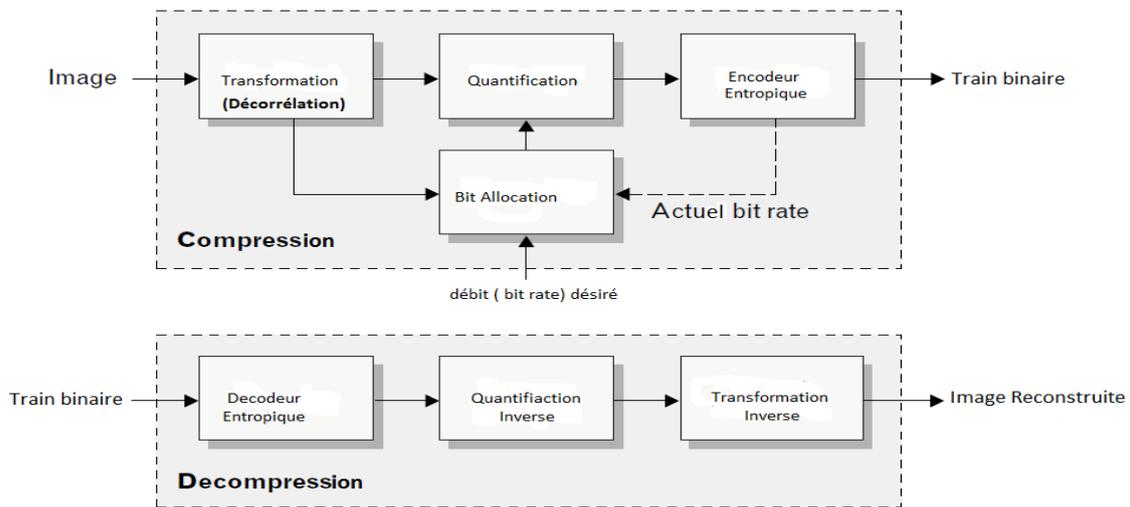


Figure 2.1. Schéma globale de la compression des images basée sur les transformations [53]

2.1.1 La décorrélation

Les pixels composant l'image numérique sont généralement corrélés, c'est à dire que les pixels adjacents sont statistiquement dépendants. Ainsi, ceci génère une information redondante qu'il faut exploiter pour la compression afin de diminuer la taille de l'image sans pour autant l'altérer. Une étape de décorrélation visant à supprimer cette redondance spatiale ou intra-image, s'avère alors très bénéfique pour l'amélioration des taux de compression. Il existe deux approches différentes qui permettent d'effectuer cette décorrélation, à savoir la technique basée sur la prédiction et celle basée sur une transformation orthogonale [8,53-54,58] sur laquelle nous nous focalisons essentiellement dans la suite de ce chapitre.

2.1.1.1. Prédiction

La décorrélation par prédiction, souvent généralisé sous le nom de codage DPCM (*Differential Pulse Code Modulation*) [59], consiste à estimer la valeur probable \hat{x}_i d'un échantillon x_i à l'aide d'une information causale, et à ne conserver que l'erreur de cette prédiction qui sera quantifiée et codée pour former une valeur numérique (voir figure 2.2).

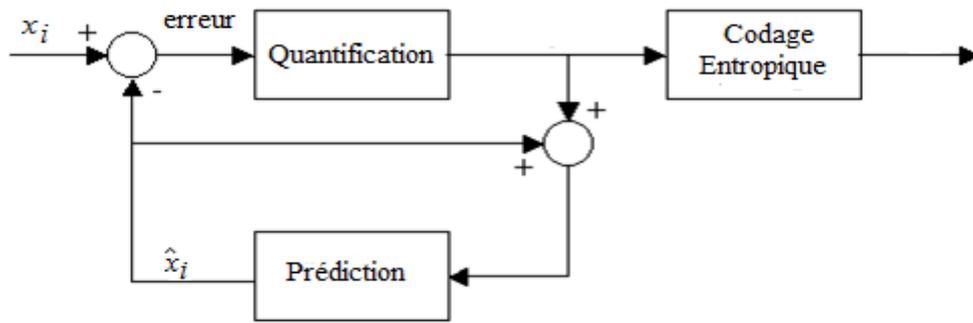


Figure 2.2. Schéma de compression par prédiction : Algorithme DPCM

2.1.1.2. Transformations

La transformation est un processus essentiellement sans perte qui convertit les données d'image à chaque pixel en une représentation qui tend à rendre plus facile l'élimination de la redondance spatiale. La transformée $T:f \rightarrow \hat{f}_T$ permet de projeter l'image sur une base de fonctions dont les propriétés doivent être capable de tenir compte des caractéristiques structurelles propre à l'image. Pour la compression, on utilise la transformée inversible T^{-1} telle que $T^{-1}(T(f))=f$ avec de bonnes propriétés de décorrélation afin que l'information contenue dans l'espace transformée soit plus compacte que dans l'espace original. Dans la plupart des systèmes de compression, la transformation produit des ensembles de coefficients appelés *sous-images de coefficients* qui doivent être moins corrélés et moins redondants. Compte tenu d'un taux total souhaité de bits, le dispositif d'attribution de bits détermine le nombre de bits qui doit être attribué à chaque sous-image de coefficients. Après la transformation, une grande partie de l'énergie de l'image est concentrée sur un faible nombre de coefficients.

Le choix de la transformée dépend du compromis débit-distorsion désiré. Il existe plusieurs types de transformées qui tendent vers une meilleure prise en compte des caractéristiques de l'image. Par exemple, la transformée en ondelettes utilisée dans les standards de compression JPEG 2000 et CCSDS (*Consultative Committee for Space Data Systems*).

Récemment, il est devenu évident que la transformée en ondelettes séparables n'est pas adaptée à la représentation des régularités géométriques dans l'image. Cette limitation est due essentiellement au fait que les seules directions bien appréhendées sont les directions horizontales et verticales. Beaucoup de coefficients d'ondelettes de forte amplitudes sont nécessaires à la représentation des structures géométriques. Afin de mieux les représenter avec plus de parcimonie, de nouvelles transformées dérivées des ondelettes ont été proposées (voir chapitre 1). Dans ce chapitre, on s'intéressera essentiellement à la compression progressive des images en utilisant la transformée en bandelettes.

2.1.2. Quantification

L'étape de quantification compte parmi les sources de perte d'information dans un processus de compression de manière irréversible [43,49,52,53,58-61]. C'est une phase cruciale pour une meilleure qualité visuelle, et quantitative du processus de par le fait qu'elle permet une réduction plus au moins importante du débit binaire (nombre de bit par pixel). Il existe deux types de quantification : la quantification scalaire (QS) et la quantification vectorielle (QV). La fonction de quantification Q appliquée à un coefficient ou un bloc de coefficients d'une fonction f est définie par :

$$Q: \mathbb{R} \rightarrow C$$
$$f \rightarrow Q(f) = \arg \min_{f_A \in C} D = \arg \min_{f_A \in C} \|f - f_A\|^2 \quad (2.2)$$

avec D : la distorsion engendrée par la quantification, C : le dictionnaire de l'ensemble des approximations f_A .

Dans ce qui suit, nous développerons la quantification scalaire qui sera utilisée dans les méthodes de compression par ondelettes et ondelettes géométriques.

Quantification scalaire

La quantification scalaire est réalisée indépendamment pour chaque élément, elle consiste à diviser l'intervalle du signal en un certain nombre d'intervalles $[b_i, b_{i+1}]$ où chaque intervalle est représenté par un indice (voir figure 2.3. (a)). Il existe deux types de quantification scalaire, la quantification scalaire uniforme et la quantification scalaire non uniforme. Dans le cas d'une quantification uniforme les représentants du dictionnaire sont répartis uniformément sur l'axe des réels, entre autres la distance q entre deux représentants aussi dite "*pas de quantification*" est constante. Le quantificateur Q est donné par la formule ci-après :

$$Q(x) = \text{signe}(x) \cdot \left\lfloor \frac{x}{q} \right\rfloor \quad (2.3)$$

où $\lfloor \cdot \rfloor$ représente la fonction associant à un réel son plus proche entier, et x un échantillon.

Les quantificateurs scalaires uniformes à zone morte (Dead zone) occupent une place centrale dans la compression par ondelettes et ondelettes géométriques. Ils permettent d'éliminer les coefficients non significatifs de faibles magnitudes au profit des coefficients significatifs quantifiés plus finement. En effet, ces quantificateurs possèdent un intervalle autour de zéro plus

large (zone=2.q) ce qui permet de ramener toutes les faibles valeurs quantifiées à zéro (voir figure 2.3. (b)).

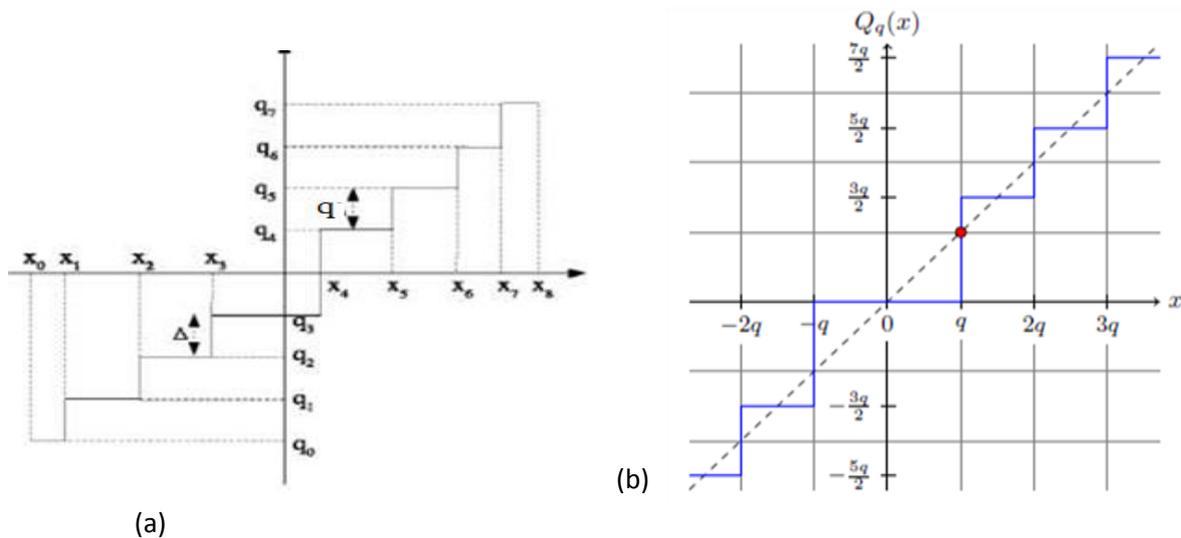


Figure 2.3. Quantification scalaire, (a) : uniforme ; (b): uniforme avec zone morte [52]

2.1.3. Codage entropique

Les valeurs qui se dégagent du processus de quantification sont généralement représentées par un code binaire de longueur fixe (N bits / symbole). L'étape de codage permet de réduire le nombre moyen de bits alloués à chaque valeur quantifiée. Il existe deux grandes familles de codage sans perte et réversibles: les codeurs entropiques et les codeurs par plage.

Les bases de codage entropique statistique ont été posées par D. Huffman [55]. Dans ce codage, on assigne à chaque symbole un code de taille inversement proportionnelle à sa probabilité. Le codeur de Huffman est un arbre binaire ordonné par tous les symboles et par leurs fréquences d'apparition; il est utilisé dans le standard de compression JPEG [62]. Une autre approche du codage entropique est le codage arithmétique utilisé dans la norme JPEG 2000. Le codage arithmétique est généralement plus performant que le codage de Huffman, il tend vers la limite inférieure théorique de Shannon, et il permet d'atteindre des taux très proches de l'entropie théorique. Le codage arithmétique a la capacité de coder chaque symbole sur un nombre non entier de bits, il n'assigne pas un mot de code à chaque symbole, mais il associe un point de l'intervalle [0,1] à un ensemble de symboles.

D'autres méthodes de codage basées sur la dépendance statistique ont été développées telles que le codeur RLE (Run-Length-Encoding) qui tient compte de la corrélation entre les

coefficients successifs [63]. En effet, dans le codage RLE, le signal est découpé en plages de symboles successifs et identiques, auxquelles on associe leur cardinal. Cela fonctionne bien dans le cas où la série constituée de coefficients comporte une forte dépendance markovienne (par exemple de longues plages de zéros entre lesquelles s'intercalent quelques valeurs non nulles).

2.1.4. Mesure des performances en compression

Afin d'évaluer et de sélectionner un algorithme de compression, il est nécessaire de mesurer les performances de ces algorithmes en considérant les points de vue visuel et quantitatif. De nombreuses métriques peuvent être utilisées pour mesurer la quantité de compression achevée, la qualité objective et subjective des images reconstruites [50,52,58-61,64,65]. Parmi ces critères, on utilise les mesures définies dans les sous-sections suivantes.

2.1.4. Taux de compression et débit binaire

Le taux de compression est le rapport entre la taille de l'image originale et la taille de l'image compressée en bit. Une autre variante du rapport de compression peut-être liée directement au débit binaire (*bit Rate*) exprimé en bits/pixel (bpp) qui mesure le nombre moyen de bits pour représenter un pixel de l'image. Le débit binaire correspond aussi au volume des données compressées divisé par le nombre de pixels dans l'image, ou celui entre la résolution numérique de l'image divisée par le taux de compression. Le taux de compression RC est donné par :

$$RC(\%) = \frac{\text{nombre de bits codés}}{\text{nombre de bits de l'image originale}} \times 100 \quad (2.4)$$

Le débit binaire est alors exprimé par :

$$\text{débit}(bpp) = \frac{\text{nombre de bits par pixel dans l'image originale}}{RC} \quad (2.5)$$

Le taux de compression que l'on peut atteindre par une méthode de compression sans perte est totalement dépendant de l'image d'entrée. Il est limité par l'entropie de l'image suivant le théorème de Shannon.

2.1.4.2. Le compromis débit-distorsion

Le taux de compression n'est pas le seul critère de performance d'un système de compression. Dans le cas d'une compression avec pertes, la qualité de l'image reconstruite doit aussi être prise en compte. Il y a donc un compromis à trouver entre le taux de compression et la

qualité de l'image, on parle alors du compromis débit-distorsion. Pour un signal f donné, soit $R(Q)$ le débit pour une configuration des paramètres de compression Q (généralement le quantificateur) et $D(Q)$ la distorsion entre l'image originale et l'image reconstruite après compression, alors la fonction débit-distorsion $D(R_m)$ est définie d'une manière théorique comme la solution au problème de minimisation de la distorsion pour un débit maximal R_m , et la fonction distorsion-débit $R(D_m)$ est définie comme celle du plus petit débit pour une distorsion maximale D_m . Ces problèmes d'optimisation sous contraintes peuvent s'écrire :

$$\begin{aligned} D(R_m) &= \min_{Q \in \mathbb{Q}, R(Q) \leq R_m} D(Q) \\ R(D_m) &= \min_{Q \in \mathbb{Q}, D(Q) \leq D_m} R(Q) \end{aligned} \quad (2.6)$$

La distorsion est une mesure de l'erreur commise entre l'image originale et l'image reconstruite. De nombreuses métriques peuvent être utilisées pour mesurer les distorsions, la majorité repose sur la norme de Minkowski [64]. La plus utilisée pour l'optimisation débit-distorsion est l'erreur quadratique moyenne EQM (MSE : *Mean Squared Error*) entre l'image originale f_o et l'image compressée f_c de taille $M \times N$; formellement elle est définie par :

$$EQM(f_o, f_c) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f_o(i, j) - f_c(i, j))^2 \quad (2.7)$$

Les valeurs de la EQM étant souvent très élevées à bas débit, le PSNR (Peak Signal to Noise Ratio) qui s'exprime en décibel (dB) est souvent préféré pour comparer des courbes de distorsions, il est exprimé par :

$$PSNR(f_o, f_c) = 10 \log_{10} \frac{(2^r - 1)^2}{EQM} \quad (2.8)$$

où r est la résolution numérique de l'image. Ces critères ne sont pas forcément représentatifs des évaluations subjectives visuelles humaines, cela est dû au fait que le système visuel humain ne considère pas l'image point par point, mais extrait certaines propriétés spatiales, chromatiques et temporelles (préservation des contours et des textures, présence d'un effet de bloc, lissage de l'image, masquage des détails, etc.). Toutefois, EQM et $PSNR$ ont l'avantage d'être très faciles à implémenter et ils sont considérés comme une référence pour la mesure des performances en compression des images.

2.2. Compression progressive des images basée sur la transformée en ondelettes

2.2.1. Introduction

Un grand nombre de méthodes de compression d'images parmi les plus récentes et les plus efficaces reposent sur les ondelettes. La plupart de ces méthodes sont basées sur l'exploitation des redondances entre les coefficients obtenus par une transformation en ondelettes. La transformée en ondelette est efficace pour concentrer l'énergie de l'image dans un nombre relativement restreint de coefficients; en plus leurs positions dans l'image transformée sont corrélées. On s'intéresse dans cette section à l'application des ondelettes introduites dans le chapitre 1, dans le domaine de compression des images fixes (Image SAR et image médicale). L'idée générale est d'utiliser la transformée en ondelette qui soit quasiment réversible et qui permette de concentrer l'énergie de l'image sur un nombre suffisamment réduit de coefficients. Cette phase est suivie d'une quantification des coefficients, une exploitation des corrélations est mise en avant par la structure de la transformation, et finalement un codage entropique des coefficients quantifiés. Par ailleurs, les transformées en ondelettes offrent naturellement une représentation progressive de l'image, il est intéressant de conserver cette propriété lors du codage des sous-bandes [66,67]. Ainsi, dans les codeurs emboîtés (*embedded*), la quantification et le codage sont également réalisés de manière progressive, en commençant par coder partiellement les coefficients de plus forte amplitude, puis en raffinant la quantification de ces derniers et en encodant de nouveaux. Nous présentons ici quelques codeurs progressifs basés sur des structures d'arbres ou de blocs.

Ces méthodes appliquent une quantification par approximations successives pour améliorer la précision de la représentation des coefficients d'ondelettes et pour faciliter le codage emboîté. Avec cette approche, la signification des coefficients d'ondelettes pour une série monotone décroissante de seuils T_n est enregistrée dans un ensemble de cartes binaires, appelées cartes de signification, correspondant chacune à un plan de bits. La Figure 2.4 illustre cette notion de plan de bits.

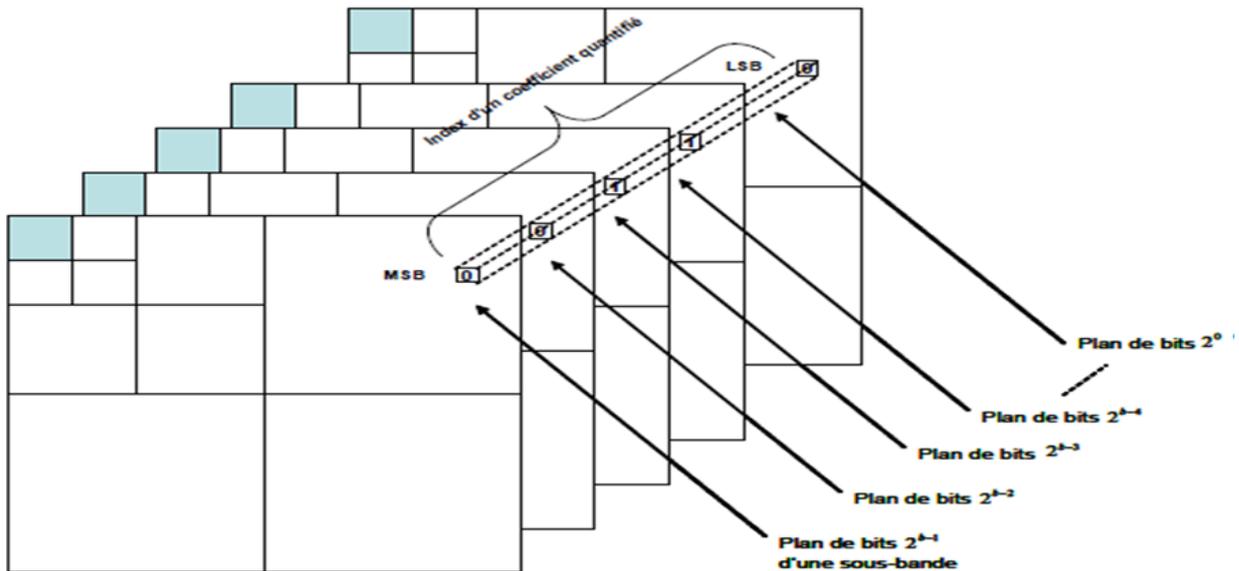


Figure 2.4. Représentation des plans de bits allant du bit le plus significatif (MSB) vers le moins significatif (LSB) pour une transformée en ondelettes 2D [64]

Nous distinguons deux types d'approches pour les standards de compression actuels basés sur la transformée en ondelette : l'approche inter-bandes qui utilise les redondances inter-échelles entre les sous-bandes pour coder les coefficients d'ondelettes, et celle intra-bandes dans laquelle les sous-bandes sont codées de façon indépendante [66-70].

Etant donné une image f ayant une régularité géométrique C^α , la meilleure approximation f_M avec M coefficients dans une base d'ondelettes satisfait la condition suivante [6-8]:

$$\|f - f_M\|_{L^2([0,1]^2)}^2 \leq CM^{-1} \quad (2.9)$$

avec C une constante qui ne dépend que de f .

2.2.2. Codage Inter sous-bandes

Les premiers codeurs de sous-bandes issues d'une décomposition spatiale par ondelette, tentent d'utiliser la redondance d'information existant entre les coefficients d'ondelettes de sous-bandes successives et de même orientation, à travers les différentes échelles afin d'améliorer la compression.

2.2.2.1. L'algorithme EZW

EZW (*Embedded Zerotree Wavelet*) est le premier algorithme de codage progressif inter-bandes, proposé par J. Shapiro [68]. Le principe de ce codeur repose essentiellement sur l'exploitation de la notion multiéchelle associée aux ondelettes, pour caractériser les dépendances

inter-bandes parmi les coefficients d'ondelettes localisés dans les sous-bandes ayant la même orientation. Son schéma de codage utilise un modèle basé sur l'hypothèse des arbres de zéros (*zerotree*) qui suppose que si un coefficient d'ondelettes c_w est non significatif pour un seuil donné T , c'est à dire $|c_w| < T$, alors tous les coefficients de la même orientation dans la même localisation spatiale à des résolutions plus fines sont supposés non significatifs pour ce même seuil T (voir figure 2.5).

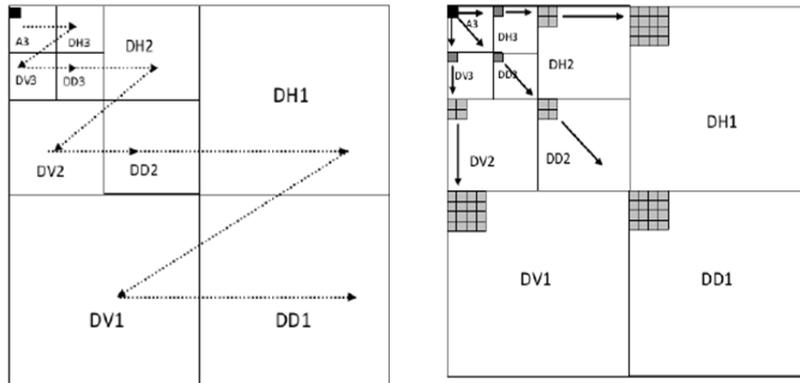


Figure 2.5. Relation entre les coefficients d'ondelettes dans les différentes sous-bandes; (a): Ordre de balayage des coefficients, (b): La dépendance parents-enfants dans des coefficients d'ondelettes.

Chaque coefficient parent à une résolution donnée, engendre quatre coefficients "enfants" à la résolution supérieure. La résolution parent-enfant pour le codeur EZW est exprimée comme suit:

$$O(i, j) = \{(2i, 2j), (2i, 2j+1), (2i+1, 2j), (2i+1, 2j+1)\} \quad (2.10)$$

L'algorithme commence d'un seuil $T_n = 2^n$, avec $n = \log_2 \left(\left\lceil c_{w_{\max}} \right\rceil \right)$ avec $c_{w_{\max}}$: la valeur maximale en valeur absolue sur toutes les sous-bandes. Une fois que le seuil est fixé, le codeur EZW commence par tester si les coefficients d'ondelettes sont significatifs ($(|c_w| \geq T_n)$), alors ils seront codés suivant leur signe par POS (positif) et NEG(négatif), et ensuite ils seront mis dans une carte significative. Si les coefficients sont testés non significatifs, tous leurs descendants seront examinés à leur tour pour tester leurs significations. Dans le cas où aucun descendant n'est significatif un arbre de zéros (AZ) est codé, si un descendant significatif apparaît, il sera codé un zéro isolé (ZI). Les coefficients qui descendent d'une racine d'arbre de zéros ne seront pas codés. Cette étape est appelée la passe dominante ou passe de signification. L'étape suivante, sera la passe de raffinement, elle s'opère d'une façon itérative sur les coefficients de la carte de signifiante. Pour chaque coefficient de cette liste, le bit situé dans le plan de bit inférieur (plan de bit 2^{n-1}) est codé. L'encodeur divise le seuil T_n par 2 ($T_n \leftarrow T_{n-1} / 2$) et exécute une nouvelle passe

dominante et de raffinement, cette procédure se poursuit jusqu'à ce que l'on atteigne le débit voulu.

2.2.2.2. L'algorithme SPIHT

L'algorithme SPIHT (*Set Partitioning In Hierarchical Trees*) [69] reprend les principes évoqués dans EZW tout en proposant de partitionner récursivement les arbres de coefficients. Ainsi, là où EZW codait un coefficient non significatif isolé ('ZI'), SPIHT effectue un partitionnement récursif de l'arbre de manière à déterminer la position des coefficients significatifs dans la descendance du coefficient considéré. Le codeur SPIHT est essentiellement basé sur une approche différente pour la transmission des cartes significatives. Pour un coefficient (i, j) donné, on définit les sous-ensembles :

- $D(i, j)$: l'ensemble de sa descendance ;
- $O(i, j)$: ses descendants directs (ses "fils") ;
- $L(i, j)$: le reste de sa descendance (petits-enfants et leur éventuelle descendance), d'où $L(i, j) = D(i, j) - O(i, j)$

Pour un coefficient donné on vérifie si $D(i, j)$ contient au moins un coefficient significatif. Dans ce cas, on le décompose en $O(i, j)$ et $L(i, j)$. On teste ensuite la significativité de chacun de ses fils (les éléments de $O(i, j)$), si l'élément de $L(i, j)$ est significatif, on le divise en sous-ensembles $D(i', j')$, où (i', j') sont les éléments de $O(i, j)$. Ensuite, on réitère le processus. Donc une partition des sous-bandes en sous-ensembles regroupant des coefficients non significatifs est effectuée selon un procédé différent de l'algorithme EZW, une telle méthode permet en fait de tirer partie des corrélations entre sous-bandes aussi bien entre deux niveaux consécutifs de résolution (relations père-fils), qu'entre sous-bandes voisines (relations de cousins) qui présentent également des concentrations d'énergie en des zones similaires. On isole ainsi des zones de coefficients non significatifs plus vastes que dans l'algorithme EZW; c'est ce qui permet de dépasser les performances de cet algorithme.

La figure suivante illustre les résultats obtenus en comparant les deux encodeurs inter-bandes sur une image du toolbox Matlab.

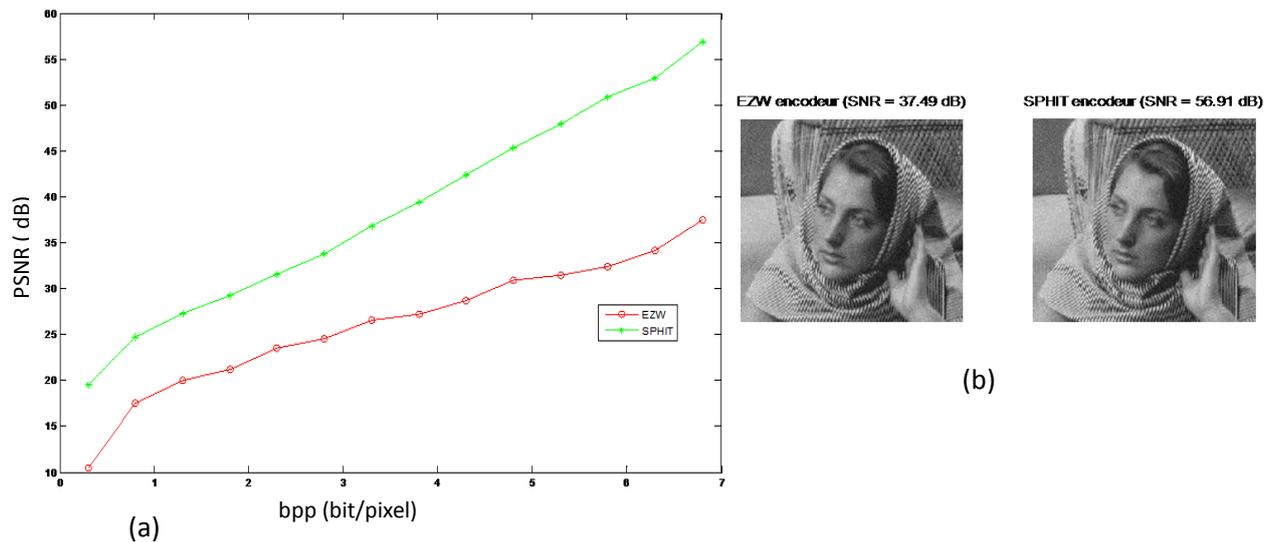


Figure 2.6. Etude comparative entre EZW et SPHIT, (a): PSNR en fonction bpp et (b) : analyse visuel pour différents PSNR

2.2.3. Codage Intra sous-bandes

Le codage intra-bandes ne prend pas en compte la redondance inter-bande et tente d'optimiser la compression des différentes sous-bandes indépendamment.

Le codeur EBCOT

L'algorithme EBCOT (*Embedded Block Coding with Optimized Truncation*) est issu des travaux de Taubman [70] mais n'appartient pas à la famille des schémas de codage à arbre de zéros. Dans EBCOT, chaque sous-bande est tout d'abord découpée en petits blocs indépendants (typiquement la taille d'un code bloc est de 64×64 coefficients, mais elle peut prendre toutes valeurs de $4.2^{e_1} \times 4.2^{e_2}$ tant que $e_1 + e_2 \leq 8$): les codeblocks (voir figure 2.7 (a)). Ces derniers sont alors codés de façon progressive, par plans de bits et au moyen d'un codeur arithmétique contextuel. Enfin, une procédure d'optimisation débit-distorsion est utilisée pour déterminer le choix optimal des codeblocks à conserver pour assurer une qualité maximale pour un ensemble de débits donnés et connus à l'encodage : les points de troncatures. Les débits intermédiaires restent toutefois accessibles mais n'offriront peut-être pas une qualité de reconstruction optimale.

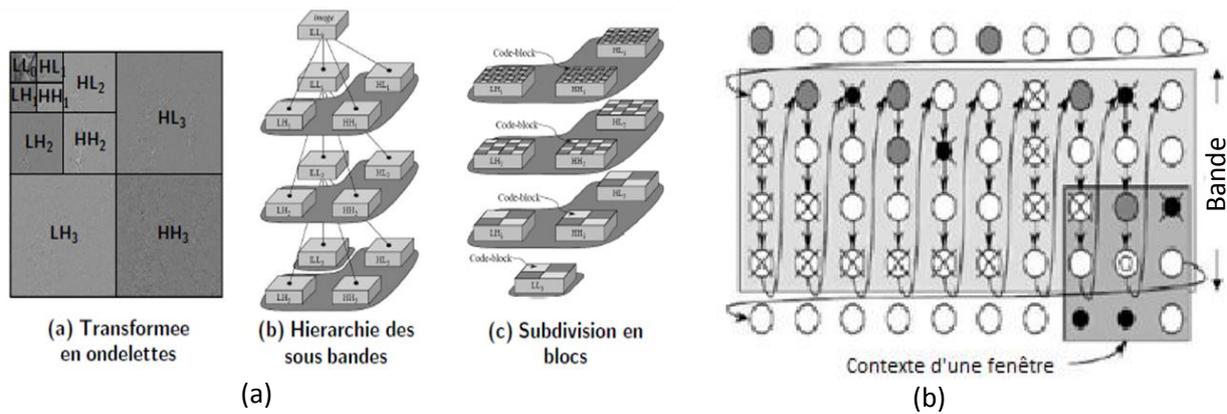


Figure 2.7 . Structure du codec EBDCOT [12], (a): Subdivision en blocs 64×64 des coefficients d'une transformée en ondelettes; (b) : Déplacement de la fenêtre de contexte lors du codage d'un plan de bits [8].

Pour chaque bloc, un flot de bits est produit pour coder de façon progressive les coefficients à l'intérieur du bloc. Des seuils dyadiques de la forme $T_k = 2^{-k} T_0$ sont utilisés pour quantifier les coefficients. Chaque coefficient c_n , quantifié à la précision la plus fine T_0 , peut être représenté sous la forme :

$$Q_{T_0}(c_n) = \sigma(c_n) \sum_{k \geq 0} c_n^k \cdot 2^k \cdot T_0 \quad \text{avec} \quad \begin{cases} \sigma(c_n) \in \{-1, 1\} \\ c_n^k \in \{0, 1\} \end{cases} \quad (2.11)$$

Pour un k fixe, l'ensemble des bits $c(k)$ forment un plan de bits et EBCOT code cet ensemble de bits en les parcourant de manière séquentielle (voir figure 2.7. (b)).

2.2.4. Exploitation de la dépendance entre les coefficients d'ondelettes

Les transformées en ondelettes ont été utilisées pour décorrélérer et réduire ainsi les redondances contenues dans les images naturelles. Il a été montré que cette décorrélation n'est pas totale, et qu'il fallait utiliser des codeurs en sous-bandes pour exploiter la corrélation restante afin d'augmenter l'efficacité de la compression [71]. Liu et Moulin [72] ont estimé ces dépendances statistiques intra et inter échelles en mesurant l'information mutuelle entre les coefficients d'ondelettes dans un voisinage très local. Delaunay et al. [52,73], ont proposé une autre approche basée sur l'estimation de la corrélation pour mesurer les dépendances statistiques entre les coefficients d'ondelettes voisins. Ils ont défini un ensemble de sept voisins d'un coefficient c_w (représenté sur la figure 2.8). Pour cela, ils ont utilisé l'information mutuelle relative $I_r(X, Y)$ qui peut être interprétée comme la proportion d'information qu'il est possible d'économiser pour le codage de la variable aléatoire X (le coefficient c_w), si la variable aléatoire

Y (coefficient voisin) est parfaitement connue. $I_r(X, Y)$ est exprimée en fonction de l'entropie d'ordre 0 par :

$$I_r(X, Y) = \frac{2(H_0(X) - H_0(X/Y))}{H_0(X) + H_0(Y)} \quad (2.12)$$

Il a été montré dans les travaux de Delaunay que la corrélation augmente avec l'échelle de la transformée et qu'il est possible d'exploiter les relations entre coefficients d'ondelettes dans un voisinage très local. La matrice de coefficients de corrélation calculée pour chaque classe de blocs, représente les corrélations intra-bloc entre toutes les paires de coefficients d'ondelettes. Le coefficient de corrélation $r_{X,Y}$ entre deux coefficients d'ondelettes X et Y d'une classe de n blocs est défini par :

$$r_{X,Y} = \frac{1}{(n-1)} \frac{1}{\sigma_x \sigma_y} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (2.13)$$

où x_i, y_i : représentent respectivement les valeurs des coefficients d'ondelettes X et Y dans le bloc numéroté $i = 1, 2, \dots, n$; \bar{x}, \bar{y} : sont respectivement les moyennes des x_i et y_i ; σ_x, σ_y : les variances des x_i et y_i .

En effet, dans les sous-bandes à la résolution la plus fine, l'information mutuelle relative approche les 15% pour les voisins 1 et 2. Elle est autour de 5% pour le voisinage inter- sous-bandes (voisins 5, 6 et 7). Cette étude a montré que la plupart des dépendances statistiques très locales liées à la structure des images naturelles sont observables à l'intérieur d'une même sous-bande.

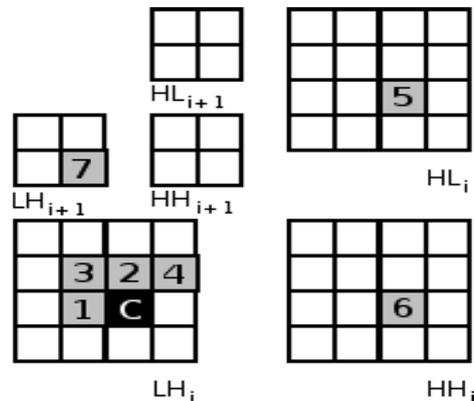


Figure 2.8. Définition d'un voisinage dans le domaine ondelettes

Pour des raisons de simplicité et d'efficacité, les ondelettes ont souvent été utilisées de manière séparable sur les axes horizontal et vertical. Il en résulte une décorrélation partielle de l'image, qui se traduit par la présence de nombreux coefficients de forte énergie le long des contours. Bien que la dépendance résiduelle soit néanmoins réduite et partiellement exploitée par les codeurs de

sous-bandes dans le cas de la compression d'image. Il semble plus intéressant de chercher une transformée qui résout ce problème en filtrant directement le long des contours de l'image. La transformée en ondelettes est optimale pour la représentation de singularités ponctuelles. Malheureusement, dans les images, la plupart des singularités rencontrées sont des contours. En effet, cette transformée n'est capable d'exploiter les régularités géométriques que dans les directions des deux axes, puisque les filtres sont appliqués horizontalement et verticalement, alors que les contours interviennent dans toutes les orientations.

2.3. Compression des images basée sur l'utilisation de la transformée en bandelettes

2.3.1. Introduction

La transformée en ondelettes possède des propriétés intéressantes du point de vue compression. Cependant, elle a le désavantage de ne pas prendre en compte efficacement les structures géométriques, et de présenter des dépendances statistiques entre ces coefficients. Pour palier à ces problèmes, Mallat et al. [6-8,41,42] ont introduit de nouvelles transformations dérivées des ondelettes qui permettent d'exploiter au mieux la régularité géométrique le long des contours, et d'éliminer les dépendances statistiques dès l'étape de la transformée. Il existe deux générations de la transformée en bandelettes (voir chapitre 1). La première génération a été proposée par Le Pennec et Mallat [6] et permet d'éliminer les redondances en effectuant une bandelétisation. Les contours sont représentés par des courbes paramétriques le long desquelles une ondelette séparable est déformée. Les bandelettes sont des ondelettes déformées le long des lignes de flots, l'opération revient à effectuer une transformée en ondelettes 2D séparable sur une région déformée afin que les lignes de flots soient horizontales ou verticales. La seconde génération a été développée par Peyré et Mallat [8], elle est construite en se basant sur la transformée en ondelettes 2D. Ainsi, ils peuvent exploiter les régularités géométriques le long des contours dans les sous-bandes et ceci en effectuant une approximation polynomiale par morceau sur les bandes qui suivent la géométrie. Peyré a utilisé comme approximation polynomiale la transformée d'Alpert étendue au cas 2D à la place des polynômes 1D [6]. Il existe deux versions de cette transformée en bandelette pour la compression des images. La première est la transformée en bandelettes d'ordre zéro, qui considère les géométries linéaires à l'intérieur des blocs issus d'une segmentation en *quadtree* des sous-bandes de la transformée en ondelettes de l'image. La transformée de Haar monodimensionnelle (équivalente à la transformée d'Alpert d'ordre 0) est effectuée sur les coefficients d'ondelettes de chaque bloc réordonnés selon la direction de la

géométrie. La seconde version est la transformée en bandelette par regroupement, elle ne possède pas la propriété de multirésolution sur les bandes car elle ne repose pas sur l'analyse multirésolution par ondelettes d'Alpert. Dans cette version, les sous-bandes de la transformée en ondelettes sont segmentées en blocs de taille 4×4 ce qui permet d'économiser le calcul et la transmission du *quadtree*.

Dans la section suivante nous donnerons un aperçu sur la construction des bases des deux versions basée sur l'optimisation du critère débit-distorsion. Pour la partie expérimentale, nous avons choisi de nous intéresser à la transformée en bandelette d'ordre 0 (voir figure 1.4 chapitre 1) qui constitue un moyen efficace pour l'amélioration des résultats de compression par rapport à ceux obtenus par la transformée en ondelettes.

2.3.2. Construction de bases de bandelettes par optimisation débit-distorsion

2.3.2.1 Transformée en Bandelettes d'ordre 0

La transformée en bandelette repose essentiellement sur l'application de la transformée en ondelettes. On remarque sur la figure 2.9 que seules les coefficients non nuls sont situés au voisinage de la singularité géométrique. Donc, il faut se concentrer sur ces coefficients à l'échelle 2^j et orientation k (avec $k = H, V, D$) fixés pour obtenir une meilleure approximation de l'image f .

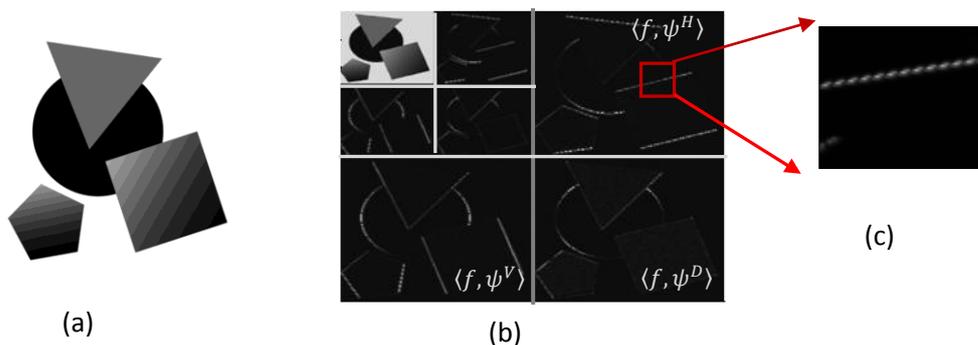


Figure 2.9. Exemple d'une image avec sa régularité géométrique, (a) : image originale, (b): Décomposition en ondelette $j=2$ et (c) : zoom sur les coefficients non nuls le long de la géométrie

L'étape de la bandelétisation permet d'enlever la corrélation qui existe entre les coefficients en ondelettes près des singularités [6,8]. Le long de la régularité de la sous-bande, la fonction transformée $f * \psi_j^k$ varie lentement dans la direction de la géométrie. Donc, cette régularité sera utilisée pour comprimer les coefficients en ondelettes non nuls restants. L'expression de la fonction transformée est donnée formellement par :

$$f_j^k(n_1, n_2) = \langle f, \psi_{jn}^k \rangle = f * \psi_j^k(2^j x_1 - n_1, 2^j x_2 - n_2) \quad (2.14)$$

où $\psi_j^k(x) = 2^{-j} \psi^k(-2^{-j}x)$, $k \in \{V, H, D\}$ et ψ_{jn}^k : la fonction ondelette dont le support est inclut dans $[0,1]^2$. La base de bandelette discrète pour une fonction f ayant une régularité géométrique C^α est définie par :

$$B^*(\Gamma) = \arg \min_{B^*(\Gamma) \in D_{T^2}} (\mathcal{L}(f, B(\Gamma), T)) \quad (2.15)$$

avec \mathcal{L} : le lagrangien , D : le dictionnaire de toutes les bases $B(\Gamma_j^k)$ obtenues à l'aide de toutes les segmentations en quadtree, et de toutes les géométries à l'intérieur du carrée S de largeur $b = L.2^j / L$: une puissance de 2.

il est à noter que l'approximation f_M de f pour un seuil de quantification T dans la base $B^*(\Gamma)$ doit vérifier la condition suivante:

$$\|f - f_M\|^2 \leq C.M^{-\alpha} \quad (2.16)$$

où $M = M_B + M_s + M_G$ avec M_B : le nombre de coefficients en bandelettes supérieurs à T , M_s : le nombre nécessaire pour spécifier les quadtrees Q_j^k , M_G : le nombre de coefficients pour spécifier la géométrie dans le carré $S \in Q_j^k$ et C : une constante qui ne dépend que de f .

Pour compresser une fonction f dans une base de bandelette $B(\Gamma)$, les coefficients en bandelettes quantifiés et les paramètres de la géométrie seront codés avec un codage entropique pour former le train binaire compressé. La quantification est une quantification uniforme de pas T avec une zone morte de $2T$ définie par :

$$\begin{cases} Q_T(x) = 0 & \text{si } |x| < T \\ Q_T(x) = \text{signe}(x)(k+1/2).T & qT \leq |x| \leq (k+1).T \end{cases} \quad (2.17)$$

Dans ce qui suit, l'algorithme de compression à base de transformée en bandelette d'ordre 0 qui inclut à la fois la meilleure base de bandelette, la décomposition de la fonction f et le calcul de l'approximation pour un seuil de quantification T fixé par l'utilisateur.

Algorithme de la transformée en bandelette d'ordre 0

Etape 1 Transformée en ondelettes 2D

Calcul de la transformée en ondelette 2D séparable sur l'image originale f de taille $N \times N$, les vecteurs de coefficients en ondelettes $\{f_j^k\}$ obtenus pour chaque échelle 2^j et orientation $k \in (V, H, D)$ sont stockés dans une unique image ayant la même taille que l'image originale.

Etape 2 Sélection de chaque carrée dyadique

Subdivision du support $[0,1]^2$ en quatre sous-blocs de tailles identiques pour définir les carrés dyadiques S . La largeur de chaque carré est L pixels tel que $4 \leq L \leq 2^{-j/2}$. Pour chaque carré dyadique S , on regroupe les coefficients f_j^k dans un vecteur f_S . Les étapes suivantes sont répétées pour chaque dyadique S et à chaque échelle et chaque orientation.

Etape 3 Sélection de chaque géométrie

Dans cette étape, on considère toutes les directions d joignant un couple de points $(2^j n, 2^j n') \in S$ de taille L (le nombre potentiel de directions est inférieur à $2L^2$). Pour récupérer la régularité géométrique qui existe autour de la direction, il faut réordonner les positions des points échantillons dans la grille. Pour cela, pour chaque direction d parallèle à la géométrie, la position x de chaque échantillon est projetée sur la ligne perpendiculaire à d pour obtenir un nouveau point \tilde{x} (voir figure 2.10). L'ensemble des numéros des nouveaux points formera le signal discret 1D f_d donné par : $\forall i; f_d(i) = f(x_i)$.

La régularité directionnelle 2D de f_S devient une régularité isotrope 1D de f_d .

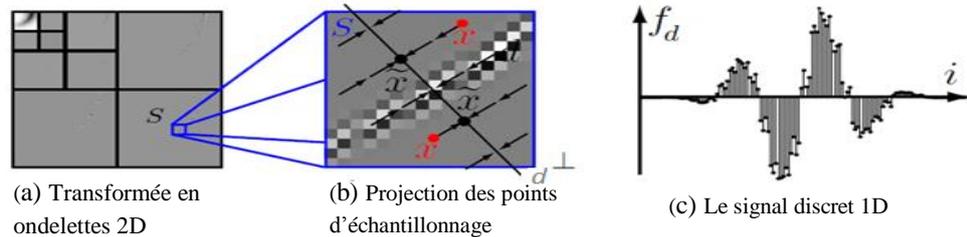


Figure 2.10. Réordonnement discret des points d'échantillonnage [8]

Etape 4 Transformée en ondelette 1D

Application de la transformée de Haar monodimensionnelle au signal discret f_d pour obtenir les coefficients en bandelettes b_k . La transformée de Haar est choisie pour son efficacité à capturer les régularités géométriques 1D.

Etape 5 Sélection de la meilleure géométrie

Pour sélectionner la meilleure géométrie, on choisit la direction d qui minimise le lagrangien. Pour la compression il est donné par :

$$\mathcal{L}(f_d, f_{dR}) = \|f_d - f_{dR}\|^2 + \lambda T^2 (R_G + R_B) \quad (2.17)$$

avec f_{dR} : le signal contenant les coefficients en bandelettes quantifiés $\{Q_T(b_k)\}$,

$R = R_G + R_B$: le nombre de bits nécessaires au codage et $\lambda = 3/8$ [6].

les coefficients de bandelettes du carré S seront stockés dans une image 2D de même dimension.

Etape 6 Construction du quadtree

Une fois que l'on a calculé les coefficients pour chaque carré dyadique, on doit choisir la meilleure segmentation en carré. Ceci revient à conserver uniquement un sous-ensemble de carrés dyadiques de façon à obtenir une segmentation en quadtree.

Etape 7 Fin de l'algorithme

2.3.2.2. Transformée en Bandelettes par regroupement

Dans cette version, on n'utilise pas de quadtree ce qui simplifiera les calculs. Après l'application de la transformée en ondelettes, les sous-bandes sont segmentées en blocs 4×4 et un dictionnaire de bases orthonormales est défini. Chaque bloc de coefficients d'ondelettes est alors transformé dans chacune des bases de ce dictionnaire. Pour un bloc de coefficients d'ondelettes, la meilleure représentation de ce bloc en termes de débit-distorsion est retenue pour la compression. Les coefficients de bandelettes étant les coefficients obtenus par projection des coefficients d'ondelettes de chaque bloc dans une base polynomiale, la reconstruction des coefficients d'ondelettes à partir des coefficients de bandelettes quantifiés est une approximation polynomiale des coefficients d'ondelettes originaux. En plus, cette approximation polynomiale est directionnelle. La figure suivante illustre le concept de la transformée en bandelette par bloc [8,52].

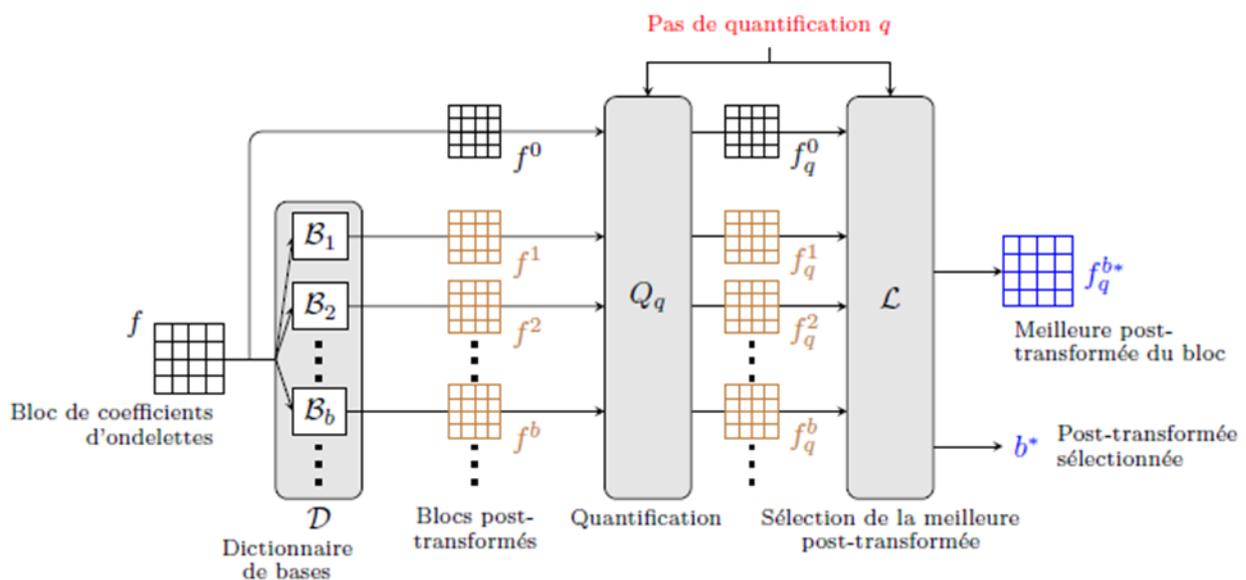


Figure 2.11. Processus de la transformée en bandelettes d'un bloc f de coefficients d'ondelettes [52]

2.3.3. Études expérimentales des performances de la transformée en bandelettes d'ordre 0

Dans cette section, nous présenterons les résultats expérimentaux obtenus à l'issue d'une étude comparative entre la transformée en ondelettes 2D séparables et la transformée en bandelettes 2G avec une quantification scalaire à zone morte, et un codage entropique classique sur un ensemble d'images médicales "Dicom" de la base d'image Medeisa et d'images naturelles. Par ailleurs, nous élargissons cette comparaison en utilisant le codeur SPHIT intégré avec les ondelettes et le codeur SPHIT modifié avec les bandelettes d'ordre 0. Pour la compression, nous

avons utilisé les filtres d'ondelettes biorthogonales 9/7 "CD 9x7" et un niveau de décomposition $j=3$. Les évaluations subjective et objective seront basées sur le PSNR pour différent débit (bpp), et l'analyse visuelle sur les images reconstruites en mettant l'accent sur l'effet d'artefact.

D'après les résultats, présentés (figure 2.12 et tableau 2.1), nous notons que les valeurs des *PSNR* obtenues par la transformée en bandelettes 2G sont meilleures par rapport aux résultats obtenus à partir des ondelettes. Pour les différents débits, nous avons obtenu d'une manière globale des améliorations en termes de *PSNR*, alors que dans certains cas particuliers les améliorations en termes de *PSNR* ont été relativement importantes.

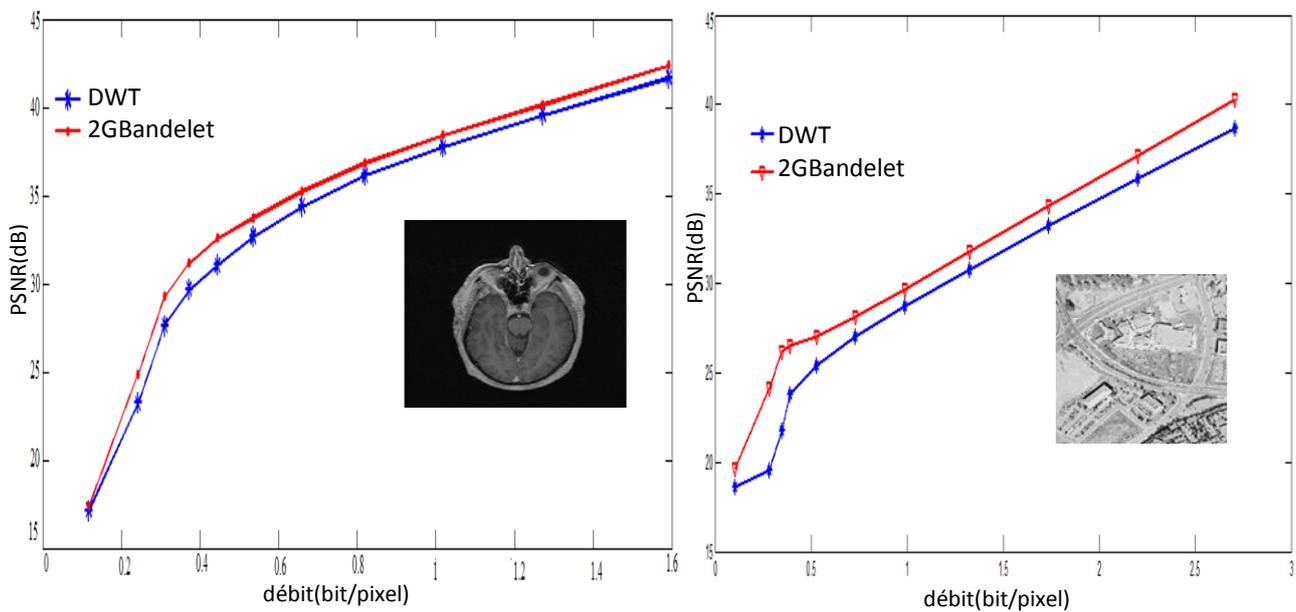


Figure 2.12. Courbe PSNR- débit pour différentes images testées, (a) : image médicale et (b): image aérienne

La figure 2.13 illustre le problème d'artefact lors de la reconstruction des images en utilisant les transformées en ondelettes et en bandelettes.

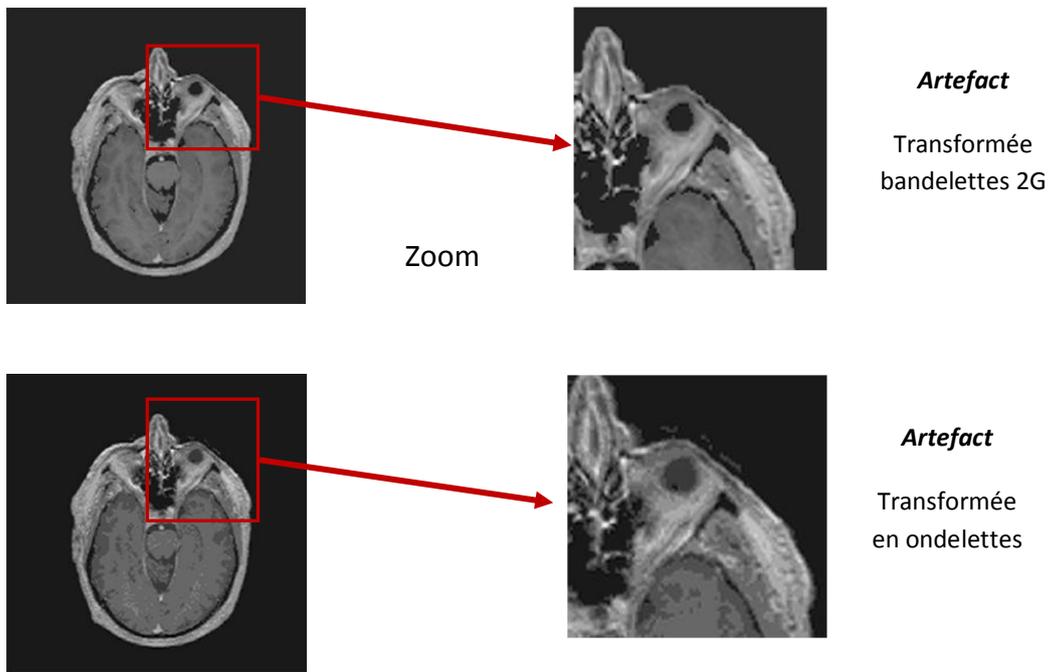


Figure 2.13. Reconstruction de l'image par bandelette 2G et ondelettes pour un débit $\text{bpp} = 0.73$ (bit/pixel): Problème d'artefact

Pour évaluer la robustesse de la transformée en bandelette d'ordre 0, nous avons proposé un schéma de compression progressive des images basé sur la transformée en bandelettes 2G combinée avec l'encodeur progressif SPIHT qui est généralement intégré avec la transformée en ondelettes. La figure suivante illustre les résultats obtenus pour différents PSNR et avec des seuils de quantification variables.

| Transformée en Ondelettes + SPHIT | | | |
|--------------------------------------|--------------|---------------|---------------|
| Psnr = 22.46 | Psnr = 26.44 | Psnr = 31.545 | Psnr = 34.931 |
| | | | |
| Transformée en bandelette 2G + SPHIT | | | |
| Psnr = 20.59 | Psnr = 28.42 | Psnr = 32.079 | Psnr = 36.153 |
| | | | |

Figure 2.14. Résultats comparatifs entre SPHIT+Bandelettes 2G et SPHIT + Ondelettes

L'algorithme SPIHT a été appliqué sur les coefficients de bandelettes pour obtenir une reconstruction progressive de l'image avec minimisation du compromis distorsion-débit. La transformée en bandelettes est appliquée sur chaque sous-bande de la décomposition en ondelettes et la structure arborescente est préservée. Ainsi, l'algorithme SPIHT sera adapté pour compresser les coefficients.

2.4. Etude comparative de la compression des images satellitaires basée sur les transformations directionnelles

Dans cette section, nous présentons une étude comparative pour la compression des images SAR, basée sur les transformées en ondelettes géométriques adaptatives (2G bandelettes) et non adaptatives (FRIT) avec une faible complexité. Nous évaluerons les performances des techniques de compression sur plusieurs images SAR, dont: Northrop Grumman SABR Radar image SAR de taille 512x512, image SAR avec mini-SAR (4 mètres dans la résolution) de Sandia National Laboratory de la taille 512x512, et Aerial image (image aérienne). Nous notons que la représentation géométrique multiéchelle pour la compression donne des résultats prometteurs en terme de taux de compression à faible débit binaire, et assure une réduction efficace du bruit multiplicatif "*speckle*". Dans le cas où l'image a des structures directionnelles dominantes; Les ondelettes géométriques sont robustes et efficaces; parce qu'elles peuvent mieux représenter les régularités géométriques. Alors que les ondelettes séparables sont plus adaptées à la représentation des images où les zones homogènes sont dominantes. La figure 2.15, représente la compression de l'image à plusieurs bits à débits : (bpp = 0,28 bits/pixel et bpp= 1,96 bits/pixel).

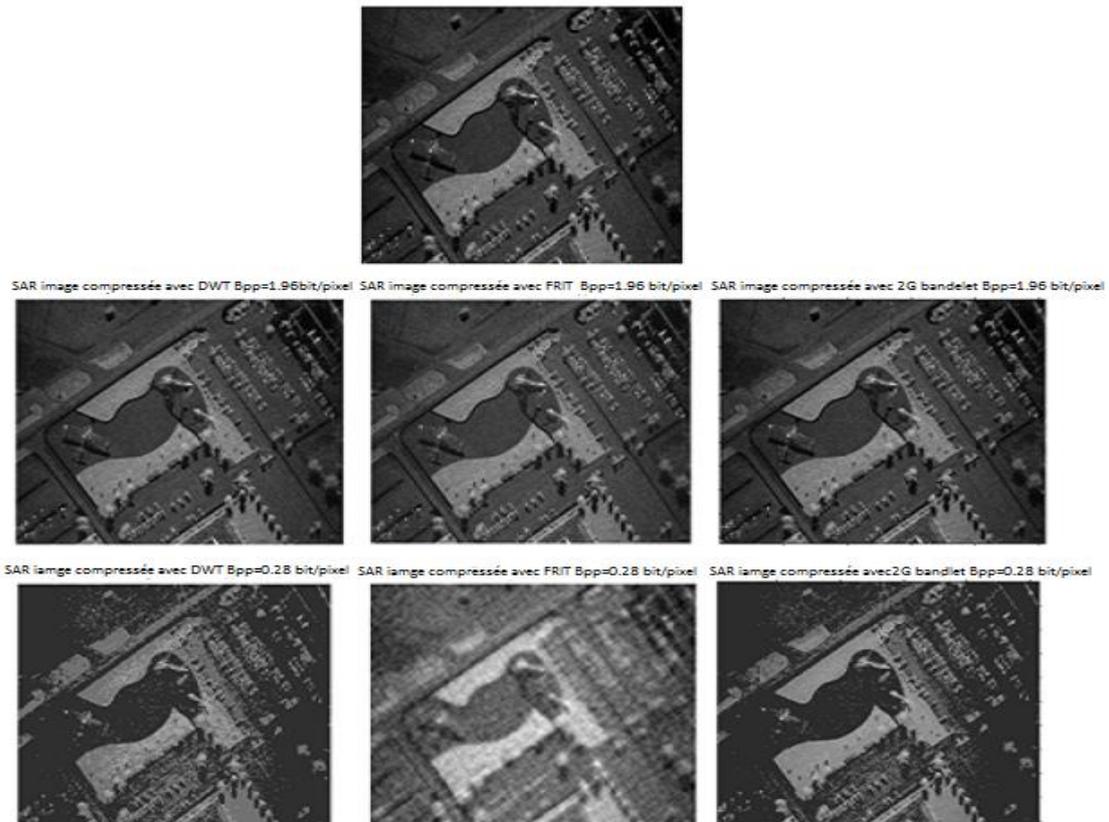


Figure 2.15. Comparaison de la compression de l'image SAR à $\text{bpp}=1.96$ bit /pixel (DWT PSNR =35.86 dB, FRIT PSNR = 34.71 dB et 2GBandelet PSNR = 37.18dB) et à $\text{bpp}=0.28$ bit/pixel (DWT PSNR =19.55 dB, FRIT PSNR = 18.85 dB et 2G Bandelet PSNR = 24.10dB)

La figure 2.16 illustre le problème d'artefact obtenu pour les différentes transformées utilisées pour la compression. Les transformations géométriques multi-échelles montrent une bonne qualité visuelle avec moins de faux artefacts et de structures dégradées.

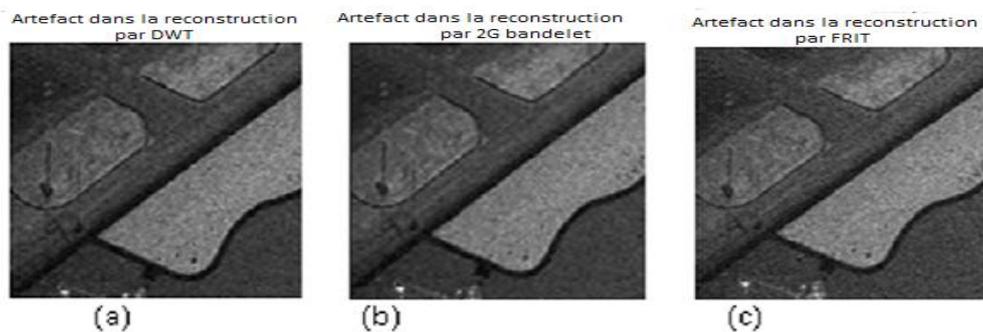


Figure 2.16. Problème d'artefact, (a) : dans DWT, (b) dans bandelettes 2G et (c) dans FRIT

Les résultats quantitatifs sont illustrés par le tableau 2.1, qui présente les valeurs du PSNR (dB) des images compressées avec différents débits binaires.

| bpp (bit/pixel) | 2.52 | 1.47 | 0.83 | 0.28 | 0.17 | 0.08 |
|--------------------|-------|-------|-------|-------|-------|-------|
| PSNR dB (DWT) | 37.71 | 32.33 | 29.06 | 17.56 | 15.88 | 14.94 |
| PSNR dB (2G BT) | 38.86 | 33.18 | 30.01 | 7.19 | 15.39 | 14.07 |
| PSNR dB (FRIT) | 37.80 | 30.76 | 25.40 | 22.19 | 21.76 | 21.26 |

Tableau 2.1. Résultats du PSNR(dB) pour la compression des images SAR avec différents bpp (nombre bits/pixel).

Les figures 2.17 et 2.18, illustrent la variation de PSNR avec bpp (bits / pixel) pour plusieurs images SAR.

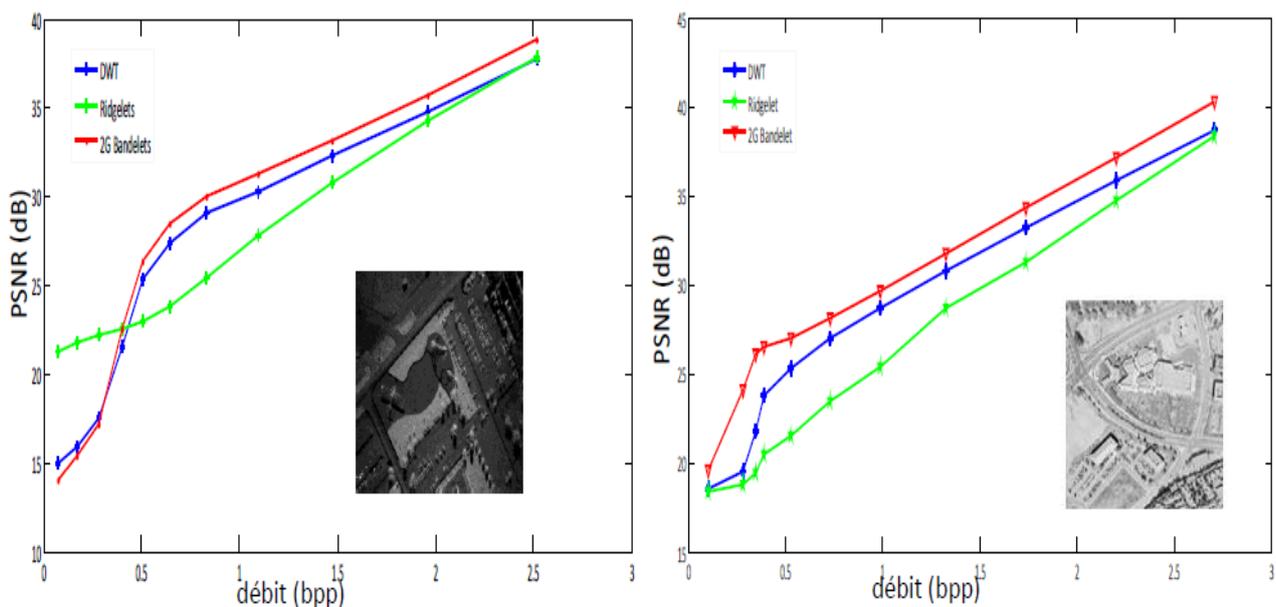


Figure 2. 17. Résultats pour la compression des images SAR utilisant les transformées directionnelles

D'après les résultats représentés dans la figure ci-dessous, nous avons constaté que dans le cas où l'image a des structures directionnelles dominantes; les ondelettes géométriques sont robustes et efficaces; ceci est dû au fait qu'elles peuvent mieux représenter les régularités géométriques. Alors que les ondelettes séparables sont plus adaptées à la représentation des images où les zones homogènes sont dominantes (voir figure 2. 18).

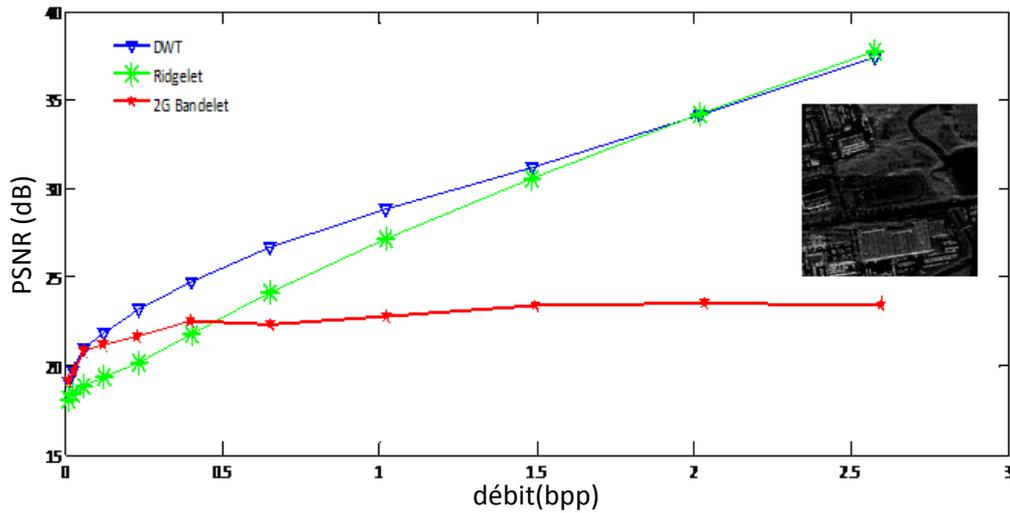


Figure 2.18. Résultats de compression obtenus pour l'image SAR SABR Radar homogène

L'un des principaux problèmes en traitement des images SAR est le chatoiement (speckle) ou bruit cohérent qui peut être modélisé comme un bruit multiplicatif. Un bon filtrage (despeckling) dans les zones homogènes, est fourni par une valeur de coefficient de variation C_v bas, une diminution de la valeur de l'écart type σ , une conservation globale de la moyenne μ , et une augmentation de la valeur de l'ENL (voir tableau 2.2). Nos résultats expérimentaux ont montré que durant la phase de compression, le speckle est éliminé des images SAR, induisant ainsi une amélioration et une efficacité de codage.

Le nombre équivalent de regards (ENL : Equivalent Number of Looks) est utilisé comme critère pour mesurer la suppression du speckle dans les images SAR. Il est calculé à partir du coefficient de variation locale de C_v dans une zone homogène. L'ENL est défini par :

$$ENL = \frac{1}{C_v} = \frac{\mu^2}{\sigma^2} \quad (2.17)$$

où μ et σ sont respectivement la moyenne et l'écart-type estimés dans une zone homogène.

| Bpp(bit/pixel) = 2.10 | | | |
|------------------------------|-----------------|------------|-------------------------|
| | PSNR(dB) | ENL | C_v |
| 2GBT | 37.18 | 20.75 | 0.219 |
| DWT | 35.86 | 20.62 | 0.220 |
| FRIT | 34.71 | 20.75 | 0.219 |
| Bpp (bit/pixel)=1.47 | | | |
| 2GBT | 33.18 | 6.94 | 0.375 |
| DWT | 32.33 | 6.87 | 0.372 |
| FRIT | 30.76 | 6.93 | 0.379 |

Tableau 2.2. Mesure d'évaluation qualitative de la compression d'images SAR et Despeckling

2.5. Conclusion et perspectives

Dans ce chapitre, nous avons présenté la transformée en bandelettes qui exploite avantageusement les structures géométriques des images et leurs orientations. L'inconvénient majeur de cette transformée est la complexité de sa conception car elle nécessite des étapes de détection de géométrie. Néanmoins, des versions récentes ont été proposées, elles sont plus souples et plus modérées en temps de calcul par rapport aux transformées en ondelettes, qui demeurent un outil puissant pour la représentation creuse de l'image. La transformée en bandelettes d'ordre 0 est relativement peu complexe permettant d'améliorer les résultats de compression par rapport à la transformée en ondelettes. Dans la construction de la transformée en bandelette deux étapes sont fondamentales à l'amélioration de la compression: la constitution d'une base de bandelettes optimales pour décrire la meilleure géométrie et l'optimisation du critère d'allocation débit-distorsion (le Lagrangien) utilisé pour choisir la meilleure direction de la géométrie. Plusieurs perspectives sont possibles pour améliorer cette méthode, par exemple l'optimisation du paramètre T qui est généralement choisi par l'utilisateur; la constitution d'un dictionnaire de base de bandelettes de départ pour une représentation parcimonieuse de l'image, etc..

Dans les chapitres suivants, on s'intéressera à la biométrie palmaire et palmaire multispectrale en vue de l'amélioration de l'identification des individus. A cet effet, nous avons utilisé des transformées en ondelettes géométriques non adaptatives pour caractériser les images palmaires. D'abord, nous commencerons par aborder les généralités sur la biométrie palmaire, puis nous présenterons les techniques de fusion et de classification, pour finalement exposer les contributions développées par nos soins en utilisant principalement la transformée de Radon et la transformée en Ridgelets.

Chapitre 3

Introduction à la multibiométrie

3.1. Généralité sur le système biométrie

La biométrie est entrain de devenir la nouvelle référence en matière de vérification de l'identité personnelle d'un individu en utilisant une ou plusieurs caractéristiques physiques ou comportementales dites modalités biométriques [74-78]. Il existe trois catégories de modalités: biologique (exemple: ADN, les signaux physiologiques, etc.), comportementale se basant sur l'analyse du comportement d'un individu (exemple: la démarche, la voix, la dynamique de frappe au clavier, etc.) et morphologique se basant sur les traits physiques qui sont généralement choisis parmi les traits permanents et uniques pour toute personne (empreintes digitales, empreintes palmaires, iris, etc.) (voir figure 3.1). Afin, qu'une modalité biométrique soit utilisée par un système biométrique sécurisé, elle doit satisfaire les conditions suivantes [74-78]:

- Universalité : toutes les personnes à identifier doivent la posséder ;
- Unicité : l'information doit être aussi dissimilaire que possible entre les différentes personnes ;
- Permanence : l'information collectée doit être présente tout au long de la vie de l'individu;
- Collectabilité : l'information doit pouvoir être collectée et mesurée afin de pouvoir être utilisée comme moyen de comparaison;
- Acceptabilité : le système doit respecter certains critères (facilité d'acquisition, rapidité, etc.) pour que les utilisateurs l'acceptent et l'utilisent.

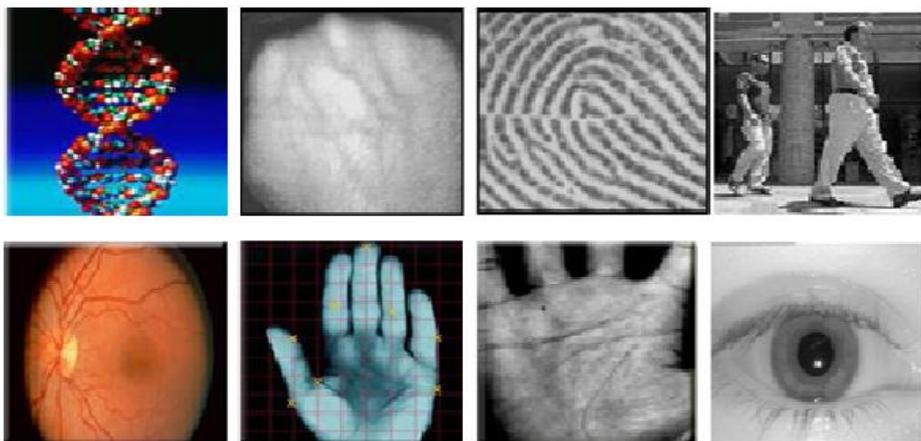


Figure 3.1. Quelques modalités biométriques

Une brève comparaison entre les différentes modalités est illustrée dans le tableau 3.1 ci dessous.

| Modalité | <i>Universalité</i> | <i>Unicité</i> | <i>Permanence</i> | <i>Collectabilité</i> | <i>Acceptabilité</i> |
|-----------------------------------|---------------------|----------------|-------------------|-----------------------|----------------------|
| Modalités biologiques | | | | | |
| ADN | Oui | Oui | Oui | Faible | Faible |
| Groupe sanguin | Oui | Non | Oui | Faible | Non |
| Signal du cerveau (EEG) | Oui | Oui | Oui | Faible | Non |
| Modalités comportementales | | | | | |
| Démarche | Oui | Non | Faible | Oui | Oui |
| Signature dynamique | Oui | Oui | Faible | Oui | Oui |
| Dynamique de frappe | Oui | Oui | Faible | Oui | Oui |
| Voix | Oui | Oui | Faible | Oui | Oui |
| Modalités morphologiques | | | | | |
| Iris | Oui | Oui | Oui | Oui | Un peu |
| Rétine | Oui | Oui | Oui | Oui | Un peu |
| Visage | Oui | Non | Faible | Oui | Oui |
| Géométrie de la main | Oui | Non | Oui | Oui | Oui |
| Veines de la main | Oui | Oui | Oui | Oui | Oui |
| Empreinte palmaire | Oui | Oui | Oui | Oui | Oui |
| Empreinte digitale | Oui | Oui | Oui | Oui | Oui |

Tableau 3.1. Propriétés des diverses modalités biométriques.

3.2. Fonctionnement d'un système biométrique

Un système biométrique est essentiellement un système de reconnaissance de formes qui utilise les informations biométriques d'un individu. Selon le contexte de l'application, un système biométrique peut fonctionner en mode d'enrôlement, en mode de vérification, ou bien en mode d'identification (voir figure 3.2):

3.2.1. L'enrôlement

C'est la première étape d'un système biométrique où une personne est enregistrée pour la première fois. Elle consiste à effectuer les tâches suivantes :

- L'acquisition d'une ou de plusieurs caractéristiques à partir d'un capteur biométrique,
- Le prétraitement en vue de l'extraction des caractéristiques (*features*) discriminantes. Ceci en utilisant un algorithme spécifique à chaque application; dont le but est de générer un modèle biométrique propre à chaque individu, souvent basé sur l'apprentissage automatique. Dans cette phase, la quantité de données à stocker dans la base de données est réduite ce qui facilite le processus de la vérification ou de l'identification.

3.2.2. La vérification

La vérification d'identité consiste à comparer une référence biométrique acquise par le système avec le modèle biométrique correspondant stocké dans la base de donnée. Dans ce cas, le

système opère un test "1 contre 1" et renvoie une décision binaire pouvant être pondérée. Considérant le vecteur d'entrée C_U définissant les caractéristiques biométriques de l'utilisateur U extraites par le système, et M_U son modèle biométrique stocké dans la base de données. Le processus de vérification retourne une valeur booléenne suite au calcul de la fonction f définie par [77-79]:

$$f(C_U, M_U) = \begin{cases} 1 & \text{si } S(C_U, M_U) \geq \tau \\ 0 & \text{ailleurs} \end{cases} \quad (3.1)$$

où $S(C_U, M_U)$: la fonction de mesure de similarité (*matching score*) qui définit la correspondance entre les deux vecteurs biométriques et τ : le seuil de décision à partir duquel les deux vecteurs sont considérés identiques.

3.2.3. Identification

En mode identification, le système biométrique détermine l'identité d'un individu à partir d'une base de donnée d'identités $[I_1, I_2, \dots, I_N]$ préalablement enrôlées dans le système. Ainsi, le système opère le test "Un contre tous" (1:N), pour attribuer une identité à l'individu inconnu. Afin de déterminer l'identité I_k ($k \in [1, 2, \dots, N]$), le processus d'identification peut être formalisé ainsi :

$$f(C_U) = \begin{cases} I_k & \text{si } \max_{1 \leq k \leq N} S(C_U, M_k) \geq \tau \\ I_0 & \text{sin on} \end{cases} \quad (3.2)$$

avec I_0 : l'identité inconnue et M_k : le modèle biométrique correspondant à l'identité I_k .

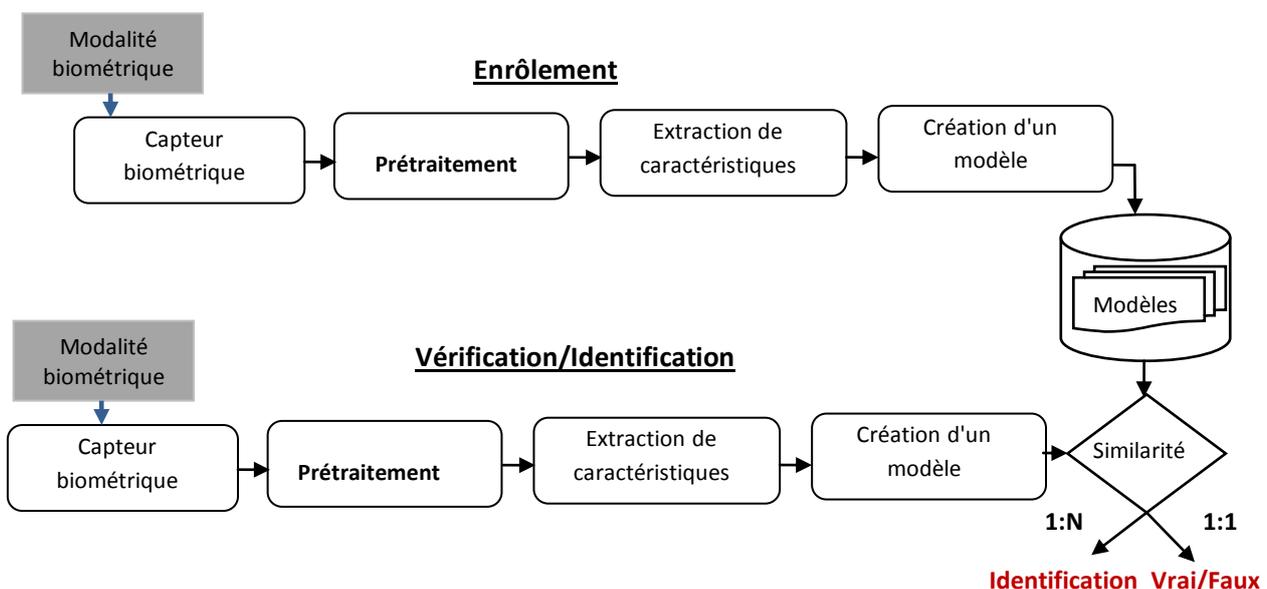


Figure 3.2. Schéma fonctionnel d'un système biométrique

3.3. Performance d'un système biométrique

L'efficacité et la fiabilité d'un système biométrique dans un contexte donné, sont mesurées par un ensemble de métriques de diverses natures. Issus de la littérature, on peut citer comme exemples : les mesures de taux d'erreur, les mesures liées au temps de traitement, les courbes de performances, ainsi que les points de fonctionnement spécifiques. Selon l'organisation Internationale de Normalisation ISO/IEC 19795-1 [82], les mesures des taux d'erreurs sont divisées en trois classes : les taux d'erreur fondamentales, les taux d'erreur des systèmes d'authentification et les taux d'erreur des systèmes d'identification. En plus des taux d'erreur fondamentale, tels que : le taux d'échec à l'acquisition (*Failure To Acquire rate* : FTA) et le taux d'échec à l'enrôlement (*Failure To Enroll rate* : FTE). Le système de vérification biométrique est soumis à deux types d'erreurs [74-83]:

- Le taux de faux rejets (*False Reject Rate* : FRR) qui correspond à la proportion des demandeurs légitimes rejetés par erreur.
- Le taux de fausse acceptation (*False Accept Rate* : FAR) qui correspond à la proportion des imposteurs acceptés par erreur.

Mathématiquement, les erreurs dans le système de vérification peuvent être formulées comme suit [74-78] :

Soit le modèle biométrique M_U stocké de l'utilisateur et l'entrée C_U acquise par le système.

les deux hypothèses liées au processus de vérification sont :

- H_0 : l'entrée C_U ne provient pas de la même personne que le modèle.
- H_1 : L'entrée provient de la même personne que le modèle.

Les deux décisions associées au système de vérification sont:

- D_0 : la personne n'est pas ce qu'elle prétend être.
- D_1 : la personne est ce qu'elle prétend être.

La règle de décision du système est : " si le score de correspondance $S(C_U, M_U)$ est inférieur au seuil τ du système, alors décidé D_0 sinon décidé D_1 ". Ainsi, les taux d'erreurs FAR et FRR peuvent s'écrire comme suit (voir figure 3.3):

$$\begin{aligned} FAR &= P(D_1 / H_0) = \int_{\tau}^{\infty} P(S(C_U, M_U) / H_0) dS \\ FRR &= P(D_0 / H_1) = \int_{-\infty}^{\tau} P(S(C_U, M_U) / H_1) dS \end{aligned} \quad (3.3)$$

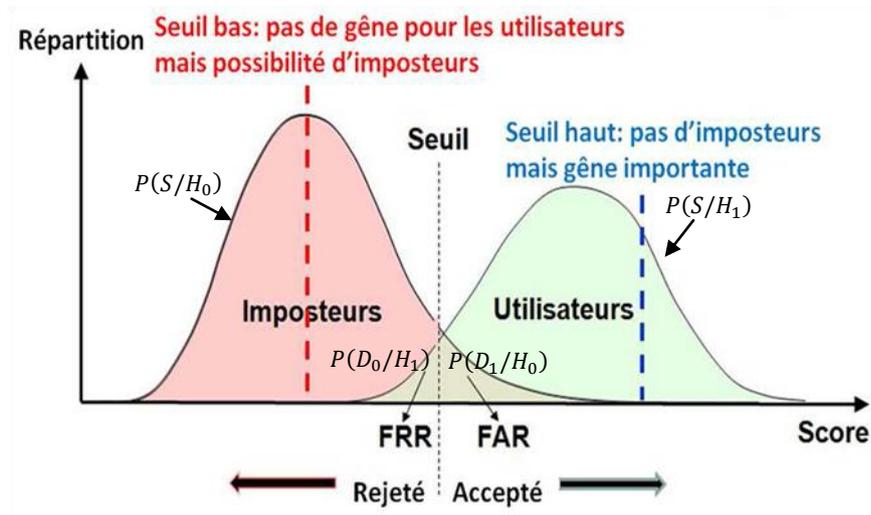


Figure 3.3. Distributions des taux de vraisemblance des utilisateurs légitimes et des imposteurs d'un système biométrique [79]

En plus des métriques, on a souvent recouru à des courbes pour la représentation des performances globales des systèmes biométriques. La plus couramment utilisée est la courbe ROC (*Receiver operating characteristic curve*) détaillée au chapitre 5.

3.4. La Multibiométrie

Le système biométrique optimal et robuste doit impérativement avoir les propriétés d'universalité, d'acceptabilité, d'unicité, de collectabilité et de sécurité. Mais comme il n'existe pas de modalités qui puissent avoir toutes ces propriétés, la solution est d'utiliser plusieurs caractéristiques biométriques dans un unique système pour améliorer ses performances. Grâce à des solutions de sécurité fiables et efficaces dans les applications critiques de sécurité, les systèmes biométriques multimodaux ont récemment émergé dans la communauté scientifique, ils sont devenus une alternative robuste aux systèmes traditionnels unimodaux. comme exemple, nous citons la reconnaissance basée sur les empreintes palmaires associée à la reconnaissance basée sur les veines palmaires. La multibiométrie combine les informations issues soit de plusieurs capteurs biométriques, soit d'algorithmes ou de décisions pour améliorer les performances des applications biométriques, dans l'optique de permettre une large couverture de la population, d'accroître la dissuasion contre l'usurpation d'identité, et finalement de permettre une meilleure tolérance aux pannes [75,79].

Les systèmes biométriques basés sur la nature des sources d'informations biométriques (voir figure 3.4), peuvent être classés dans l'une des cinq catégories suivantes [74-81,83 -85]:

- 1- Système multi-capteurs (*Multi-sensors*) : plusieurs capteurs sont utilisés pour acquérir une seule caractéristique biométrique d'un individu afin d'extraire diverses informations provenant de l'enregistrement des images.
- 2- Système multi-algorithmes (*Multi-algorithms*) : dans ces systèmes, plusieurs algorithmes d'extractions de caractéristiques et/ou plusieurs algorithmes de calcul de similarité sont utilisés sur la même modalité biométrique. Par exemple, des algorithmes d'analyse de texture et de minuties peuvent être associés pour traiter la même image d'empreinte digitale afin d'extraire diverses caractéristiques qui peuvent améliorer la performance du système. Ainsi, ce genre de système ne nécessite pas de capteurs supplémentaires et n'oblige pas l'utilisateur à interagir avec de multiples capteurs, d'où l'amélioration de la commodité d'utilisation.
- 3- Système multi-instances (*Multi-instances*) : dans ces systèmes, plusieurs instances de la même caractéristique biométrique sont utilisées pour l'authentification. Par exemple, l'image de la main gauche et celle de droite d'un sujet peuvent être utilisées. Ces systèmes ne nécessitent généralement pas l'introduction de nouveaux capteurs, ni le développement de nouveaux algorithmes d'extraction de caractéristiques ou de reconnaissance; ce qui les rend par conséquent rentables.
- 4- Système multi-échantillons (*Multi-samples*) : un seul capteur peut être utilisé pour acquérir plusieurs échantillons du même trait biométrique dans le but de prendre en compte les variations qui peuvent se produire au sein de ce trait, ou pour obtenir une représentation plus complète du caractère sous-jacent. Par exemple, un seul capteur peut être utilisé pour un système de reconnaissance faciale pour capturer différentes images d'expression du visage d'un sujet.
- 5- Systèmes multimodaux (*Multimodal system*) : dans ce cas, plusieurs modalités biométriques sont combinées afin d'établir l'identité d'un individu. Le nombre de traits biométriques utilisés dans une application spécifique est également limité par des considérations pratiques, telles que le coût de déploiement, le temps d'enrôlement, le temps de retour ("*throughput time*") ou encore le taux d'erreur attendu.

Finalement, il est à noter que l'on utilise souvent le terme de *systèmes hybrides* pour se référer aux systèmes qui intègrent un sous-ensemble des cinq scénarios que nous venons d'évoquer.

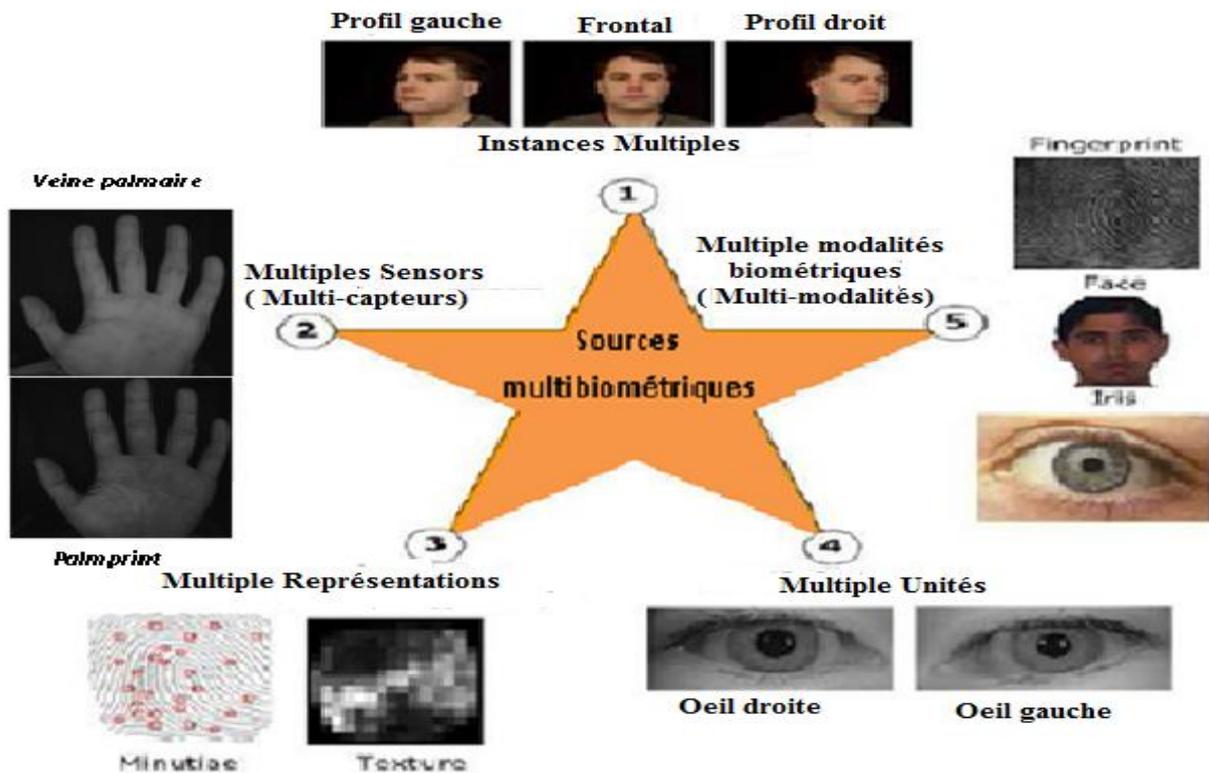


Figure 3.4. Les différentes sources dans les systèmes multibiométriques.

Les systèmes multibiométriques offrent plusieurs avantages par rapport aux systèmes classiques. Ils fournissent une amélioration substantielle en terme de précision, aussi ils abordent efficacement la question de non universalité du fait qu'ils combinent plusieurs traits biométriques, et présentent l'avantage de diminuer considérablement les risques d'usurpation d'identité.

3.5. Niveau de fusion d'information dans le système multibiométrique

Au cours du processus d'un système biométrique la quantité d'information disponible devient compressée au fur à mesure qu'on progresse dans ses étapes, allant de l'acquisition à la décision. Relativement au type d'information dans chaque module, plusieurs niveaux de fusion peuvent être définis. En effet, deux grandes familles de niveaux de fusion ont été définies [75-79, 84-87] à savoir: la fusion pré-classification ou avant l'appariement (*matching*) et la post-classification ou après appariement. Dans ce qui suit, nous détaillerons ces deux familles de fusion.

3.5.1. Fusion pré-classification

Avant l'appariement, l'intégration des informations peut avoir lieu soit au niveau capteur (*Sensor level fusion*), soit aux niveaux caractéristiques (*Feature Level fusion*).

3.5.1.1. Fusion niveau capteur

Le premier niveau de fusion se situe au niveau capteur dans le cadre des systèmes multi-capteurs ou multi-échantillons. L'objectif est de créer une nouvelle donnée à traiter en combinant les données brutes issues de plusieurs capteurs compatibles entre eux, de la même modalité biométrique ou de plusieurs instances de cette modalité à partir d'un seul capteur. Il s'agit d'une technique appelée aussi fusion bas-niveau ou fusion niveau pixel. la figure 3.5, schématise le processus de fusion de ce type de niveau.

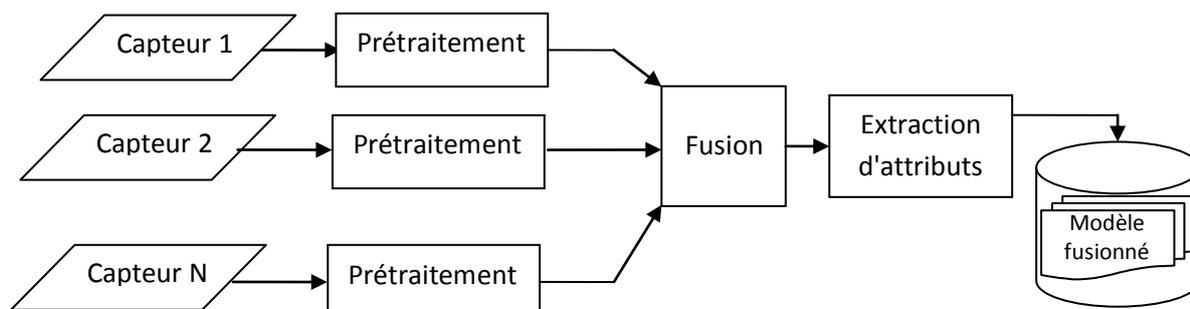


Figure 3.5. Schéma de fusion niveau capteur

3.5.1.2. Fusion niveau caractéristiques

La fusion au niveau caractéristiques consiste à combiner différents vecteurs de caractéristiques (*feature vectors*) qui sont obtenus à partir de l'une des sources suivantes : plusieurs capteurs du même trait biométrique, plusieurs instances du même trait biométrique, plusieurs unités du même trait biométrique ou encore plusieurs traits biométriques. Lorsque les vecteurs de caractéristiques sont homogènes (par exemple, plusieurs images de l'empreinte palmaire d'un utilisateur), un unique vecteur de caractéristiques résultant peut être calculé comme une somme pondérée des vecteurs de caractéristiques individuels. Lorsque les vecteurs de caractéristiques sont hétérogènes (par exemple, des vecteurs de caractéristiques de différentes modalités biométriques comme les veines palmaires et les empreintes palmaires), ils peuvent être concaténés pour former un seul vecteur de caractéristiques. La figure 3.6 illustre le cas de la fusion des caractéristiques par concaténation. La concaténation de deux vecteurs de caractéristiques peut engendrer un vecteur de caractéristiques ayant une grande dimension, menant au fameux problème de la "*malédiction de la dimensionalité*" [75-79] ce qui engendre des difficultés au niveau de l'apprentissage. Il est souvent intéressant de réduire la taille du vecteur obtenu après fusion, soit par des méthodes de réduction basées sur l'analyse statistique (telle que l'ACP et la LDA (voir chapitre 4)) ou par la sélection des attributs les plus pertinents avant la fusion.

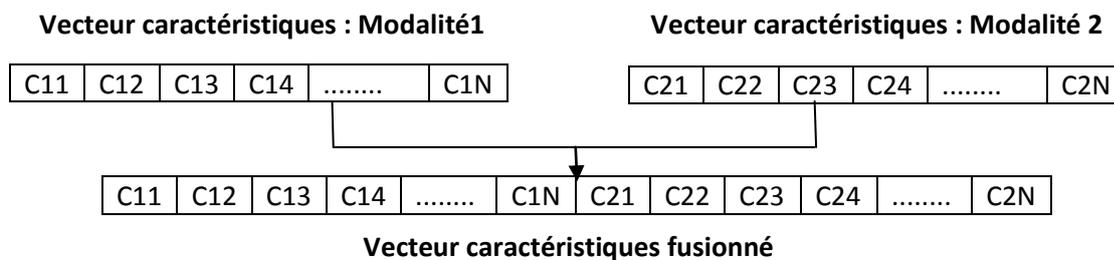


Figure 3.6. Concaténation des vecteurs caractéristiques

3.5.2. Fusion post-classification

Les schémas d'intégration de l'information après l'étape de la classification ou de correspondance peuvent être divisés en quatre catégories : sélection dynamique de classificateurs, fusion au niveau décision, fusion au niveau rang et enfin la fusion au niveau score.

3.5.2.1. Fusion niveau score

Les différents sous-systèmes du système multibiométrique produisent des scores après l'étape de la comparaison. Après les vecteurs de caractéristiques, les scores d'appariement contiennent l'information la plus riche à propos du modèle biométrique. Par conséquent, la fusion au niveau score donne le meilleur compromis entre la richesse d'information et la facilité d'implémentation. Il existe deux approches pour fusionner les différents scores obtenus (voir figure 3.7), l'une basée sur la combinaison des scores pour générer un nouveau score et l'autre basée sur la classification pour générer une classe finale à partir de ces scores.

3.5.2.1.1. Approche basée sur la combinaison des scores

Dans cette approche, les scores de correspondance individuels sont combinés de manière à former un unique score qui est ensuite utilisé pour rendre la décision finale. Deux étapes sont nécessaires pour établir la fusion des scores :

- La normalisation des scores
- La combinaison des scores normalisés

La normalisation permet de rendre les distributions des scores des différents sous-systèmes d'authentification compatibles, c'est à dire de modifier les paramètres des distributions de scores de correspondance (moyenne et écart-type) de manière à ce qu'ils soient dans un même intervalle. Pour avoir un bon schéma de normalisation, les estimateurs de paramètres de distribution de scores de correspondance doivent être robustes (insensible au données

aberrantes) et efficaces, l'une des méthodes de normalisation les plus simples est la normalisation Min-Max. Elle est la plus adaptée lorsque les valeurs des bornes (minimum et maximum) de la distribution des scores sont connues. Dans ce cas, on peut facilement traduire les scores minimums et maximums respectivement vers 0 et 1. Soit s_{ij} le $j^{\text{ème}}$ score de correspondance de la $i^{\text{ème}}$ modalité, où $i = 1, 2, \dots, R$ et $j = 1, 2, \dots, M$ (R : le nombre de modalités et M : nombre de scores disponibles dans l'ensemble d'entraînement). Le score normalisé s_{ik}^N par Min-Max du score test s_{ik} est défini par :

$$s_{ik}^N = \frac{s_{ik} - \min(\{s_{i.}\})}{\max(\{s_{i.}\}) - \min(\{s_{i.}\})} \quad (3.4)$$

avec $\{s_{i.}\} = \{s_{i1}, s_{i2}, \dots, s_{iM}\}$. Cette méthode n'est pas robuste et très sensible aux données aberrantes. D'autres méthodes plus au moins robustes et efficaces existent, nous en citons la méthode z-score [76,86] , la méthode utilisant la médiane et l'écart absolu à la médiane (MAD), la technique de normalisation basée sur les estimateurs de Hampel (estimateurs tanh) [76,86], etc.

L'ensemble des règles de combinaison des scores a été développé par Kittler et al [87] (voir figure 3.7). Le but est d'obtenir un score de classification unique en utilisant plusieurs schémas de combinaison tels que la règle (*sum rule*), la règle du produit (*product rule*), la règle du maximum (*max rule*), la règle du minimum (*min rule*) et la règle de la médiane (*median rule*). Ils considèrent le problème de classer un modèle d'entrée X dans l'une des m classes (w_1, w_2, \dots, w_m) possibles (dans un système de vérification, $m = 2$) d'après les scores fournis par les R différents classificateurs. Soit \vec{x}_i un vecteur de caractéristiques (provenant du modèle d'entrée X) présenté au $i^{\text{ème}}$ sous-système. Considérons la probabilité de chaque sous-système (score) $P(w_j / \vec{x}_i)$, c'est-à-dire, la probabilité a posteriori de la classe w_j sachant \vec{x}_i . Soit c le score final (score de classification) obtenu par combinaison des différents scores des sous-systèmes. Les règles suivantes peuvent alors être utilisées pour estimer c [76,86,87]:

- **La Règle du produit (*Product Rule*)** : Cette règle est basée sur l'hypothèse d'indépendance statistique des représentations ($\vec{x}_1, \vec{x}_2, \dots, \vec{x}_R$). Dans ce cas, le score final est donné par :

$$c = \arg \max_j \prod_{i=1}^R P(w_j / \vec{x}_i) = \prod_{i=1}^R s_i \quad (3.5)$$

- **La Règle de la somme (*Sum Rule*)** : La règle de la somme est plus efficace que la règle du produit lorsqu'il y a un fort niveau de bruit menant à une certaine ambiguïté dans le problème de classification. La règle de la somme est définie comme suit :

$$c = \arg \max_j \sum_{i=1}^R P(w_j / \vec{x}_i) = \sum_{i=1}^R s_i \quad (3.6)$$

- **La Règle du maximum (Max Rule)** : La règle du maximum approxime la moyenne des probabilités a posteriori par la valeur maximum. Elle est définie par :

$$c = \arg \max_j \left(\max_i P(w_j / \vec{x}_i) \right) = \max_i (s_i) \quad (3.7)$$

- **La Règle du minimum (Min Rule)** : La règle du minimum est obtenue en bornant le produit des probabilités a posteriori. Elle est formulée par :

$$c = \arg \max_j \left(\min_i P(w_j / \vec{x}_i) \right) = \min_i (s_i) \quad (3.8)$$

3.5.2.1.2. Approche basée sur la classification des scores

Dans cette approche, plusieurs classificateurs sont utilisés pour combiner les scores issus des différents modules de reconnaissance biométrique, afin d'arriver à une décision finale. Les règles de combinaison précédentes permettent de calculer le score final à l'aide des scores des différents sous systèmes (voir figure 2.7).

3.5.2.2. Fusion niveau décision

La fusion de décision (Decision Level Fusion) intervient dans une problématique d'authentification ou de vérification d'identité. L'intégration d'information à ce niveau est totalement abstraite, dans le sens où seules les décisions finales (accepté ou rejeté) de chaque sous-système sont considérées. Pour obtenir la décision finale, il existe plusieurs méthodes qui permettent de fusionner ces décisions, seules les méthodes les plus utilisées sont présentées et nous en citons :

- La règle "**ET**" logique: introduite dans [83], elle consiste à accepter l'utilisateur si et seulement si tous les sous-systèmes l'ont reconnu.
- La règle "**OU**" logique: introduite aussi dans [83], elle consiste à accepter l'utilisateur si au moins un seul sous-système l'a reconnu.
- La méthode des votes majoritaires (*majority voting*) : elle a été introduite dans [88]. Pour chaque échantillon d'entrée chaque expert ou classificateur produit une décision unique quant à l'identité de l'échantillon. Dans le cas où la décision contient plusieurs choix, le premier choix sera sélectionné. En combinant les décisions des experts, l'échantillon est assigné à la classe pour laquelle il existe un consensus (vote majoritaire).

D'autres méthodes plus complexes utilisent à priori des informations sur les performances des différents sous-systèmes biométriques et comportent une phase d'apprentissage. On peut citer la méthode basée sur la théorie Bayes [89,90]; la théorie de l'évidence de Dempster-Shafer [90] ou la méthode BKS (*Behavior Knowledge Space*) [88].

3.5.2.3. Fusion niveau rang

La fusion de rang (Rank Level Fusion) concerne le processus d'identification d'un individu parmi tous les individus autorisés. Chaque sous-système fournit au système global d'identification une liste de classements qui sont triés par ordre décroissant de confiance ou de probabilité. Plusieurs méthodes peuvent être utilisées pour combiner les différents rangs [89,91]. On cite la méthode de plus haut rang (*highest rank method*), la méthode "*Borda count*" qui repose sur un processus de vote pondéré à partir de la somme des rangs donnés par chaque sous-système, et la méthode de régression logistique qui est une généralisation de la méthode "*Borda count*" où une somme pondérée des rangs individuels est calculée et les poids sont déterminés par régression logistique.

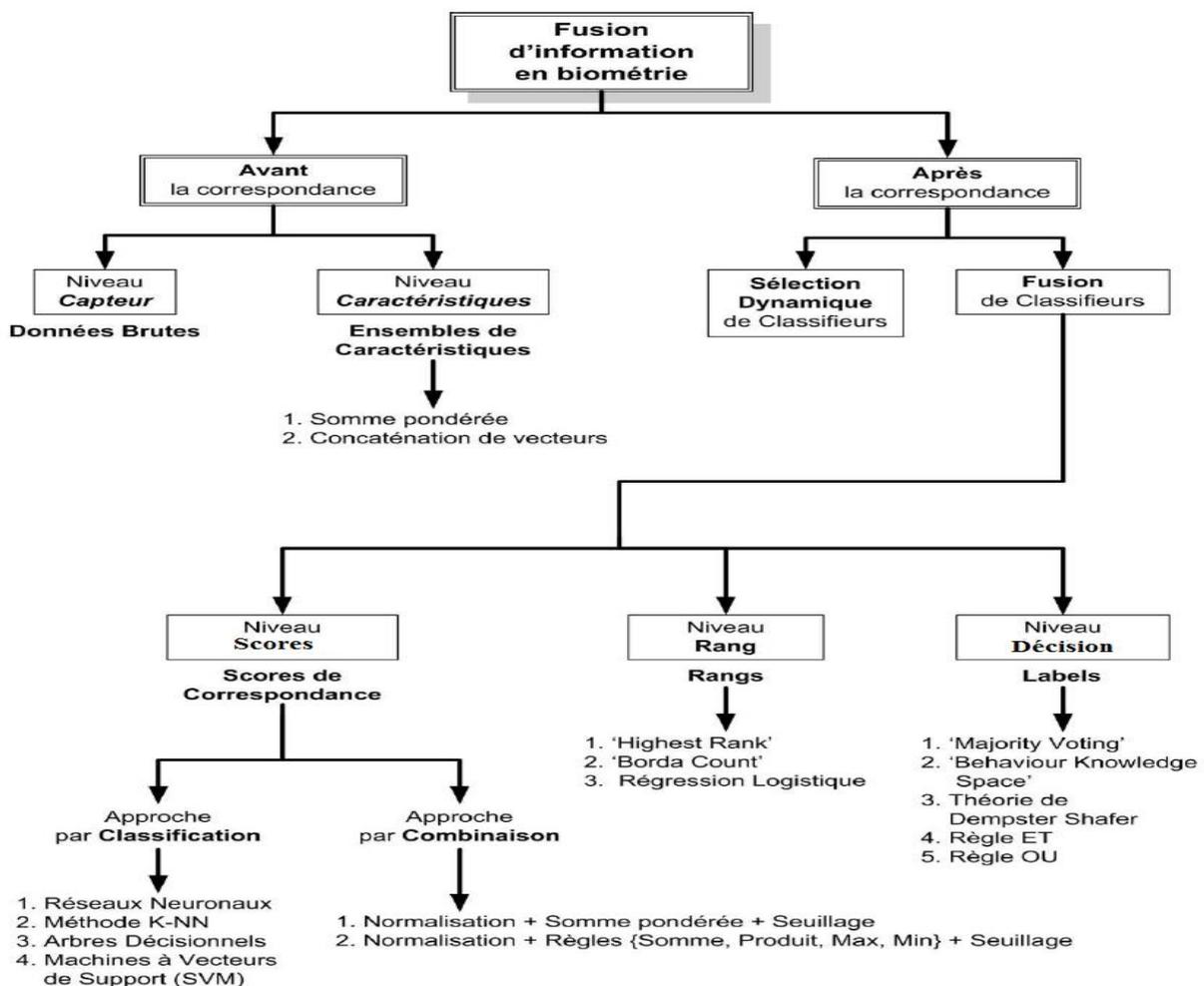


Figure 3.7. Les différents niveaux de fusion dans un système multibiométrique [76]

De nombreux systèmes biométriques multimodaux avec diverses méthodes et stratégies ont été proposés au cours de la dernière décennie pour atteindre un taux de précision plus élevé, et d'augmenter la robustesse contre les attaques frauduleuses et des conditions extérieures. Le tableau suivant résume quelques systèmes.

| Année | Modalités utilisées dans la fusion | Les auteurs | Fusion | Technique de Fusion |
|-------|--|-------------------------|----------|--|
| 1998 | Visage et empreinte digitale | Hong et Jain | Score | Règle Produit |
| 2000 | Visage, voix et mouvement | Frischholz et Dieckmann | Décision | Le vote majoritaire, la règle de la somme pondérée |
| 2003 | Visage, fingerprint et géométrie de la main | Ross et Jain | Score | Règle somme, l'arbre de décision |
| 2004 | Visage palmprint | Feng et al | Feature | Concaténation |
| 2005 | Visage, fingerprint et géométrie de la main | Jain et al | | Somme, max et min règle |
| 2009 | Hand biometrics (fingerprint, palmprint et géométrie des doights) | Yu et al | Décision | Règle ET, Règle OU et le vote majoritaire |
| 2010 | Deux palmprint images | Kumar et Shekhar | Rang | Borda count, maximum rang, BKS |
| 2011 | Palmprint et palmveine | Zhang et al | Score | Somme pondérée dynamique |

Tableau 3.2. Quelques exemples de systèmes biométriques multimodaux avec leur technique de fusion

3.6. Conclusion

Dans ce chapitre, nous avons abordé le principe des systèmes multibiométriques et les techniques de fusion. Il existe différents types d'architecture ainsi que différents niveaux de fusion le long de la progression dans le système biométrique, partant du capteur biométrique jusqu'à la décision finale d'accepter ou de rejeter un individu. La multibiométrie permet d'améliorer efficacement les performances et la robustesse du processus d'authentification en combinant différentes informations. La recherche dans le domaine de la biométrie multimodale est relativement récente. Depuis 1995, date à laquelle les premiers travaux de fusion du visage et de la voix ont été menés, de nombreuses études ont été conduites en associant différentes modalités, en faisant varier le niveau de fusion des données et en testant plusieurs règles de fusion. Notre système multimodal développé dans le cadre de cette thèse fait intervenir les modalités empreintes palmaires et les veines palmaires et deux niveaux de fusion, le niveau pixel en utilisant les transformées en ondelettes, et le niveau caractéristiques par concaténation.

Chapitre 4

Techniques de classification pour la biométrie

4.1. Introduction

L'apprentissage automatique est une technique qui permet à des machines d'apprendre à partir d'exemples ou de l'expérience. Dès lors, qu'il n'existe pas de modèles mathématiques pour résoudre certains problèmes tels que la reconnaissance des formes, alors une approche naturelle s'impose pour palier aux limitations de la programmation classique. Ce concept regroupe une famille d'algorithmes dont la fonctionnalité est de prédire la classe d'un objet selon les variables qui le caractérisent, en partant d'un ensemble d'objets considérés comme exemples d'entraînement décrits par les mêmes variables, et qui constituent la base d'apprentissage. L'apprentissage est dit supervisé ou classification, si les classes d'appartenance des échantillons sont connues d'emblée, sinon il est qualifié de non supervisé ou « *clustering* ». Le processus d'authentification en biométrie est un exemple d'apprentissage supervisé, dans la mesure où on dispose d'un ensemble de caractéristiques décrivant des personnes acceptées par le système, donc appartenant à une classe libellée "authentique", et d'autres personnes rejetées par ce système, appartenant à la classe "imposteur". Le but du processus est de prédire la classe de nouvelles personnes en soumettant leurs caractéristiques respectives au système. L'apprentissage non supervisé, aussi appelé segmentation, ou "*clustering*", consiste à regrouper des objets donnés en groupes communément appelé "*clusters*", en considérant des critères de similitude dont nous citons la distance euclidienne comme exemple. L'apprentissage à vu le jour avec les travaux de Rosenblatt dans les années 60, qui a proposé le premier modèle d'apprentissage concrétisant les recherches sur le neurone formel, pour résoudre un problème de reconnaissance de formes. L'étape suivante fût la mise au point au cours des années 80 de l'algorithme de back-propagation pour le calcul simultané des poids des différentes couches de neurones, donnant ainsi naissance au réseau de neurones. Les années quatre vingt dix ont vu la naissance des SVMs (pour Support Vector Machine), basée sur la théorie SRM (pour Structural Risk Minimisation) introduite par Vladimir Vapnik [92]. Le principe de la classification pour certaines techniques telles que les réseaux de neurones, consiste à ajuster les paramètres de l'algorithme dit classificateur, afin de minimiser l'erreur empirique qui représente l'écart entre les sorties réelles et les sorties désirées.

Effectivement, l'erreur de généralisation des classificateurs ne pouvait être calculée du moment qu'on ne disposait pas des probabilités de distributions des objets. Cependant dans ces travaux sur la théorie statistique de l'apprentissage, Vladimir Vapnik a borné l'erreur de généralisation, et a proposé la théorie SRM qui suggère que le meilleur séparateur linéaire est celui qui résout le compromis entre la minimisation de l'erreur empirique et la maximisation de la marge séparant deux classes d'objets.

L'objectif de ce chapitre est de présenter quelques techniques de classification parmi les plus utilisées en biométrie, la liste de ces techniques n'est pas exhaustive car elle ne constitue pas en soit l'objet de notre étude. Par ailleurs, cela a comme priorité d'abord d'exposer avec plus de détails aux lecteurs les notions utilisées dans le chapitre 5 dans lequel nous présenterons une partie de nos contributions. Plus précisément, le principe et l'algorithme de base de chaque méthode seront présentés et discutés d'une manière synthétique.

Nous nous proposons d'aborder les méthodes de classifications dites paramétriques, en l'occurrence la méthode bayésienne naïve, la discrimination linéaire de Fisher appelée communément LDA (*Linear Discriminant Analysis ou Fisher Discriminant Analysis*). Par ailleurs, nous présenterons quelques méthodes non paramétriques dont nous citerons : les *KNN* pour (*K Nearest Neighbors*) ou les *k* plus proches voisins, les réseaux de neurones dits aussi ANN pour (*Artificial Neural Network*), et finalement les machines à supports vecteurs (SVMs). Toutefois, nous jugeons nécessaire de présenter en premier lieu les techniques d'évaluation des performances des classificateurs et leurs aptitudes à généraliser, communes à toutes ces méthodes.

4.2. Aptitude de généralisation et mesures de performance

Une question majeure inhérente à l'apprentissage concerne l'aptitude des classificateurs à généraliser. En effet, l'entraînement étant effectué sur un ensemble restreint de l'ordre de quelques objets exemples, alors comment évaluer dans ce cas le pouvoir de généralisation des classificateurs ?.

Il existe une technique dite validation croisée, à laquelle on a souvent recours pour mesurer l'aptitude d'un classificateur à généraliser. Celle-ci consiste à diviser la base qui inclut à la fois les objets d'apprentissage et de test en *k* sous-ensembles, ou plus précisément répertoires dits aussi ("*folders*" en anglais) de tailles plus ou moins égales. Le processus d'apprentissage est alors réitéré en utilisant à chaque fois un répertoire comme base de test, et les *k* - 1 répertoires

restant comme base d'entraînement. Finalement, la performance du classificateur est considérée comme la moyenne de toutes les classifications impliquant toutes les combinaisons de répertoires possibles. Lors de la classification, un objet de test est dit bien classé, si sa classe d'origine correspond à la classe prédite par le classificateur ; sinon, il est considéré comme mal classé. La performance de classification ne peut être évaluée seulement par le pourcentage d'objets de test bien classés, qui n'est autre que le rapport entre le nombre des objets bien classés et le nombre total des objets de test. Considérons un exemple de classification binaire dont le but est de discriminer deux classes de personnes. Une classe libellée « Authentiques », désignant des personnes acceptées par un système biométrique donné, et une autre libellée « Imposteurs », désignant les personnes rejetées par ce système. Admettons que la base de test englobe 1030 personnes, dont 30 de la classe « Imposteurs ». Un résultat de classification correspondant à 1/30 imposteurs bien classés, et 1000/1000 personnes authentiques bien classés, reflète un taux de 97,44%. Ce taux paraît intéressant mais en réalité ce n'est pas le cas, dans la mesure où le taux réel des imposteurs bien classés est de 3,33%, ce qui représente un résultat de classification médiocre. A cet effet, il faut considérer en plus du taux de classification des deux classes confondues, d'autres mesures qui renseignent sur chacune d'elles.

Les mesures de performances empiriques les plus communément utilisées dans la classification sont : la précision, la sensibilité, et la spécificité. Si nous considérons le cas de la classification binaire d'une classe positive libellée P et d'une classe négative N , alors ces mesures sont exprimées par les équations suivantes :

$$\text{Précision} = \frac{VP + VN}{VP + FP + FN + VN} \quad (4.1)$$

$$\text{Sensibilité} = \frac{VP}{VP + FN} \quad (4.2)$$

$$\text{Spécificité} = \frac{VN}{FP + VN} \quad (4.3)$$

où :

- VP : désigne le nombre de vrais positifs. Ce sont des objets de l'ensemble de test dont la classe prédite correspond à la classe d'origine P .
- VN : désigne le nombre de vrais négatifs. Ce sont les objets de l'ensemble de test dont la classe prédite correspond à la classe d'origine N .

- FN : désigne le nombre de faux négatifs. Ce sont des objets de la classe P prédits dans la classe N.
- FP : désigne le nombre de faux positifs. Ce sont les objets de la classe N, prédits dans la classe P.

En pratique, on utilise une matrice dite de contingence à partir de laquelle (voir Tableau 4.1) on extrait les différentes mesures de performances empiriques. La précision renseigne sur le nombre de classifications correctes pour les deux classes positive et négative, cependant la sensibilité représente le nombre de classifications correctes relatives à la classe positive, et la spécificité correspond au nombre de classifications correctes de la classe négative. Dans le cas multi-classes, la classe d'origine des objets est considérée comme la classe positive, et les autres classes réunies comme classe négative.

| | | Classes prédites par le classificateur | |
|------------------|-----------------|--|-----------------|
| | | Classe positive | Classe négative |
| Classe d'origine | Classe positive | Vrais positifs | Faux négatifs |
| | Classe négative | Faux positifs | Vrais négatifs |

Tableau 4.1. Matrice de contingence d'une classification binaire.

D'autres statistiques utilisées en biométrie sont dérivées des mesures vues précédemment à savoir: le taux de fausses acceptations (FAR , *False Acceptance Rate*) et le taux de faux rejets (FRR , *False Rejection Rate*). Formellement, les deux statistiques sont exprimées par :

$$FAR = \frac{FP}{\text{nombre des imposteurs}} \quad (4.4)$$

$$FRR = 1 - \frac{FN}{\text{nombre des utilisateurs authentiques}} \quad (4.5)$$

En plus des métriques, on a souvent recourt à des courbes pour la représentation des performances globales des systèmes biométriques. La courbe ROC (*Receiver Operating Characteristic curve*) est la plus utilisée dans la littérature, elle reflète la relation entre le taux de fausses acceptations (FAR) et le taux de faux rejets (FRR). L'avantage de cette représentation graphique est qu'elle permet une représentation compacte de la performance globale d'un

système biométrique pour ses différents paramètres (voir figure 4.1). L'intersection de la courbe ROC et la droite d'équation $FAR = FRR$ fournit une indication appelée taux d'égale erreur (EER, *Equal Error Rate*), cette mesure n'a pas d'utilisation pratique mais elle renseigne sur la performance. En effet, plus le EER est faible, plus le système est performant.

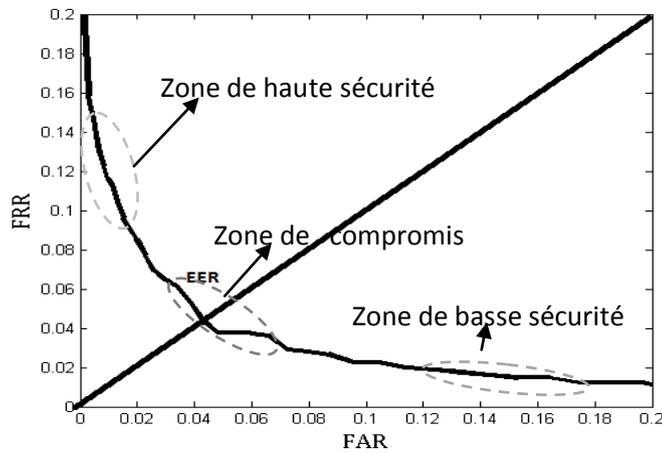


Figure 4.1. Représentation d'une courbe ROC

4.3. Approches paramétriques

4.3.1. Classification naïve Bayésienne

La classification bayésienne est une méthode probabiliste qui permet de définir une règle de décision optimale basée sur le théorème de Bayes, afin de déterminer les probabilités d'associer certaines classe c_i ; à certains objets selon des données d'apprentissage [93]. Donc, l'idée principale de la classification bayésienne d'objets représentés par un ensemble de caractéristiques $f = (v_1, v_2, \dots, v_p)$, est de considérer la classe et les caractéristiques comme étant des variables aléatoires, puis calculer la distribution de la probabilité conditionnelle $P(c_i/f)$. Cette probabilité à posteriori peut être évaluée selon la règle de Bayes pour chaque classe c_i $i \in \{1, 2, \dots, C\}$ par :

$$P(c_i / f) = \frac{P(f / c_i)P(c_i)}{P(f)} \quad (4.6)$$

La règle de décision optimale est définie comme étant le principe d'affecter un label c_i à la classe dont la probabilité à posteriori $P(c_i / f)$ est maximale [93, 94]. Etant donné un ensemble d'apprentissage $F = [f_1, f_2, \dots, f_n]$ où chaque $f_i = (v_{i1}, v_{i2}, \dots, v_{ip})$ désigne un vecteur de

caractéristiques, considérons un nouvel élément $f = (v_1, v_2, \dots, v_p)$ à classer, la décision optimale est donnée par :

$$\hat{c} = \arg \max_c P(c_i / f) = \arg \max_c \frac{P(f / c_i)P(c_i)}{P(f)} = \arg \max_c P(f / c_i)P(c_i) \quad (4.7)$$

La difficulté de l'estimation de cette probabilité réside principalement dans le calcul des probabilités *à postériori* [93,95], une solution naïve serait de considérer que les attributs sont indépendants les uns des autres. En réalité l'hypothèse d'indépendance des variables et des hypothèses sur la distribution n'est pas toujours vérifiée, d'où l'appellation classification bayésienne naïve. Sachant que les caractéristiques sont indépendantes pour une classe donnée c_i on peut en déduire :

$$P(f / c_i) = \prod_{k=1}^n P(f_k / c_i) \quad (4.8)$$

Par ailleurs, si on considère cette hypothèse, l'algorithme Naïf de Bayes, se résume au problème suivant :

$$\hat{c}_{NB} = \arg \max_c P(c_i) \prod_k P(f_k / c_i) \quad (4.9)$$

Dans la pratique, le classificateur naïf de Bayes est aussi performant lorsque la condition d'indépendance conditionnelle des variables prédictives n'est pas vérifiée. Le processus de classification de cet algorithme est constitué de deux phases essentielles :

- *Phase d'entraînement* : Dans cette phase, il s'agit d'estimer les paramètres de la probabilité de distribution des caractéristiques f_k pour chaque classe donnée, et chaque instance de l'ensemble d'apprentissage F.

- *Phase de Prédiction* : il s'agit dans cette phase de calculer la probabilité à postériori pour chaque échantillon inconnu de l'ensemble de test, et pour chaque classe. Finalement, comme résultat L'algorithme classe l'échantillon de test selon la plus grande probabilité à postériori.

Le classificateur Naïf de Bayes s'attèle d'abord à estimer les densités de probabilités conditionnelles des attributs pour une classe donnée. En pratique, pour les variables continues la proposition la plus courante consiste à utiliser une distribution gaussienne [93, 95]. A cet effet, pour chaque classe et pour chaque attribut continu, on doit estimer la moyenne $\mu_{k,i}$ et l'écart-type $\sigma_{k,j}$ de l'attribut k de la classe i . La dernière étape est des plus simples est systématique, elle

consiste à estimer le maximum de vraisemblance par l'équation 4.9. Formellement, la distribution gaussienne de $P(f_k/c_i)$ s'écrit :

$$P(f_k/c_i) = \mathcal{N}(f / \mu_{k,i}, \sigma_{k,i}) = \frac{1}{\sigma_{k,i}} \cdot \exp\left(-\frac{1}{2\sigma_{k,i}^2} (f_k - \mu_{k,i})^2\right) \quad (4.10)$$

Il est à noter que l'hypothèse de normalité est efficace tant que les conditions de symétrie et d'uni-modalité des distributions conditionnelles sont vérifiées. Cependant, Lorsqu'elles sont multimodales avec en plus un chevauchement entre les modalités, toute approximation par une distribution gaussienne uni-modale est absolument inadaptée et produira un classificateur de mauvaise qualité. La règle de décision pour la classification bayésienne naïve pour la loi normale s'écrit finalement sous la forme suivante:

$$\hat{c}_{NB} = \arg \max_C P(c_i) \prod_k \mathcal{N}(f / \mu_{k,i}, \sigma_{k,i}) \quad (4.11)$$

La transformation de l'équation (4.9) en passant au logarithme engendre la règle de décision de Bayes sous sa forme définitive suivante :

$$d_{NB} = 2 \ln P(c_i) + \sum_k \left(-\ln \sigma_{k,i} - \left(\frac{f_k - \mu_{k,i}}{\sigma_{k,i}} \right)^2 \right) \quad (4.12)$$

4.3.2. Analyse discriminante Linéaire (LDA)

Introduite en 1936 par Fisher, l'analyse discriminante est une méthode de classification supervisée, basée sur des modèles probabilistes qui permettent de modéliser d'une façon statique la classification optimale à partir d'un ensemble de descripteur prédictifs [93, 96]. Étant donné que l'analyse discriminante est une méthode de classification par réduction d'espace, le but principal de cette méthode est de trouver une base réduite de projection qui optimise la fonction de Fisher. La projection sur cette base de dimension inférieure par rapport à l'espace original multidimensionnel, réduit considérablement la complexité des calculs tout en conservant l'information discriminante. L'analyse discriminante linéaire (LDA) consiste à rechercher dans l'espace des données les axes qui permettent de discriminer au mieux les différentes classes (plutôt que les axes qui décrivent au mieux les données, ce qui est le cas de l'ACP). La LDA permet de créer une combinaison linéaire d'un certain nombre de paramètres indépendants qui aboutit à un maximum de distance entre les moyennes des différentes classes.

Etant donné un ensemble de N échantillons (observations) $\{x_1, x_2, \dots, x_N\}$ et c classes $\{\chi_1, \chi_2, \dots, \chi_c\}$, le principe de la LDA consiste à trouver les meilleurs axes sur lesquels la projection des échantillons maximise la distance entre les moyennes de leurs classes respectives. Toutefois, la distance entre les moyennes projetées n'est pas une mesure très fiable puisqu'elle ne prend pas en compte l'écart-type au sein des classes. La solution proposée par Fisher est de maximiser une fonction qui représente la différence entre les moyennes normalisées en utilisant la covariance. Formellement, la moyenne d'une classe, la moyenne globale et la matrice de covariance S_i de la classe i sont données par les équations suivantes :

$$\mu_i = \frac{1}{N_i} \sum_{k \in \chi_i} x_k \quad (4.13)$$

$$\mu = \frac{1}{N} \sum_{k=1}^N x_k \quad (4.14)$$

$$S_i = \sum_{k \in \chi_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad (4.15)$$

Entre autres, la méthode de Fisher utilise la matrice de covariance intra-classe S_W , la matrice de covariance totale S_T , et l'éloignement entre les classes ou covariance interclasses S_B sont définies par :

$$S_W = \sum_{i=1}^c S_i \quad (4.16)$$

$$S_T = S_W + S_B \quad (4.17)$$

$$S_B = \sum_{i=1}^c |\chi_i| (\mu_i - \mu)(\mu_i - \mu)^T \quad (4.18)$$

L'idée principale est de maximiser une fonction qui va engendrer une grande séparation entre les classes d'objets projetées, tout en donnant une faible variance dans chaque classe et de façon à minimiser le chevauchement entre les classes. La projection W des échantillons de l'espace original de grande dimensions sur l'axe de direction dans un espace transformé de dimension réduite, est donnée par :

$$y_k = W^T x_k \quad k = 1, 2, \dots, N \quad (4.19)$$

Les matrices de covariance après projection seront exprimées par $\widetilde{S}_B = W^T S_B W$ et $\widetilde{S}_W = W^T S_W W$; où W est la matrice de transformation qui maximise le critère de Fisher $J(W)$. Ce dernier représente le rapport entre le déterminant de la matrice de covariance interclasses des échantillons projetés, et le déterminant de la matrice de covariance intra-classe. Formellement, il est exprimé par :

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|} \quad (4.20)$$

où l'opérateur $| \cdot |$ désigne le déterminant de la matrice. L'optimisation du critère de Fisher n'est possible que si S_W est non singulière, la projection optimale W_{opt} qui maximise le rapport donné par l'équation (4.18), autrement dit qui maximise les distances entre les centres des classes sur le nouvel axe est exprimée par l'équation:

$$\begin{aligned} W_{opt} &= \arg \max_W [J(W)] = \arg \max_W \frac{|\widetilde{S}_B|}{|S_W|} \\ &= \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1 \ w_2 \ \dots \ w_m] \end{aligned} \quad (4.21)$$

Par ailleurs, afin de trouver ce point critique, il suffit de résoudre l'équation $\frac{d}{dW} [J(W)] = 0$; en divisant ensuite par le terme $W^T S_W W$, on obtient le problème de valeurs propres généralisées suivant :

$$S_B w_i = \lambda S_W w_i \quad i = 1, 2, \dots, m \quad (4.22)$$

Finalement, la solution optimale est calculée à partir des valeurs et vecteurs propres généralisées de S_B et S_W . La figure 4.2, illustre un exemple de classification entre deux classes, l'axe de projection de la LDA est l'axe optimal qui maximise la séparation entre les classes.

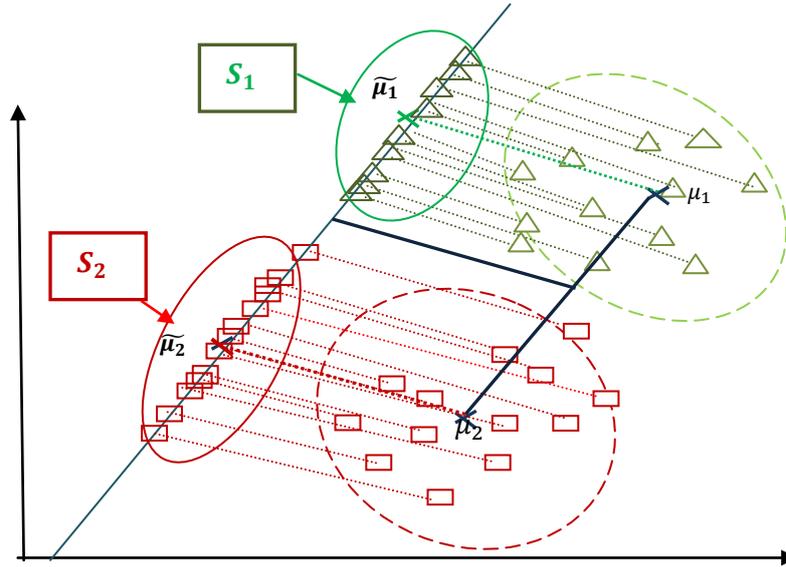


Figure 4.2. Illustration de la discrimination linéaire (LDA) pour 2 classes (projection sur l'axe optimal).

L'analyse discriminante linéaire (LDA) est largement utilisée dans plusieurs applications avec une grande dimension ; comme la reconnaissance faciale. Néanmoins, cette technique est limitée à cause du problème de singularité de la matrice S_W dont le rang est : $\text{rang}(S_W) \leq N - C$. Intuitivement, cette limitation existe lorsque le nombre d'observations est petit par rapport à la dimension de l'espace. Les auteurs dans [96] proposent une alternative au critère de Fisher, appelée *Fisherface* ; cette approche consiste à combiner l'ACP (Analyse en Composante Principale) à l'algorithme de la LDA classique qui réduit la dimension à $C - 1$. La projection par l'ACP dans l'espace de dimensions $N - C$ est exprimée par les équations suivantes :

$$z_k = W_{ACP}^T x_k \quad (4.23)$$

où

$$W_{ACP} = \arg \max_w |W^T S_T W| \quad (4.24)$$

Finalement, la projection par la LDA dans l'espace de dimension $C - 1$ est formulée par :

$$y_k = W_{LDA}^T z_k \quad (4.25)$$

avec

$$W_{LDA} = \arg \max_w \frac{|W^T W_{ACP}^T S_B W_{ACP} W|}{|W^T W_{ACP}^T S_W W_{ACP} W|} \quad (4.26)$$

A l'instar de l'ACP, la LDA est une technique qui est souvent utilisée dans la littérature en combinaison avec d'autres classificateurs dans le but de la réduction des caractéristiques. Par

ailleurs, elle fût étendue pour le cas non linéaires, plusieurs extensions ont vu le jour, telles que l'Analyse Discriminante Quadratique (QDA), l'analyse discriminant flexible (FDA) et les méthodes à noyau basées sur la LDA.

4.4. Approches non paramétriques

4.4.1. Classificateur K-NN

L'algorithme des K-NN (*K-Nearest Neighbors*) est une technique de discrimination non paramétrique, elle fait partie de la famille des algorithmes d'apprentissage supervisé. Cette méthode de classification est très simple à mettre en œuvre, dans la mesure où elle se base uniquement sur la mesure de similarité en terme de distance entre les objets de la base d'apprentissage et l'objet à tester. L'idée principale de l'algorithme des K-NN est d'affecter une classe à un nouvel échantillon en se basant sur les classes de ses k plus proches voisins dans la base d'apprentissage. En effet, cette classe correspondra à la classe majoritaire parmi celles de ses voisins. Par ailleurs, le paramètre k est une donnée définie par l'utilisateur dont l'ajustement est empirique et dépend souvent de l'expérimentation. Etant donné un ensemble d'apprentissage (Base d'apprentissage) $T = \{(x_i, c_j); i = 1, \dots, N\}$ où $c_j \in C, j = 1, \dots, J$ représente l'étiquette de la classe d'appartenance de l'échantillon $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$; décrit par p attributs. L'algorithme K-NN, calcule la distance d entre l'échantillon test x et tous les éléments de la base d'apprentissage T , ensuite il détermine les k plus proches voisins x_k qui satisfont à l'équation:

$$d(x, x_k) = \min(d(x, x_i)) \quad (4.27)$$

Finalement, la classe optimale prédite $\hat{c} = c_k$ correspond à la classe majoritaire parmi celles de ses k plus proches voisins [94]. Autrement dit, la règle de décision du classificateur K-NN consiste à attribuer à l'échantillon test un label par un simple vote sur toutes les étiquettes des k plus proches voisins, formellement défini par :

$$\hat{c} = \arg \max_c \sum_j I(v = classe(c_j)) \quad (4.28)$$

où $I(.)$ est la fonction indicatrice booléenne.

Par ailleurs, les performances du classificateur K-NN dépendent principalement de plusieurs choix à savoir :

- Le paramètre k introduit par l'utilisateur, en effet, celui-ci ne doit pas être trop large pour ne pas introduire de faux voisinage, ni trop petit pour ne pas être sensible au bruit. L'estimation de la meilleure valeur du paramètre k peut être obtenue par diverses heuristiques techniques telles que la validation croisée [94].

- La règle de décision ou règle de vote : dans la pratique, l'utilisation du principe du maximum de votes pose un problème lorsque les voisins les plus proches sont très variables relativement à leurs distances. Afin de remédier à cet inconvénient les travaux dans [94], suggèrent une technique moins sensible au choix du paramètre k . Celle-ci consiste à affecter un poids de pondération à chaque distance mutuelle entre chaque voisin et l'échantillon test, relativement à sa grandeur. Ce facteur de pondération est généralement considéré comme l'inverse de la distance au carré $w_k = 1/(d(x, x_k))^2$. Cette nouvelle approche est appelée la règle des k -plus proches voisins avec distances pondérées. Formellement, cette règle d'affectation est exprimée par :

$$\hat{c} = \arg \max_c \sum_j w_j I(v = classe(c_j)) \quad (4.29)$$

- Finalement, un des paramètres à considérer est qui n'est pas des moindres est le type de la distance utilisée comme mesure de similitude (voit Tableau 4.2). Plusieurs distances de similarité ont été utilisées dans la littérature, nous en citerons à titre d'exemple la distance de Minkowski, la distance Euclidienne, etc..

| Mesure | Formulation | Description |
|----------------------|--|---|
| Distance Minkowski | $D_{ij} = \left(\sum_{l=1}^d (x_{il} - x_{jl})^{1/n} \right)^n$ | Invariante à la translation et à la rotation |
| Distance Euclidienne | $D_{ij} = \left(\sum_{l=1}^d (x_{il} - x_{jl})^{1/2} \right)^2$ | Très utilisée : cas particulier de la distance de Minkowski $n = 2$ |
| Distance City-block | $D_{ij} = \sum_{l=1}^d x_{il} - x_{jl} $ | Cas particulier de la distance de Minkowski $n = 1$ |

Tableau 4.2. Exemple de distances utilisées pour le classificateur K -NN

4.4.2. La classification en utilisant les réseaux de neurones (ANN)

4.4.2.1. Structure d'un neurone formel

La notion de réseaux de neurones artificiels découle des travaux des chercheurs Warren Mc Culloch et Walter Pitts en 1943 [97]. Concrètement, la première application fût menée par Frank Rosenblatt qui a construit le perceptron, un réseau de neurones pour la reconnaissance des formes. Le neurone artificiel a été inspiré du neurone biologique, le modèle mathématiquement du neurone formel (voir figure. 4.3), est composé d'un ensemble d'entrées auxquelles sont affectés des poids respectifs, d'un intégrateur dont le rôle est d'effectuer la combinaison linéaire des entrées et de leurs poids respectifs, et d'une fonction d'activation qui permet de décider à partir d'un seuil donné si le neurone est activé ou non [98].

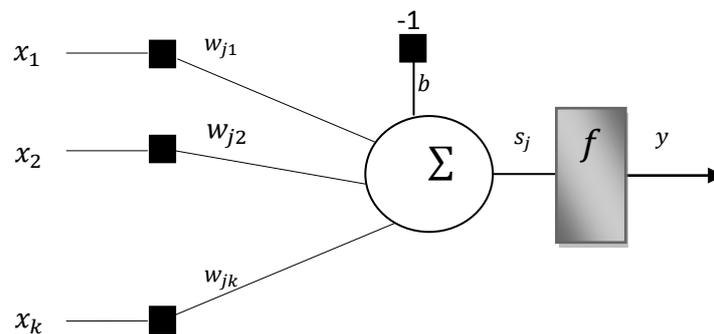


Figure 4.3. Modèle d'un neurone artificiel

La sortie s de l'intégrateur ou niveau d'activation du neurone, représente la somme pondérée des entrées du neurone et de leurs poids respectifs. L'entrée particulière -1 , à laquelle est associée un poids b est appelé seuil d'activation ou biais. Ce poids particulier tire son nom du fait que le neurone n'est activé que si la somme pondérée des entrées autres que l'entrée -1 , et de leurs poids respectifs soit supérieure à b . Etant donné un vecteur x de composantes (x_1, \dots, x_k) représentant les entrées du réseau de neurones, w le vecteur de coefficients de pondération ou coefficients synaptiques (w_{j1}, \dots, w_{jk}) qui reflète l'efficacité de la connexion synaptique, la sortie du neurone est alors donnée par :

$$y = f(s_j) \quad (4.30)$$

où

$$s_j = w^T x - b = \sum_{i=1}^k w_{ji} \cdot x_i - b \quad (4.31)$$

avec w_{ji} le poids associé à la $i^{\text{ème}}$ entrée du neurone j . La notion d'activation du neurone dépend du choix de sa fonction d'activation f dont il existe plusieurs types. Nous en citons les fonctions

seuil, linéaire, et sigmoïde qui sont généralement les plus utilisées. Les neurones binaires ont des fonctions d'activation discontinues suivant que les états soient (0,1) ou (-1,1), les fonctions seuils peuvent être soit la fonction échelon ou fonction de Heaviside. Dans le cas d'une fonction seuil la valeur de sortie $s = 1$ correspondra à l'état activé du neurone artificiel, et la valeur $s = 0$ correspondra à l'état non activé. Par contre, dans les cas où les états des neurones sont continus, on utilise les fonctions sigmoïdes (voir Tableau 4.3).

| Fonction d'activation f | Expression | Nom MATLAB |
|-----------------------------------|---|------------|
| Seuil : fonction échelon | $f(n) = \begin{cases} 1, & n \geq 0 \\ 0, & x < 0 \end{cases}$ | hardlim |
| Seuil symétrique : fonction signe | $f(n) = \begin{cases} 1, & n \geq 0 \\ -1, & x < 0 \end{cases}$ | Hardlims |
| Linéaire | $f(n) = n$ | Purelin |
| Sigmoïde | $f(n) = \frac{1}{1 + \exp(-n)}$ | Logsig |

Tableau 4.3. Fonction d'activation d'un modèle de neurone artificiel.

4.4.2.2. Architecture d'un réseau de neurones

Un réseau de neurones appelé aussi Perceptron multicouche (PMC), est organisé en plusieurs couches de neurones, chacune composée de k neurones. Les différentes couches sont interconnectés, en effet les entrées d'une couche donnée constituent les sorties de celle qui la précède, à l'exception de la première et de la dernière couche. Effectivement, la première couche est reliée directement aux entrées qui sont assimilées à l'excitation du réseau par un stimuli, alors que les sorties de la dernière ne sont reliées à rien dans la mesure où elles représentent la réponse du réseau de neurones au stimulus. Les couches situées entre la première et la dernière sont dites couches cachées (voir figure 4.4), car le réseau de neurones peut être perçu comme une boîte noire dont l'interface est constituée de la première et la dernière couche. L'entraînement des réseaux de neurones comprend deux phases, une phase d'apprentissage et une phase d'opération. L'apprentissage est un processus à travers lequel, les paramètres (poids de connexion) du réseau sont ajustés.

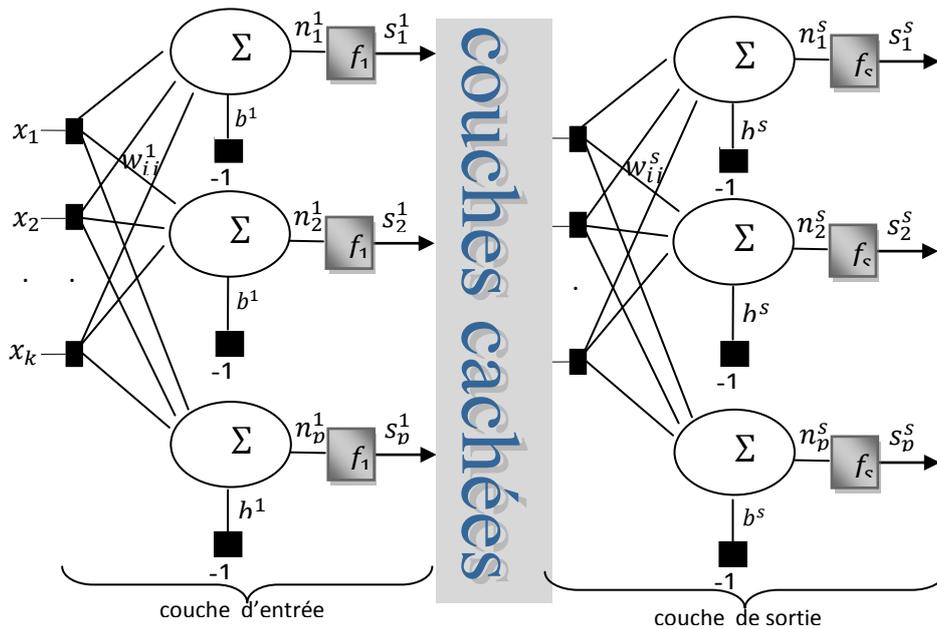


Figure 4.4. Modèle du Perceptron multicouche

Une fois le réseau est entraîné, il représente une base de connaissance statique qui peut être appelé durant la phase d'opération. L'entraînement du réseau nécessite en général :

- Une base de données d'apprentissage pour le calcul des poids synaptiques.
- Une base de données de test, pour évaluer la qualité de l'apprentissage.

Dans ce qui suit, nous allons aborder un réseau spécifique qui a prouvé son efficacité dans la pratique pour des problèmes de classification et de prédiction. Il s'agit du réseau multicouche avec son algorithme d'apprentissage de retro-propagation d'erreur (*Back-propagation*) qui est une généralisation de l'algorithme d'apprentissage LMS (Least Mean Square) en utilisant les fonctions d'activation linéaires [93, 98]. Dans le processus de classification, on utilise généralement des réseaux à deux ou trois couches sigmoïdes. Le choix d'une fonction d'activation non linéaire de type sigmoïde, n'est pas arbitraire, elle permet de procéder à un apprentissage automatique.

4.4.2.3. Algorithme de BACK-PROPAGATION

L'algorithme de Back-propagation est un algorithme d'apprentissage supervisé avec une fonction d'activation sigmoïde continue et dérivable. La topologie du réseau auquel cet algorithme est adopté (voir figure 4.5), est formée de plusieurs couches de neurones, sachant que plus le nombre de couche est grand, plus le réseau apprend vite.

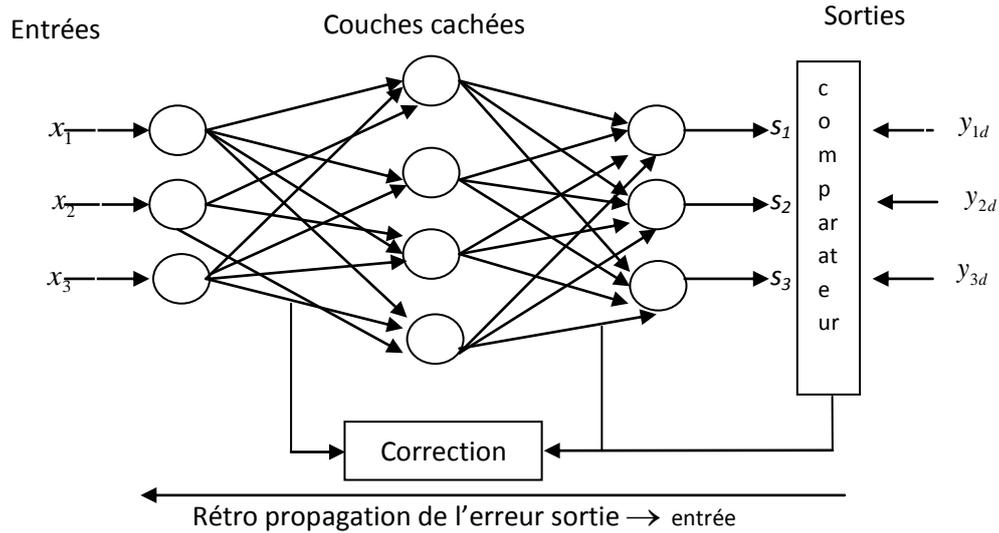


Figure 4.5 Synoptique de la méthode de Back propagation

La méthode d'apprentissage est basée sur l'algorithme de la descente du gradient qui permet de trouver les poids optimaux en minimisant la valeur moyenne de l'erreur quadratique sur l'ensemble des échantillons d'entraînement du réseau. Pour converger vers le minimum d'erreur, la technique ajuste les poids dans la direction opposée à celle du gradient instantané de la fonction erreur. Ainsi la fonction coût à minimiser est :

$$E(w) = \sum_{p=1}^M E_p(w) \quad (4.32)$$

où M représente le nombre d'exemples d'apprentissage et $E_p(w)$ le carré de l'erreur associée au $p^{\text{ème}}$ exemple.

$$E_p(w) = \frac{1}{2} \sum_{j=1}^N (d_{pj} - y_{pj})^2 \quad (4.33)$$

N : étant le nombre de neurones dans la couche de sortie, d_{pj} : la sortie désirée au nœud de sortie et y_{pj} la sortie du réseau au nœud j . Pour un réseau multicouche à R entrées, M sorties ; composé de L couches cachées. La somme des entrées de la $j^{\text{ème}}$ neurone caché est donnée par :

$$s_{pj}^L = \sum_{i=1}^R w_{ji}^L \cdot x_{pi} + b_j^L \quad (4.34)$$

où s_j^L est le potentiel somatique du neurone de $L^{\text{ème}}$ couche cachée, x_i : le vecteur de la $i^{\text{ème}}$ entrée $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]$, L : l'indice de la couche cachée et w_{ji}^L les coefficients synaptiques de la connexion de la $i^{\text{ème}}$ entrée du neurone j . Formellement, la sortie non linéaire du neurone j dans la couche cachée L correspond à l'équation:

$$y_j^L = f(s_j^L) = f\left(\sum_{i=1}^R w_{ji}^L \cdot x_{pi} + b_j^L\right) \quad (4.35)$$

La fonction utilisée dans cet algorithme est une fonction non linéaire de type sigmoïde de la forme $f(x) = \frac{1}{1+e^{-x}}$, $-1 < f(x) < 1$, de ce fait l'équation de la couche de sortie est déduite par :

$$s_{pj} = \sum_{i=1}^N w_{ji} \cdot y_j^L + b_j \quad (4.36)$$

N : étant le numéro de la couche de sortie. Par ailleurs, l'ajustement des coefficients est réalisé par la méthode du gradient basée essentiellement sur la formule itérative suivante :

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n) \quad (4.37)$$

où $\Delta w_{ji}(n) = -\mu \frac{\partial E_p(w)}{\partial w_{ji}(n)}$, n : numéro d'itération, μ : pas d'apprentissage désignant la vitesse de convergence, dont la valeur est généralement choisie expérimentalement. Le rapport $\frac{\partial E_p(w)}{\partial w_{ji}(n)}$ et calculé conformément à l'équation :

$$\frac{\partial E_p(w)}{\partial w_{ji}(n)} = \left(\frac{\partial E_p}{\partial y_{pj}}\right) \cdot \left(\frac{\partial y_{pj}}{\partial s_{pj}}\right) \cdot \left(\frac{\partial s_{pj}}{\partial w_{ji}}\right) \quad (4.38)$$

sachant que $\frac{\partial y_{pj}}{\partial s_{pj}} = f'(s_{pj})$, $\frac{\partial s_{pj}}{\partial w_{ji}} = x_{pj}$. Par ailleurs, le neurone appartenant à la couche de sortie est caractérisé par l'équation $\frac{\partial E_p}{\partial y_{pj}} = -(d_{pj} - y_{pj})$, Tandis que le neurone appartenant à la couche cachée satisfait l'équation $\frac{\partial E_p}{\partial y_{pj}} = -\sum_{k=1}^N (d_{pk} - y_{pk}) f'(s_{pk}) \cdot w_{pj}$. A cet effet, la règle d'ajustement des coefficients sera définie par l'équation définitive suivante :

$$w_{ji}(n+1) = w_{ji}(n) - \mu \cdot \Delta j \cdot x_{pi} \quad (4.39)$$

avec $\Delta j = -\left(\frac{\partial E_p}{\partial y_{pj}}\right) \cdot f'(s_{pj})$ et μ devant être petit, de l'ordre 0.05 à 0.25 pour assurer la convergence du réseau.

Afin d'optimiser l'algorithme en terme de rapidité et de vitesse de convergence, les auteurs dans [93] proposent d'utiliser la technique dite de "Momentum". A cet effet, la règle d'ajustement des coefficients considérée est décrite par l'équation:

$$w_{ji}(n+1) = w_{ji}(n) - \mu \cdot \Delta j \cdot x_{pi} + \alpha \cdot (w_{ji}(n) - w_{ji}(n-1)) \quad (4.40)$$

En respectant la contrainte $0 < \alpha < 1$.

- Algorithme de Back-propagation

L'algorithme de Back-propagation est détaillé ci-après par les étapes suivantes :

-
- 1- Initialisation des poids w_{ji} et les seuils internes des neurones à de petites valeurs aléatoires,
 - 2- Présenter le vecteur d'entrée $X_p(x_{p1}, x_{p2}, \dots, x_{pn})$ et la sortie désirée $D_p(d_{p1}, d_{p2}, \dots, d_{pn})$,
 - 3- Calcul des paramètres du PMC
 - La somme des entrées des neurones de la couche cachée (Eq. 4.35)
 - Les sorties des neurones de la couche cachée (Eq. 4.36)
 - La somme des entrées de la couche de sortie (Eq. 4.37)
 - La sortie du réseau (Eq. 4.37)
 - 4- Calcul des termes de l'erreur pour les neurones de la couche de sortie
 - 5- Réinjection de l'erreur de sortie dans le réseau et calcul des termes de l'erreur pour les neurones de la couche cachée
 - 6- Ajustement des coefficients synaptiques (poids) de la couche de sortie et de la couche cachée, en utilisant soit l'Equation 4.39 ou l'équation 4.40.
 - 7- Si la condition sur l'erreur ou sur le nombre d'itérations est atteinte, alors on arrête, sinon revenir à l'étape 3 jusqu'à la stabilisation du système.
-

4.4.3. Les Machines à Vecteurs supports

4.4.3.1. Le concept

Les machines à vecteurs supports (SVMs : Support vector Machines) est une technique de classification supervisée introduite par *Vladimir Vapnik* [92] dans les années 60. La vocation originelle de cette méthode fût la séparation linéaire d'objets appartenant à deux classes différentes par un hyperplan optimal. Elle a été étendue plus tard pour prendre en compte le cas multi-classes.

Soit $\{x_i\}_{i=1}^{i=m} \in \mathbb{R}^k$, l'ensemble d'apprentissage composé d'objets caractérisés par un vecteur de k composantes chacun, et appartenant à deux classes différentes : une classe positive libellée $+1$ et une classe négative libellée -1 . Soit également $y_i \in \{1, -1\}$, la variable représentant les étiquettes des deux classes. Le but des SVMs est de déterminer la frontière linéaire définie par l'équation $wx + b$ entre les objets positifs et les objets négatifs (voir figure 4.6), où w représente le vecteur normal à ce séparateur linéaire, et b le biais. Par ailleurs, déterminer un hyperplan revient à définir le couple de valeurs (w, b) qui le caractérise. Dans la mesure où plusieurs solutions sont possibles, quel sera donc le critère qui définira la séparation optimale?.

Les SVMs sont basés sur la théorie baptisée SRM (Structural Risk Minimization) [92]. Cette technique consiste à choisir l'hyperplan optimal parmi tous les hyperplans séparateurs possibles; celui-ci doit résoudre le compromis entre la minimisation de l'erreur empirique, et la maximisation de la marge séparant les deux classes. Contrairement à d'autres techniques de classification qui ne considèrent que la minimisation de l'erreur d'apprentissage.

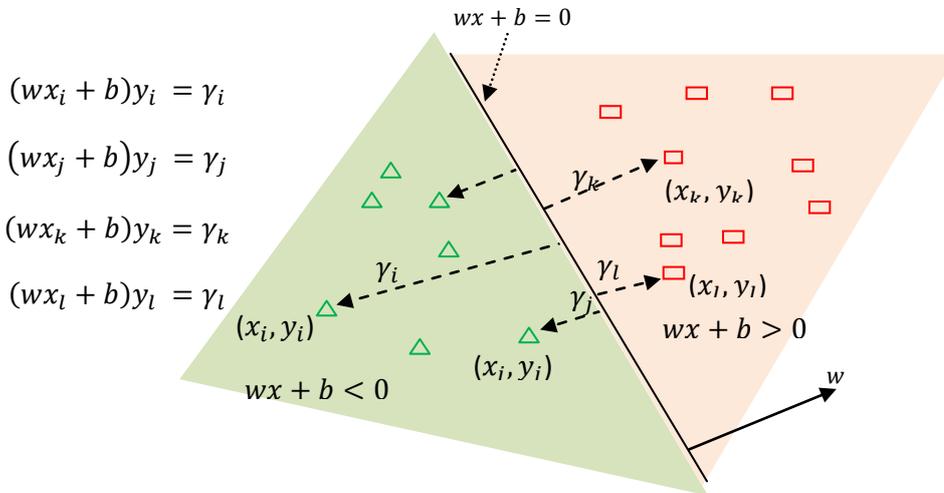


Figure 4.6. Séparation linéaire des classes carré (étiquetée +1), et triangle (étiquetée -1), respectivement positive et négative par un hyperplan. w représente le vecteur normal à l'hyperplan, les γ_i représentent la marge séparant chaque objet de l'hyperplan.

La marge γ_i qui sépare chaque objet x_i étiqueté y_i de l'hyperplan séparateur est formulée par l'équation suivante :

$$(wx_i + b)y_i = \gamma_i \quad (4.41)$$

L'hyperplan optimal doit maximiser la plus petite marge le séparant d'un objet, définie par l'équation :

$$\gamma = \min_{i=1..m} \gamma_i \quad (4.42)$$

de telle sorte que la marge de chaque objet de l'ensemble d'apprentissage soit supérieure ou égale à celle-ci, conformément à l'équation :

$$\begin{cases} \max_{w,b} \gamma \\ (wx_i + b)y_i \geq \gamma, \forall i=1, \dots, m \end{cases} \quad (4.43)$$

Tout objet x_i étiqueté +1 situé exactement sur l'hyperplan H_2 (voir figure 4.7) obéit à l'équation :

$$wx_i + b = 1 \quad (4.44)$$

Similairement, tout objet x_j étiqueté -1 situé sur l'hyperplan H_1 (voir figure 4.7) obéit à l'équation :

$$wx_j + b = -1 \quad (4.45)$$

En soustrayant l'équation (4.45) de l'équation (4.44) on obtient :

$$w(x_i - x_j) = 2 \quad (4.46)$$

Donc l'écart séparant les deux objets x_i et x_j s'exprimera par l'équation :

$$\|x_i - x_j\| = \frac{2}{\|w\|} \quad (4.47)$$

Par conséquent, la séparation ou la classification se résume à la maximisation de la grandeur $2/\|w\|$ qui représente la marge séparant les objets des deux classes positive et négative. Le problème quadratique formulé par Cortes et Vapnik [99] exprimé ici sous une forme plus simplifiée qui ne traite que le cas des objets linéairement séparables se présente comme suit :

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ y_i (w \cdot x_i + b) \geq 1 \quad \forall i = 1, \dots, m \end{cases} \quad (4.48)$$

Effectivement, maximiser la grandeur $2/\|w\|$ revient à minimiser $\frac{1}{2} \|w\|^2$, le facteur $1/2$ étant rajouté pour des convenances de calcul. L'équation (4.48) représente la formulation d'un problème d'optimisation d'une fonction objective quadratique sous contrainte. La contrainte permet d'assurer que chaque élément appartenant à la classe +1 soit positionné dans la région située au dessus de l'hyperplan H_2 , et similairement que tout objet de la classe -1 soit positionné dans la partie inférieure à l'hyperplan H_1 . Par ailleurs, afin de prendre en considération le cas des objets non linéairement séparables, ce qui équivaut à tolérer des erreurs d'apprentissage qui se traduisent par des objets positifs situés dans la région des objets négatifs et vice versa, les SVMs ont été étendues pour résoudre un compromis entre la maximisation de la marge et la minimisation des erreurs d'apprentissage. A cet effet, la formulation (4.48) a été reconsidérée comme suit :

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i & , \forall i = 1, \dots, m \\ y_i (w \cdot x_i + b) \geq 1 - \xi_i & , \forall i = 1, \dots, m . \\ \xi_i \geq 0 & , \forall i = 1, \dots, m \end{cases} \quad (4.49)$$

La fonction objective exprimée dans (4.49) est composée de la somme de deux termes. Le premier représente la marge et le second désigne la somme de tous les écarts ξ_i des objets de leurs classes d'origine (voir figure 4.7), qui définit une majoration sur le nombre des objets mal classés. Le

deuxième terme est aussi pondéré par un facteur C défini par l'utilisateur lors de la phase d'entraînement du classificateur, et qui permet de résoudre le compromis entre la marge et l'erreur. Effectivement, si C se voit affecter une petite valeur, alors la priorité sera à l'optimisation de la marge au détriment de l'erreur, sinon c'est la minimisation de l'erreur d'apprentissage qui sera privilégiée.

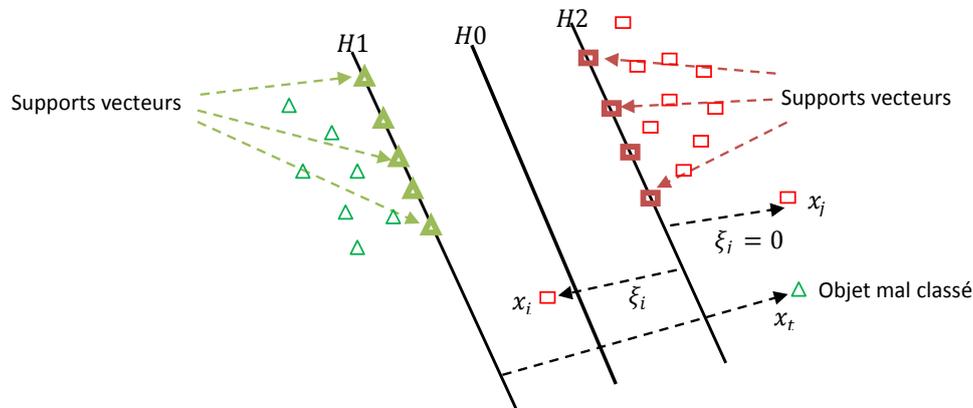


Figure 4.7. Les objets triangles et carrés représentés en gras, situés respectivement sur les hyperplans $H_1: wx + b = -1$ et $H_2: wx + b = 1$ sont appelés supports vecteurs, les objets x_i , et x_t sont considérés comme non classés, les objets positifs situés dans la partie entre H_2 et H_0 et les objets négatifs situés entre H_1 et H_0

La transformation du problème quadratique (4.49) dans sa forme duale se traduit par l'équation:

$$\begin{cases} \max \mathcal{L}_D = \sum_{i=1}^m \lambda_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j \cdot x_j \\ 0 \leq \lambda_i \leq C \\ \sum_{i=1}^m \lambda_i y_i = 0 \end{cases}, \quad \forall i = 1, \dots, m \quad (4.50)$$

Cette formulation qui peut être résolue entre autres par les techniques standards de résolution des problèmes quadratiques, découle de l'utilisation de la méthode des multiplicateurs de Lagrange pour la résolution du problème (4.49). Il en résulte aussi l'expression suivante du vecteur normal à l'hyperplan séparateur :

$$w = \sum_{i=1}^m \lambda_i x_i y_i \quad (4.51)$$

De l'équation (4.51), il en découle que le vecteur normal peut être exprimé seulement en fonction des objets de l'ensemble d'apprentissage, les multiplicateurs de Lagrange qui leur sont associés, et leurs libellés respectifs. La résolution du problème d'optimisation (4.51) montre également qu'à chaque objet bien classé correspond un multiplicateur de Lagrange nul, sauf pour les objets situés

exactement sur les hyperplans H_0 , H_1 , et H_2 appelés vecteurs supports et auxquels correspondent des multiplicateurs strictement positifs.

Par conséquent, l'hyperplan optimal séparateur s'exprimera en fonction des supports vecteurs uniquement, ce qui réduit la complexité de calcul considérablement dans le cas d'un ensemble d'apprentissage de grande taille. L'équation de l'hyperplan se réduit alors à ce qui suit :

$$\left(\sum_{j=1}^{n_{sv}} \lambda_j x_j y_j \right) x_i + b \quad (4.52)$$

avec : $b = Moyenne(y_i - \sum_{j=1}^{n_{sv}} \lambda_j x_j y_j \cdot x_i)$, $i = 1, \dots, n_{sv}$, où n_{sv} désigne le nombre de vecteurs supports. La prédiction de la classe d'un objet x_k qui correspond à la phase de test, s'effectue alors par la fonction de décision suivante :

$$classe(x_k) = \begin{cases} +1 & \text{si } wx_k + b \geq 0 \\ -1 & \text{si } wx_k + b < 0 \end{cases} \quad (4.53)$$

où

$$classe(x_k) = \text{signe}(\sum_{i=1}^{n_{sv}} \lambda_i y_i x_i \cdot x_k + b), \quad (4.54)$$

sachant que :

$$\text{signe}(f(x)) = \begin{cases} +1 & \text{si } f(x) \geq 0 \\ -1 & \text{si } f(x) < 0 \end{cases} \quad (4.55)$$

La séparation des objets par un hyperplan linéaire n'étant pas toujours réalisable, il existe des situations où elle ne peut être effectuée à cause de la distribution des objets de l'ensemble d'apprentissage. Afin de remédier à ce problème, l'idée serait d'effectuer une projection des objets de leur espace d'origine dit espace d'entrée dans un espace de plus grande dimension, voire de dimension infinie dit espace de *redescription*, par le biais d'une transformation non linéaire φ (voir figure 4.8).

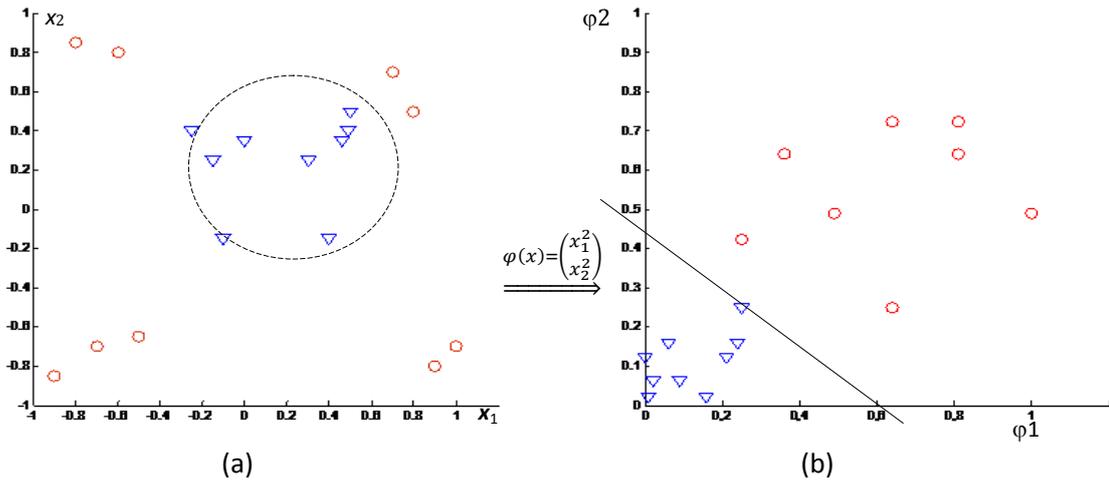


Figure 4.8. (a): représente un ensemble d'objets non linéairement séparables, (b) : montre qu'après application de la transformation non linéaire, les objets en question pouvaient être séparés par une droite dans \mathbb{R}^2 .

Le problème d'optimisation quadratique exprimé par l'équation (4.50), et l'expression de l'hyperplan séparateur décrite par (4.52) seront alors redéfinis comme suit :

$$\begin{cases} \max \mathcal{L}_D = \sum_{i=1}^m \lambda_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j \phi(x_i) \cdot \phi(x_j) \\ \lambda_i \geq 0 \\ \sum_{i=1}^m \lambda_i y_i = 0 \end{cases}, \quad \forall i = 1, \dots, m \quad (4.56)$$

$$\sum_{i=1}^{n_{sv}} \lambda_i y_i \phi(x_i) \cdot \phi(x_k) + b, \quad (4.57)$$

Le produit scalaire $\phi(x_i) \cdot \phi(x_j)$ engendre une grande complexité lorsqu'il est calculé dans un espace de haute dimension voire de dimension infinie, toutefois sous certaines conditions¹, l'utilisation de fonctions particulières dites noyaux permet de calculer ce produit scalaire dans l'espace original ou d'entrée, au lieu de l'effectuer dans l'espace de projection, conformément à l'équation :

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (4.58)$$

Ce principe est connu sous l'appellation de l'astuce du noyau. Il permet de rendre implicite la projection dans l'espace de projection ramenant ainsi le problème d'optimisation de l'équation (4.56) à sa formulation définitive suivante :

¹ Conditions de Mercer : une fonction $K: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ est une fonction noyau si elle est symétrique, i.e $K(\mathbf{u}, \mathbf{v}) = K(\mathbf{v}, \mathbf{u})$ pour tout \mathbf{u}, \mathbf{v} , et semi-définie positive, i.e pour tout $J \in \mathbb{N}$, tout vecteur $\boldsymbol{\lambda} \in \mathbb{R}^J$ et tout J-uplet $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_J$ d'éléments de \mathcal{X} , $\sum_{1 \leq j, k \leq J} \lambda_j \lambda_k K(\mathbf{x}_j, \mathbf{x}_k) \geq 0$.

$$\begin{cases} \max \mathcal{L}_D = \sum_{i=1}^m \lambda_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j K(x_i, x_j) \\ \lambda_i \geq 0 \\ \sum_{i=1}^m \lambda_i y_i = 0 \end{cases}, \quad \forall i = 1, \dots, m \quad (4.59)$$

et par conséquent, l'équation de l'hyperplan séparateur à ce qui suit :

$$\sum_{i=1}^{n_{sv}} \lambda_i y_i K(x_i, x_j) + b \quad (4.60)$$

Les fonctions noyaux sont des fonctions symétriques, semi-définies positives appliquées à deux vecteurs. Les fonctions noyaux les plus classiquement utilisées dans le domaine de la classification, spécifiquement la reconnaissance des formes sont décrites par les équations suivantes :

$$K(x, y) = x^T y \quad (4.61)$$

$$K(x, y) = ((x \cdot y) + c)^d \quad (4.62)$$

$$K(x, y) = e^{-\gamma \|x-y\|^2} \quad (4.63)$$

$$K(x, y) = \tanh(\gamma x^T y + c) \quad (4.64)$$

où γ , c , et d sont des paramètres qui doivent être ajustés afin d'améliorer les performances de classification. Les équations ci-dessus désignent respectivement les fonctions noyaux : linéaire, polynomiale, radiale ou RBF, et sigmoïde. Toutefois, on peut construire d'autres fonctions noyaux, à partir des fonctions classiques, les équations suivantes décrivent les règles de construction.

$$K(x, y) = K_1(x, y) + K_2(x, y) \quad (4.65)$$

$$K(x, y) = \beta K_1(x, y) \quad (4.66)$$

$$K(x, y) = K_1(x, y) \cdot K_2(x, y) \quad (4.67)$$

$$K(x, y) = f(x)f(y) \quad (4.68)$$

$$K(x, y) = K_1(\Phi(x), \Phi(y)) \quad (4.69)$$

$$K(x, y) = x^T A y \quad (4.70)$$

$$K(x, y) = P(K_1(x, y)) \quad (4.71)$$

$$K(x, y) = e^{K_1(x, y)} \quad (4.72)$$

où K_1, K_2 : désignent deux fonctions noyaux, β : un réel positif, f : une fonction à valeurs dans \mathbb{R} , P : un polynôme à coefficients positifs, Φ : une fonction de projection de vecteur dans un espace vectoriel, et A : une matrice semi-définie positive.

4.4.3.2. Extension des SVMs au cas multi-classes

Les SVMs ont été à l'origine conçus pour la classification binaire, l'algorithme a été étendu au cas multi-classes par des méthodes dites de décomposition. Nous citerons les travaux de Krebel [100] qui proposa la méthode baptisée *One-Against-One* (OAO), dont l'idée principale fût de décomposer le problème de classification de k classes en $k(k - 1)/2$ problèmes de classifications binaires. Cette méthode consiste à confronter les classes deux à deux, la classe élue sera alors déterminée selon le principe du maximum de votes. En effet, lors de chaque classification binaire, une classe se voit affectée un vote si sa fonction de décision renvoie la valeur 1. Par conséquent, un objet de test est affecté à la classe ayant cumulé le plus de votes.

La figure 4.9 ci-dessous illustre cette méthode en considérant trois classes. Les colonnes du tableau désignent les différentes classifications binaires possibles qui sont au nombre de 3. Ceci exclut les classifications binaires $C2/C3$, $C1/C3$, et $C1/C2$ qui ne sont pas réellement effectuées mais déduites de la relation $D_{ij} = -D_{ji}$, où D_{ij} désigne la fonction de décision de la classification C_i contre C_j . Les lignes désignent les objets de l'ensemble de test.

| | $C3/C2$ | $C3/C1$ | $C2/C3$ | $C2/C1$ | $C1/C3$ | $C1/C2$ | Classe élue |
|-------|----------|----------|-----------|----------|-----------|-----------|------------------------|
| $X1$ | D_{32} | D_{31} | $-D_{32}$ | D_{21} | $-D_{31}$ | $-D_{21}$ | $\arg \max_i D_i(X_1)$ |
| $X2$ | | | | | | | |
| | | | | | | | |
| Xm | | | | | | | |

$$D_3 = D_{32} + D_{31} \quad D_2 = -D_{32} + D_{21} \quad D_1 = -D_{31} - D_{21}$$

Figure 4.9. Principe de la méthode *One-Against-One* appliquées à trois classes.

D'autres travaux ont été basés sur le principe de décomposition. Nous en citerons finalement la méthode *One-Against-All* (OAA) introduite par Vapnik [99], et dont le principe consiste à considérer les classifications binaires de chaque classe contre les autres classes réunies. Cette technique engendre pour un problème de classification de n classes, n classifications

binaires. La fonction de décision qui permet de prédire la classe d'un objet x de l'ensemble de test, est exprimée par l'équation :

$$\arg \max_i D_i(x) \quad (4.73)$$

Où $D_i(x)$, désigne la fonction de décision de la classe i contre les autres classes réunies. Ces méthodes de décomposition sont les plus communément utilisées dans le cas multi-classes, bien qu'elles présentent l'inconvénient d'engendrer des objets non classés, problème qui découle de la possibilité de la présence de plusieurs maximums.

4.4.3.3. Les SVM dans la pratique

Dans la pratique l'utilisation des SVMs sous entend le choix de la fonction noyau, ainsi que ses paramètres. Réellement, il n'existe pas de règles bien définies pour appréhender ce problème, l'expérimentation demeure l'ultime recours, où on est souvent confronté à un compromis entre les performances de classification, et la complexité engendrée par la fonction noyau. Certaines études [101] suggèrent que la fonction radiale (RBF) est un bon premier choix, dans la mesure où elle donne les mêmes résultats de performances que la fonction linéaire pour certaines valeurs du paramètre de pénalité C , définit par l'utilisateur. La fonction radiale compte moins de paramètres à ajuster que la fonction polynômiale, aussi elle engendre moins de complexité numérique que cette dernière pour des degrés élevés du polynôme. Par ailleurs, les auteurs mentionnent que la fonction sigmoïde donne les mêmes résultats que la fonction radiale sous certains choix de leurs paramètres respectifs. Cependant, si le vecteur de caractéristiques décrivant les objets est de grande taille, ils proposent d'utiliser la fonction linéaire au lieu de la fonction radiale, qui dans ce cas ne présente pas un choix judicieux.

Les SVMs est une technique bien fondée mathématiquement, qui a démontré son efficacité dans plusieurs domaines, tels que la vision par ordinateur, la reconnaissance des formes, la biométrie, etc.. Cette technique se base sur l'optimisation d'une fonction objective convexe qui présente l'avantage de toujours converger vers un minimum global. Contrairement à d'autres techniques, telles que les réseaux de neurones où le problème des minimas locaux se pose souvent, particulièrement lorsqu'il s'agit d'une base d'apprentissage de taille réduite. Cependant, cette méthode souffre de quelques inconvénients qui peuvent altérer ses performances, essentiellement quand la base d'apprentissage est de grande taille, ce qui engendre une énorme matrice du noyau dont chaque entrée représente la valeur de la fonction noyau calculée sur les vecteurs caractérisant les objets d'apprentissage considérés deux à deux. Cette matrice renferme

$m \times m$ entrées pour une base d'apprentissage comptant m objets. Une multitude de travaux ont été réalisés pour remédier à ce problème dont l'idée principale est la réduction de la taille de la base d'apprentissage sans trop de perte d'information, autrement dit le principe est de remplacer les objets d'apprentissage par des prototypes. Nous en citons les travaux de Cervantes [102], qui proposa d'utiliser la FCM (Fuzzy C-Means) un algorithme non supervisé qui consiste à regrouper les objets en clusters, en considérant en plus de la mesure de similitude, la fonction d'appartenance floue de chaque objet, pour ne considérer finalement que les centres des clusters comme base d'apprentissage. Gilles Lebrun [103] a proposé d'utiliser les vecteurs de quantification pour compresser base d'apprentissage selon le même principe.

La sensibilité des SVMs aux aberrances appelés aussi "*outliers*", est un inconvénient majeur qui altère considérablement ses performances. Ces objets atypiques constituent un inconvénient de la plupart des méthodes statistiques. Afin de remédier à ce problème, la suppression des aberrances de la base d'apprentissage s'impose alors comme phase préalable à la classification. Malheureusement, cette solution présente l'inconvénient d'engendrer plus de complexité. Lin et al [104] ont proposé une technique baptisée FSVM (Fuzzy Support Vector Machines) qui consiste à reformuler la fonction objective en pondérant chaque objet selon son importance dans sa classe par une fonction d'appartenance floue. La complexité engendrée se trouve alors réduite, et les performances de classification significativement améliorées.

4.5. Conclusion

Dans ce chapitre nous avons passé en revue les principales techniques de classification et leurs performances respectives souvent utilisées dans le domaine de la biométrie, on retiendra que l'utilisation de ces méthodes dépend souvent de la nature du problème. En effet, chaque classificateur présente des avantages et des inconvénients, en termes de performances et de complexité de calcul. Certaines méthodes présentent une grande simplicité de mise en œuvre, dont nous citons les KNN qui ne nécessite pas une phase d'apprentissage. D'autres présentent une grande complexité dans la phase d'apprentissage, telles que les SVMs, et le classificateur bayésien naïf. Cependant ces deux méthodes présentent de bonnes performances.

Afin de tirer partie des avantages de chaque classificateur, des techniques de fusion de classificateurs ont été proposées pour améliorer la qualité des systèmes de reconnaissances biométriques.

La biométrie s'impose de plus en plus comme un moyen robuste de sécurité. Les performances d'un système biométrique, dépendent essentiellement de l'information choisie (modalité biométrique) et du classificateur mis en œuvre. La fusion de l'information et des classificateurs offre une solution prometteuse pour le développement d'un système de classification robuste et à haute performance.

Depuis les années 90, la combinaison des classificateurs (parallèle, série ou hybride), a été une des directions de recherche les plus prometteuses pour améliorer nettement les performances des systèmes. Parmi les différentes approches de fusion des classificateurs, l'architecture parallèle est de loin celle qui a donné une simplicité de mise en œuvre et une grande efficacité. Pour la prise de décision finale, plusieurs stratégies ont été proposées dans la littérature en fonction des informations fournies par les classificateurs, telles que : le vote à la majorité, la règle du produit, la règle de la somme, etc..

Chapitre 5

Les transformées géométriques pour l'identification/vérification des images palmaires

5.1. Introduction

La biométrie est une branche de la science dont le but est l'identification et la vérification d'identités d'individus en utilisant leurs caractéristiques physiologiques et comportementales. Le processus de vérification sert à déterminer si un individu appartient à la classe imposteur (non accepté par le système) ou à la classe authentique (accepté par le système). Par contre, le processus d'identification sert à identifier la personne cible. Les caractéristiques relatives à l'aspect physiologique les plus communément utilisées, et considérées comme les plus efficaces sont : l'empreinte digitale et l'empreinte palmaire. Dans ce chapitre nous allons présenter l'application des transformations géométriques à un domaine spécifique de la biométrie à savoir l'identification palmaire. En effet, le choix des images palmaires n'est pas fortuit car d'une part ces images ne nécessitent pas un système complexe pour leurs acquisitions (matériel bon marché), et d'autre part, elles offrent une vaste zone pour l'extraction des caractéristiques, à savoir la paume de la main. Par ailleurs, cette modalité offre un grand éventail de caractéristiques telles que les lignes principales, les rides, les points de minuties, les points singuliers, la texture, etc..

Nous allons détailler dans ce qui suit nos travaux dans lesquels nous soulignerons l'efficacité et la robustesse des systèmes d'identification/vérification de personnes proposés, lesquels sont basés sur les images palmaires. Ces travaux s'articulent autour de l'utilisation des transformations géométriques pour l'extraction de caractéristiques discriminatoires. Dans un premier temps, nous explorerons l'utilisation de la méthode de Radon pour la caractérisation des lignes principales de la main, d'abord ces dernières sont projetées dans le domaine de radon afin d'obtenir un ensemble de points discriminants. Le processus de vérification est ensuite conduit en utilisant une mesure de similarité qui utilise un ensemble de points pour mesurer la similarité /désimilarité, à savoir la distance de Hausdorff combinée à la triangulation de Delaunay. Dans la seconde partie de ce chapitre nous exposerons nos travaux sur l'utilisation de la transformée Ridgelets pour la caractérisation des images palmaires multispectrales. A cet effet, nous avons proposé d'utiliser la fusion d'informations pour l'amélioration de l'efficacité et la robustesse du système biométrique proposé. Effectivement, nous avons exploré deux niveaux de fusion à savoir : la fusion au niveau pixel en utilisant la transformée d'ondelettes discrètes, et la fusion des

caractéristiques via la transformation Ridgelets des différentes bandes des images palmaires acquises. Le processus d'identification/vérification pour cette approche a été développé en utilisant deux méthodes de classification à savoir les k-plus proches voisins, et les machines à vecteurs supports (SVMs). Ce choix n'est pas fortuit, car partant du principe que les caractéristiques extraites sont discriminatoires, nous avons jugé que même l'utilisation d'une méthode des plus simples pourrait donner de bons résultats, pour ensuite expérimenter une méthode très élaborée à savoir les SVMs. Les lecteurs pourront noter que les résultats obtenus et publiés de ces travaux sont très prometteurs.

5.2. Etat de l'art de l'identification/vérification des images palmaires

Cette section est consacrée à l'étude de l'art relative aux processus d'identification/vérification des empreintes palmaires. La texture et les lignes palmaires sont les caractéristiques les plus observables en basse résolution. Pour cette raison elles ont été les plus utilisées dans les applications de vérification. Une multitude de nouvelles approches ont été proposées dans la littérature pour l'identification et /ou la vérification basées sur l'empreinte palmaire. Généralement, ces approches sont basées sur la texture, les lignes, l'apparence et l'orientation de la main, voire les représentations multiples des empreintes. Il a été établi que les approches qui exploitent la texture de l'empreinte donnent de bons résultats. Zang et al. [105], ont introduit l'algorithme *Palmcode* qui utilise les filtres de Gabor 2D pour extraire différentes textures spatio-temporelles de l'empreinte palmaire. Dans [106], les auteurs ont proposé un algorithme basé sur la texture pour la reconnaissance d'empreintes palmaires en haute et basse résolutions. Les auteurs ont combiné la transformation à base d'ondelettes de Gabor 2D, et les réseaux de neurones couplés (*PCNN: Pulse Coupled Neural Network*). Dans cette approche les images ont été d'abord normalisées selon l'orientation, la position et l'illumination, pour ensuite être décomposées en cinq échelles et huit orientations par les filtres de Gabor. Finalement, l'algorithme PCNN a été utilisé pour extraire des caractéristiques basées sur l'entropie pour les images binaires. Les lignes principales ont suscité un grand intérêt de la communauté scientifique, ceci a été essentiellement dû à leurs simplicité, et les bonnes performances qui en ont résulté. Dans cette optique Huang et al. [107] ont introduit la transformée de Radon finie modifiée pour extraire les lignes principales. Zhang et al. [108] ont utilisé un modèle en ondelettes basé sur la modélisation et l'expansion d'un contexte directionnel, afin d'extraire les lignes principales. Par ailleurs, Han et al. [109] ont proposé des opérations de Sobel et morphologique dans ce même but. La plupart des techniques récentes relatives à la reconnaissance de l'empreinte palmaire sont

basées sur la fusion de plusieurs représentations de cette modalité, le but étant d'accroître les performances et la robustesse des systèmes biométriques. Ces algorithmes exploitent essentiellement trois niveaux de fusion, à savoir : la fusion au niveau pixel, celle des caractéristiques et finalement la fusion au niveau du processus de décision (fusion des résultats de classification). Kumar et Zhang [110] ont utilisé la transformation de Gabor, et deux modalités: les lignes et l'apparence de la main. Afin d'améliorer les résultats de l'authentification, ils ont effectué la fusion des scores au niveau décisionnel. Kong et Zhang [111] ont proposé une fusion des caractéristiques, à cet effet ils ont utilisé plusieurs filtres de Gabor elliptiques. Les auteurs ont baptisé cette règle de fusion *Fusioncode*. Par ailleurs, Jia et al. [112] ont proposé deux schémas d'identification d'individus à partir de l'empreinte palmaire, basés sur une stratégie de fusion. En effet, après extraction des lignes principales et appariement (*matching*) les auteurs ont sélectionné le plus grand score, et ont effectué une fusion au niveau décisionnel en combinant le score obtenu à partir des lignes et celui basé sur la préservation des projections locales (*LPP: Locality Preserving Projection*) pour l'identification proprement dite. Dans [113] les auteurs ont introduit la fusion des scores en utilisant la règle de la somme pondérée. Dans cette approche chaque image d'empreinte palmaire a été divisée en blocs superposés, dans le but d'extraire les informations issues de la différence de phases, et de leur utilisation comme caractéristiques binaires. La phase de décision consistait alors à fusionner les scores d'appariement des différents blocs de l'image. Hong et al [114] ont combiné les représentations multi-échelles et multi-directions des transformées de Gabor et Curvelets pour l'amélioration de la boîte de comptage différentiel (*IDBC: Improved Differential Box Counting*), ceci afin d'extraire des caractéristiques d'empreintes palmaires. Ces dernières années ont vu l'apparition de techniques 3D pour l'authentification biométrique, telle que "3D Palmprint" proposée par Zhang et al [115] où des caractéristiques globales ont été extraites pour décrire la forme de la main.

Dans certaines applications, les systèmes biométriques basés sur les empreintes palmaires qui agissent seulement dans la lumière visible sont souvent inappropriés pour répondre aux exigences en termes de performances [116-121]. Pour résoudre ce problème, de plus en plus de chercheurs se sont intéressés aux images des veines palmaires, qui nécessitent une acquisition dans le proche spectre infrarouge (NIR) [122-129]. Cette technique d'acquisition est une partie de la biométrie multispectrale, considérée comme une approche émergente, efficace et prometteuse. Conformément à la théorie électromagnétique, les ondes hertziennes (à partir du spectre de la lumière visible au proche infrarouge) permettent d'extraire des informations de profondeur à partir d'un objet d'intérêt. Dans ce cas, l'éclairage multispectral de l'empreinte

palmaire permet une visualisation du tissu sous-cutané, en utilisant différentes longueurs d'ondes qui offrent plusieurs types d'informations. Ces informations peuvent être traitées en utilisant une approche de fusion qui consiste à combiner les données acquises à différentes longueurs d'ondes, de sorte à augmenter la précision et la robustesse du système biométrique. L'information représentée par de multiples modalités peut être fusionnée à différents niveaux [123,124,126-129], dont le niveau d'extraction de caractéristiques, le niveau du score ou le niveau de décision. Pour la reconnaissance biométrique, la tâche la plus importante est d'extraire des caractéristiques distinctives et discriminantes du modèle biométrique choisi, que ce soit dans le domaine spatial, qui utilise les mesures statistiques des données spatiales, ou dans le domaine fréquentiel en utilisant une transformation. Dans notre approche, l'algorithme d'extraction de caractéristiques est basé sur l'analyse statistique et l'énergie des coefficients ridgelets avec une implémentation facile et une faible complexité dans les applications pratiques.

Une étude comparative de la technique de fusion au niveau pixel des images palmaires multispectrales basée sur plusieurs critères bien établis, pour l'évaluation objective de la qualité de la fusion, a été donnée en détail par Hao et al. dans [122, 128]. Dans [129], les auteurs ont comparé les performances de reconnaissance d'empreintes palmaires dans différentes longueurs d'ondes spectrales y compris le spectre rouge, vert, bleu et infrarouge. Ils ont présenté un système de reconnaissance d'empreintes palmaires multispectrales basé sur la fusion d'images en utilisant la transformée en ondelettes. Dans [123], un nouveau système de reconnaissance biométrique a été proposé, son idée principale était la fusion des images d'empreintes palmaires et des images palmaires des veines en utilisant la préservation des contours, et une amélioration du contraste en utilisant les ondelettes; ce concept est appelée méthode *Laplacianpalm*. Dans [105], Zhang et al. ont proposé une vérification en ligne conjointe des empreintes palmaires et des veines palmaires, qui comprend la détection d'activité (*liveness detection*), le codage de la texture et la fusion des scores, calculée par une somme pondérée dynamique (*dynamic weighted sum*). Dans [126,127], les auteurs ont proposé un système d'empreintes palmaires multispectrales à faible coût qui peut fonctionner en temps réel, et offrir un système de haute qualité. Dans [130], Cui a présenté une stratégie de fusion basée sur un modèle étendu des images couleurs généralement discriminantes (*GCID : General Color Image Discriminate*), le but étant de générer trois nouvelles composantes de couleurs pour l'amélioration de la performance de la reconnaissance. Les auteurs dans [131] ont proposé un schéma de fusion multispectrale basé sur la mesure de la dépendance; ils ont effectué une analyse de dépendance statistique entre ces sous-bandes pour réaliser la fusion, soit par choix ou par fusion pondérée.

5.3. Contributions

5.3.1. La transformée de Radon pour la caractérisation des images palmaires

Dans cette première contribution une nouvelle approche pour caractériser les images palmaires est proposée. Les images sont projetées dans le domaine de Radon qui a les spécificités (1) de transformer des lignes parfaites en points, (2) d'être robuste, c'est-à-dire non sensible aux bruits [15, 132-137]. Même si les lignes de la main ne sont pas rectilignes, la transformée de Radon (TR) permet une concentration d'énergie locale qui peut être utilisée pour encoder les positions de ces lignes. Afin de définir une structure topologique pour chaque individu, nous proposons de connecter les principaux points extraits par la TR en utilisant la triangulation de Delaunay. Dans le reste de cette section nous discuterons l'approche proposée dans laquelle nous détaillerons d'abord l'extraction des Régions D'Intérêt (ROI: *Region Of Interest*) et des lignes principales. Ensuite, nous développerons la technique de la triangulation de Delaunay pour l'extraction d'une signature de la main. Nous présenterons entre autres les résultats de l'expérimentation afin de prouver l'efficacité de notre approche.

5.3.1.1. Approche proposée

Dans cette section nous présenterons les différentes étapes proposées pour la caractérisation de l'empreinte palmaire à savoir : extraction du ROI, extraction des lignes principales et finalement l'étape de traitement dans le domaine de Radon, et le processus de vérification préconisé. La figure 5.1 illustre le diagramme bloc de l'algorithme proposé. Il est à noter que dans notre approche les valeurs maximales des coefficients de Radon sont utilisées pour caractériser les primitives spécifiques telles que, les lignes droites ou les arcs coniques.

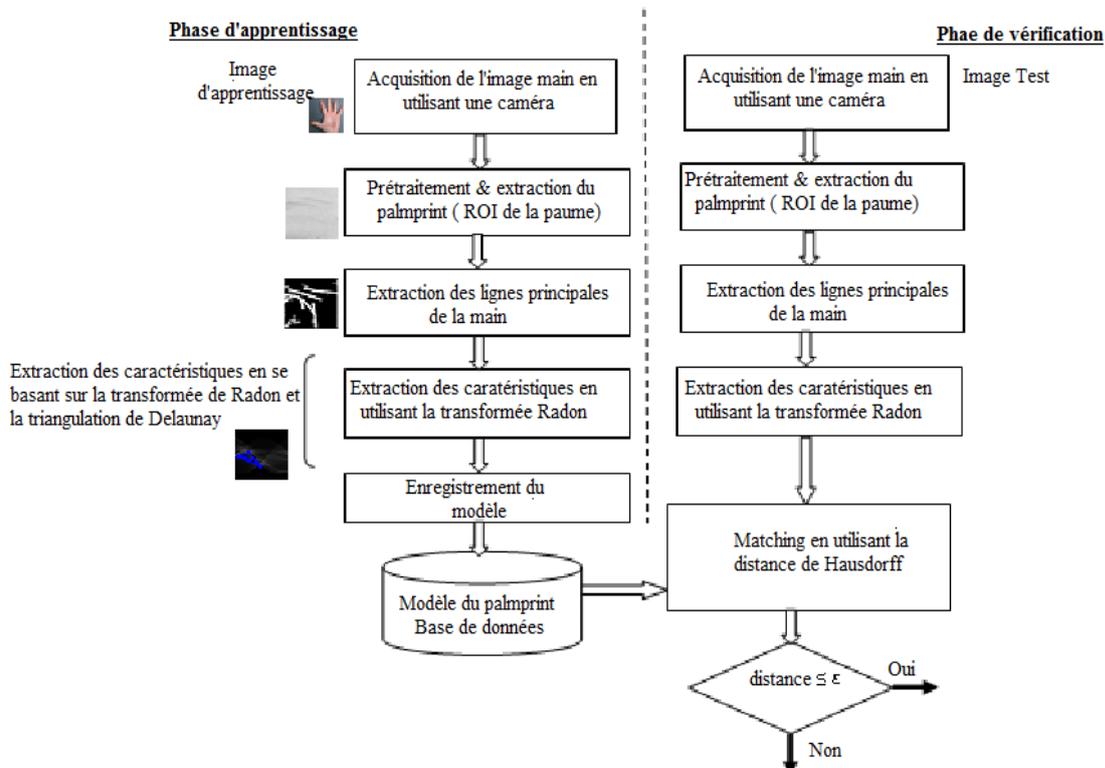


Figure 5.1. Diagramme bloc de l'algorithme proposé pour la vérification des images palmaires basée sur la transformée de Radon et la triangulation de Delaunay.

5.3.1.2. Extraction des régions d'intérêt

Afin d'extraire les régions d'intérêt des images palmaires (ROIs), il est nécessaire de procéder au préalable à un prétraitement de ces images. Cette phase de prétraitement englobe la segmentation des images de la main, et la détermination des points clés (voir figure 5.2). Extraites de la littérature, plusieurs méthodes peuvent être utilisées pour la segmentation des images de la main, ceci inclus les méthodes basées sur l'histogramme, l'entropie, la segmentation du niveau de gris et finalement, celles basées sur la cohérence spatiale et l'adaptation locale. Pour extraire la région d'intérêt deux approches peuvent être utilisées : la première permet d'extraire une région carrée, tandis que la deuxième permet d'extraire une région circulaire des images de la main.

Dans ce qui suit les régions d'intérêt sont extraites conformément aux étapes suivantes [15, 112, 117, 119,120]:

- *Etape 1* : d'abord un seuillage simple est utilisé pour séparer la région de la main du fond. A cet effet, un seuil γ est nécessaire pour convertir l'image en niveaux de gris en une image binaire (voir figure 5.2 b).
- *Etape 2* : Le contour de la forme de la main est obtenu en utilisant un algorithme de traçage des frontières, en l'occurrence *Neighborhood Border Tracing Algorithm* (voir figure 5.2 e),

- *Etape 3* : Un algorithme de détection des vallées des doigts est appliqué à l'image issue de l'étape 2 (voir figure 5.2 f), il s'agit de l'algorithme *Salien-point Detection*.
- *Etape 4* : Les vallées des doigts obtenues de l'étape 3 serviront comme points principaux pour la détection de la région d'intérêt (voir figure 5.2 f). Afin d'extraire les coordonnées des points, il est nécessaire de définir un système de coordonnées invariant par rapport à la translation et la rotation de la position de la paume de la main. Ceci permet l'extraction de la même surface à étudier quelque soit la position de la main.

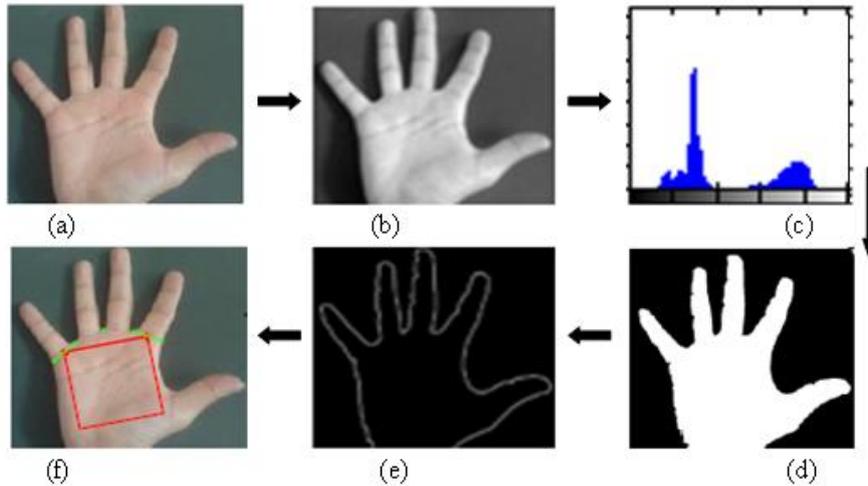


Figure 5.2. Extraction du ROI, (a) : image originale ; (b): image en niveaux de gris; (c): histogramme de l' image; (d): binarisation de l' image; (e): extraction du contour de l'image (d) et (f):Extraction du ROI

5.3.1.3. Extraction des lignes et caractérisation dans le domaine de Radon

5.3.1.3.1. Généralités sur la transformation de Radon

La transformée de Radon consiste à projeter l'image sur un certain nombre d'orientations en intégrant l'image le long de la direction orthogonale à la projection [[15, 132-137]. Soit $R_f(\beta, \theta)$ la transformée de Radon d'une fonction $f(x, y)$ continue et à support compact dans \mathbb{R}^2 ; définie par:

$$R_f(\beta, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x \cos(\theta) + y \sin(\theta) - \beta) dx dy \quad (5.1)$$

où $\theta \in [0, \pi]$ et $\beta \in [-\infty + \infty]$, $\delta(\cdot)$ représente la fonction de Dirac définie par:

$\int_{-\infty}^{+\infty} \delta(t) dt = 1$ et $\delta(t) = 0$ si $t \neq 0$. Si on considère une ligne rectiligne définie par :

$$L(x, y) = \delta(\beta^* - x \cos(\theta^*) - y \sin(\theta^*)). \quad (5.2)$$

La transformée de Radon de la fonction $L(x, y)$ est exprimée par :

$$\begin{aligned}
 R_L(\beta, \theta) &= 0; & \beta &\neq \beta^* \text{ ou } \theta \neq \theta^* \\
 R_L(\beta, \theta) &= \int \delta(0) d\theta; & \beta &= \beta^* \text{ et } \theta = \theta^*
 \end{aligned}
 \tag{5.3}$$

Sachant que dans l'espace de Radon, l'image d'une ligne paramétrée par β^*, θ^* est égale à une Dirac localisée au point $[\beta^*, \theta^*]$ (voir figure 5.3), on peut estimer les descripteurs ponctuels des lignes présents dans une image.

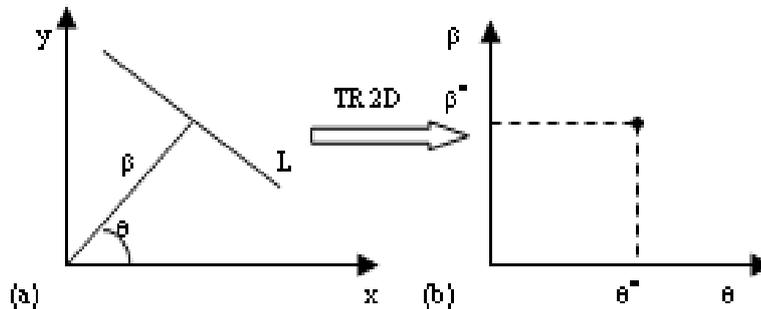


Figure 5.3. Représentation d'une ligne rectiligne ; (a) dans le domaine spatial et (b) dans le domaine de Radon

La transformée de Radon possède des propriétés intéressantes; dont certaines sont pertinentes pour la représentation de forme dans l'image à savoir:

- **La Translation**

Si on suppose qu'une la fonction $f(x, y)$ est déplacée d'une distance (x_0, y_0) ; alors la TR de cette fonction est :

$$R(\beta, \theta) [f(x - x_0, y - y_0)] = R_f(\beta - \beta_0, \theta) \tag{5.4}$$

Avec $\beta_0 = x_0 \cos \theta + y_0 \sin \theta$

- **La Rotation**

Une rotation de la l'image $f(x, y)$ d'un angle φ ; implique un décalage de la transformée de Radon en θ . L'expression de la TR est donnée par :

$$R(\beta, \theta) [f(x \cdot \cos \varphi + y \cdot \sin \varphi, -x \cdot \cos \varphi + y \cdot \sin \varphi)] = R_f(\beta, \theta + \varphi) \tag{5.5}$$

- **Le Changement d'échelle**

Un changement d'échelle de f implique un changement d'échelle dans les coordonnées β et l'amplitude de la transformée $R(\beta, \theta)$; ceci est exprimé alors par l'équation :

$$R(\beta, \theta) \left[f\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right) \right] = \lambda \cdot R_f\left(\frac{\beta}{\lambda}, \theta\right) \tag{5.6}$$

Où λ : le facteur d'échelle.

5.3.1.3.2. Extraction des lignes et de la signature de l'empreinte palmaire

De par leur unicité et stabilité, la localisation et la forme des lignes principales (telles que : la ligne du cœur, la ligne de tête et la ligne de vie.) de la paume de la main sont considérées comme des caractéristiques importantes pour l'identification des individus [15,120]. L'extraction des lignes principales de la paume requiert les étapes suivantes : l'application d'un masque moyen au ROI, suivi de l'application d'un masque pour la détection des lignes, et finalement la squelettisation des lignes.

Afin de caractériser l'image ROI $f(x, y)$ de l'empreinte palmaire dans le domaine de Radon, il est particulièrement intéressant de considérer seulement les lignes principales sans les informations aberrantes (communément appelées *outliers* dans la littérature). Formellement, cette considération est exprimée par [15,134,137] :

$$f_L(x, y) = \begin{cases} 1 & \text{si } (x, y) \in L \\ 0 & \text{ailleurs} \end{cases} \quad (5.7)$$

où L représente chacune des lignes principales de la main.

Basé sur le fait que la projection des lignes sur différents angles contient des informations cruciales pouvant être perçues visuellement, il nous a paru intéressant d'explorer cette propriété pour caractériser la structure de l'image palmaire. Effectivement, les formes géométriques telles que les lignes droites ou courbées pourront être représentées explicitement par la transformée de Radon sous forme d'énergies concentrées en quelques coefficients de grandes valeurs dans le domaine de la transformée. Ces coefficients dont les valeurs sont maximales permettront de former une signature de l'empreinte palmaire. Formellement, l'extraction de ces valeurs revient à résoudre le problème d'optimisation suivant :

$$S_R(\beta, \theta_i) = \arg \max_{\beta} \{ |R_f(\beta, \theta_i)| \} \quad (5.8)$$

Une fois les maximums détectés pour chaque orientation $\theta_i \in [0, \pi]$, l'application d'un seuillage approprié pour chaque classe d'images d'empreintes palmaires (une classe représente plusieurs acquisitions d'images palmaires d'un même individu) permettra de n'extraire que les valeurs significatives conformément à l'équation suivante :

$$M_{threshold}(\beta, \theta_i) = S_R(\beta, \theta_i) > Threshold \quad (5.9)$$

Où le calcul du seuil *Threshold* est basé sur la valeur maximale et la moyenne des coefficients de Radon. Afin d'adapter les seuils à chaque classe d'empreintes palmaires, on a introduit un facteur empirique adaptatif α . L'équation suivante donne l'expression du seuil approprié [15] :

$$Threshold = \mu + \frac{(\max - \mu)}{\alpha} \quad (5.10)$$

α : facteur adaptatif, μ et \max : représentent respectivement la moyenne et le maximum des coefficients de Radon le long de chaque projection.

Après extraction, seuillage et binarisation des coefficients de la transformée de Radon, la représentation des empreintes palmaires se résumera à un ensemble d'objets distincts. Il est primordial de pouvoir mesurer la similarité entre les différentes structures obtenues dans la perspective d'une application de vérification des individus, à cet effet, nous avons exploité les attributs géométriques de ces structures, en ne retenant que le centre de gravité de chaque région dont la concentration en énergie est plus forte. Le reste des objets est éliminé car considéré comme des aberrances (outliers) .

3. Modélisation des caractéristiques géométriques de l'image palmaire en utilisant la triangulation Delaunay

Afin de produire une structure unique pour chaque empreinte palmaire, les points dominants extraits par la méthode décrite précédemment sont reliés les uns aux autres au moyen de la triangulation de Delaunay. Soit S l'ensemble de points principaux $s_1 s_2 s_3 \dots \dots s_n$, obtenus après seuillage et binarisation de la transformée de Radon de l'image. La triangulation de Delaunay de S est obtenue en calculant d'abord le diagramme de Voronoï, ensuite en reliant les centres de chaque paire de régions voisines de Voronoï. la triangulation de Delaunay est unique, elle permet entre autres de maximiser le plus petit angle de tous les triangles. Par conséquent, on peut obtenir une triangulation régulière [15, 134, 135, 137-140]. L'idée de base étant d'établir une structure stable des caractéristiques discriminantes, il est important de souligner que la signature géométrique n'est pas sensible aux rotations, ce qui signifie que les principales caractéristiques d'une classe donnée seront conservées.

Soit C l'ensemble des centres de gravité qui forment la signature de l'empreinte palmaire:

$$C = [c_1 c_2 c_3 \dots \dots c_n] \quad (5.11)$$

où $c_n = (x_n, y_n)$ représente les coordonnées du centroïde de l'objet dans l'image. Après extraction de tous les centroïdes dans le domaine de Radon, nous appliquons la triangulation de Delaunay en considérant ces centroïdes comme sommets de triangles. Chaque empreinte palmaire est alors représentée par un graphe connexe particulier, où chaque nœud constitue un point de la transformation de Radon et chaque segment relie deux centroïdes.

Afin de comparer deux structures obtenues par triangulation de Delaunay de deux images palmaires différentes, nous avons utilisé dans notre approche la distance de Hausdorff comme une métrique de mesure de similarité.

5.3.1.4. Mesure de similarité en utilisant la distance de Hausdorff

Afin de mesurer la similarité entre les images acquises et les images de la base d'apprentissage, nous avons utilisé la distance de Hausdorff. Ce choix est justifié par le fait que les mesures de similarité d'images basées sur des caractéristiques géométriques sont plus fiables que celles basées sur l'intensité [15, 141-145]. Par ailleurs, la distance de Hausdorff engendre une mesure entre deux ensembles distincts de points.

Soient deux ensembles finis de points $C = \{c_1, c_2 \dots c_p\}$ et $C' = \{c'_1, c'_2 \dots c'_p\}$. La distance de Hausdorff est la distance maximale entre ensemble de points et le point le plus proche d'un ensemble [15, 141-145]. Formellement, la distance Hausdorff entre les ensembles de points C et C' est la fonction maximum définie par:

$$H(C, C') = \max (h(C, C'), h(C', C)) \quad (5.12)$$

$$\text{où } h(C, C') = \max_{c \in C} \{ \min_{c' \in C'} \{ d(c, c') \} \},$$

avec c, c' : les centroïdes des objets dans le domaine de Radon et $d(c, c')$: la distance entre ces points. Dans un but de simplification, $d(c, c')$ est considéré comme étant la distance euclidienne entre c et c' . Par conséquent, on peut mesurer la tolérance de la triangulation de Delaunay de deux ensembles de points C et C' en se reposant sur la mesure de la distance de Hausdorff. Soient $DT(C)$ et $DT(C')$ les triangulations de Delaunay respectives de deux ensembles C et C' . La tolérance est définie comme la borne supérieure de toutes les distances telles que $\varepsilon \geq 0$, de manière que si chaque point c_i se déplace arbitrairement, la structure $DT(C)$ ne change pas. Cette dernière est définie par l'équation :

$$tol(DT(C)) = Sup \left\{ \begin{array}{l} \varepsilon \geq 0 \mid DT(C) \sim DT(C') \text{ pour tout } C' \\ \text{tel que } h(C, C') < \varepsilon \end{array} \right\} \quad (5.13)$$

où $DT(C) \sim DT(C')$ implique que c_i et c_j sont voisins dans la structure de Delaunay $DT(C)$ si et seulement si les points c'_i et c'_j correspondants sont aussi voisins dans la structure $DT(C')$.

5.3.1.5. Expérimentation et résultats

Afin d'évaluer l'approche proposée, une base d'images de mains acquises sans contact, du Laboratoire Signaux et Images (LISSI) de l'université de Paris Est Créteil, a été utilisée. Cette base est constituée de 200 images collectées sur 20 individus. Ces images ont été collectées à une distance de 20 à 50 cm à l'aide d'une caméra digitale ayant une résolution de 1280x960 pixels. A l'origine les images sont d'une résolution de 300x300, les parties centrales des images extraites ont une taille de 128x128. La figure 5.4 ci-dessous illustre le processus d'acquisition des images au sein du laboratoire.

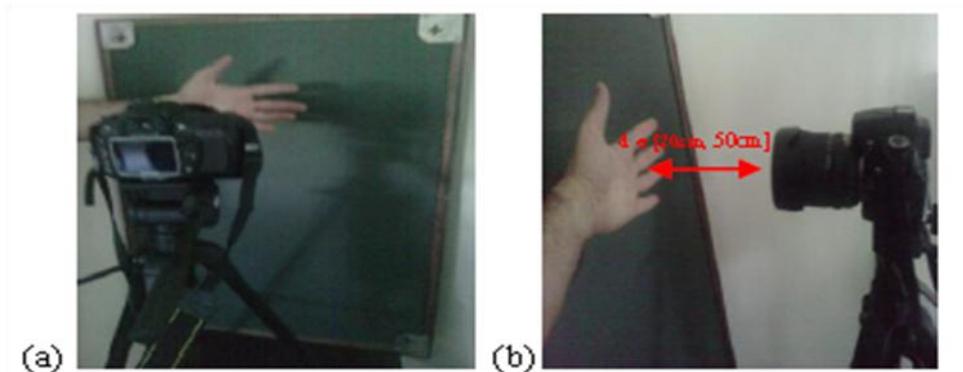


Figure 5.4. Système d'acquisition des images de la main, (a): acquisition d'une image palmaire; (b): la distance utilisée dans l'acquisition.

Une fois les régions d'intérêts des images de la main extraites, un algorithme pour la détection des lignes principales est appliqué, les figures 5.5(a) et 5.5(b) montre les empreintes palmaires et les lignes principales de plusieurs images de mains après normalisation de chaque acquisition. Les empreintes palmaires sont projetées dans le domaine de Radon, il est à noter que dans l'espace de la transformation de Radon quelques coefficients sont plus significatifs que les autres (voir figure 5.5 (c)).

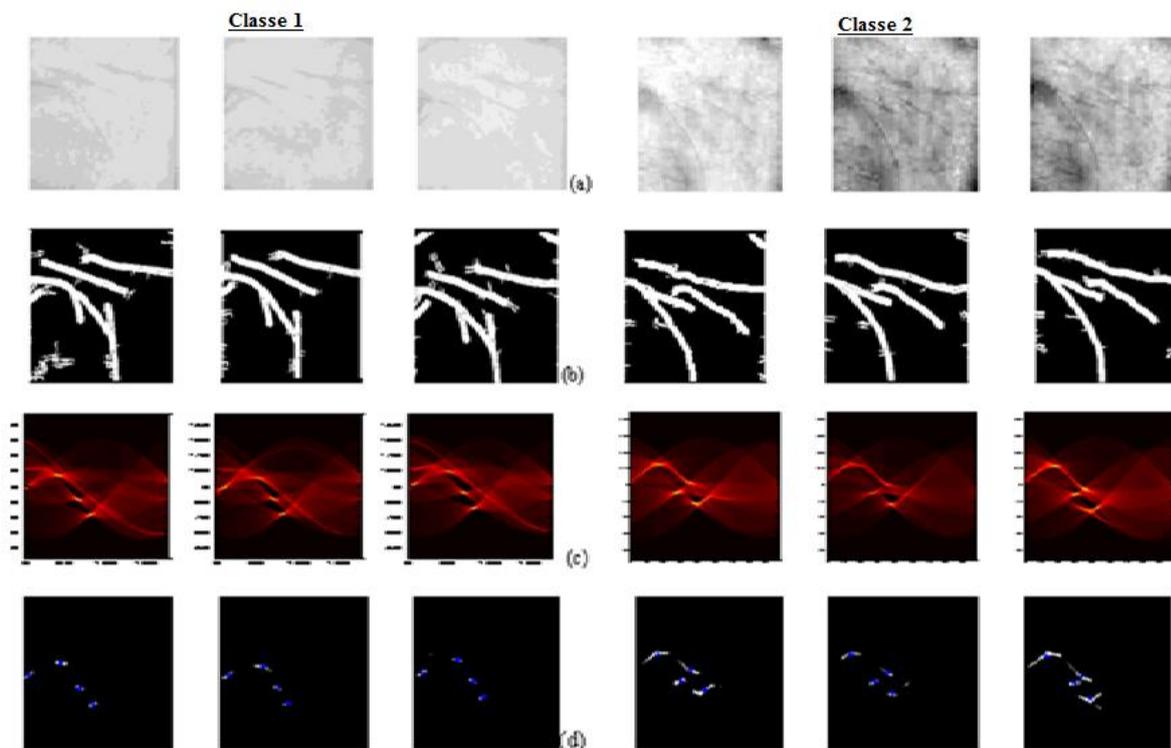


Figure 5.5. Transformation de Radon des lignes principales dans le ROI pour 2 classes d'images palmaires avec plusieurs acquisitions; (a) : images palmaires; (b) : les lignes principales extraites; (c): Transformation de Radon et (d): Extraction des centroïdes utilisant le seuillage et la binarisation des coefficients de Radon

La représentation de l'histogramme des coefficients de Radon selon plusieurs projections (voir figure 5.6), montre qu'il ya un maximum qui correspond à une grande concentration d'énergie pour chaque angle de projection.

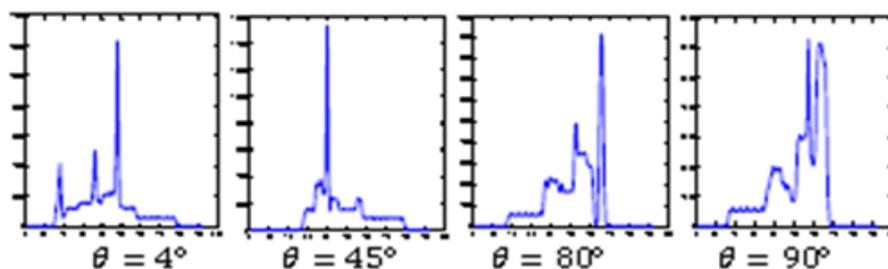


Figure 5.6. Histogramme des coefficients de Radon pour différentes projections ($\theta = 4, 45, 80$ et 90°)

Afin de ne conserver que les coefficients maximaux de la transformation de Radon pour chaque projection, nous avons appliqué un seuillage adaptatif à chaque classe d'images palmaires. Comme résultat nous avons obtenu une représentation simple d'un ensemble d'objets distincts qui peuvent être caractérisés par un ensemble de caractéristiques. Nous avons réitéré la simulation sur plusieurs classes différentes (plusieurs individus différents) avec plusieurs acquisitions, à cet effet, il nous a été donné de constater visuellement que pour un seuil approprié

pour chaque classe d'images palmaires, la signature obtenue par la triangulation de Delaunay est stable pour plusieurs acquisitions, et qu'elle est spécifique pour chaque empreinte palmaire (voir figure 5.7).

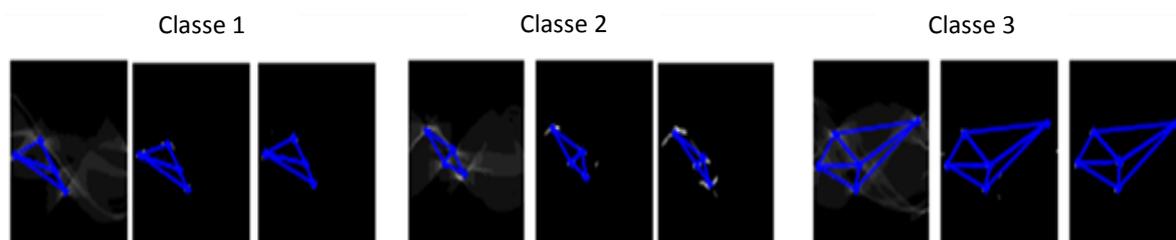


Figure 5.7. Exemples de signatures obtenues par triangulation sur trois individus

Il est avéré que la rotation de l'image palmaire d'un angle donné engendre un décalage de la transformation de Radon d'une quantité proportionnelle à cette rotation. Nos expérimentations ont montré que la caractérisation des empreintes palmaires en utilisant les centroïdes et la triangulation de Delaunay reste toujours stable (voir figure 5.8).

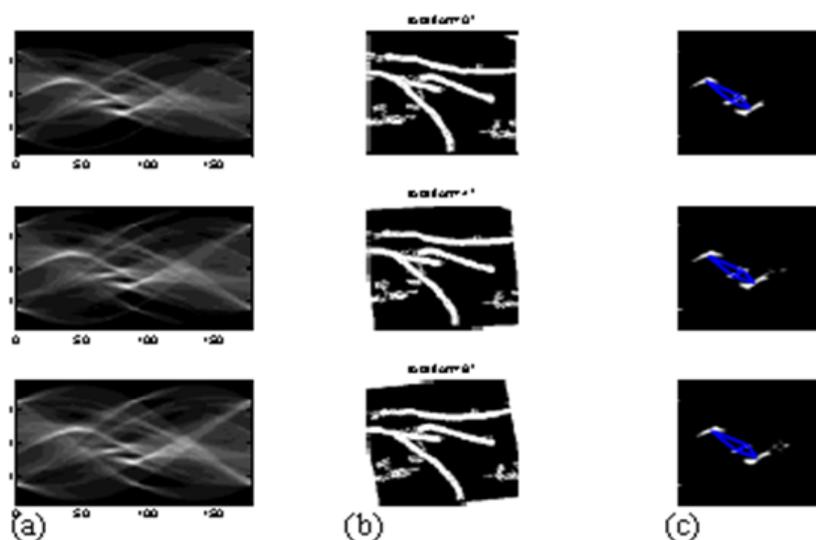


Figure 5.8. Invariance de la transformation à la rotation, (a): Coefficients de Radon ; (b) : Rotation du ROI de $\theta = 4^\circ$ et 8° et (c): Extraction des caractéristiques et triangulation de Delaunay

Afin de mesurer objectivement la similarité entre deux signatures, nous avons calculé la distance de Hausdorff entre des signatures d'images nouvellement acquises et des images références de la base d'images palmaires. On a noté que plus cette distance est petite, plus grande sera la similarité entre les deux images. Le tableau 5.1 illustre les résultats obtenus sur plusieurs images palmaires montrant que chaque signature est unique et stable. Idéalement, on peut considérer que l'image acquise correspond à l'image de référence si la distance de Hausdorff est nulle. Cependant, en pratique cette distance doit être inférieure à un seuil ε donné.

| | | | | | | | | | | | | |
|--|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | | | | | | | | | | | |
| | 0 | 7.45 | 10.33 | 8.67 | 11.18 | 38.51 | 42.97 | 35.59 | 33.13 | 41.79 | 39.69 | 38.99 |
| | | 0 | 13.19 | 16.30 | 12.15 | 33.16 | 38.13 | 29.96 | 32.63 | 37.76 | 37.04 | 32.78 |
| | 7.45 | 0 | 13.19 | 16.30 | 12.15 | 33.16 | 38.13 | 29.96 | 32.63 | 37.76 | 37.04 | 32.78 |
| | | 13.19 | 0 | 13.37 | 13.93 | 35.25 | 39.33 | 32.59 | 31.57 | 40.75 | 38.65 | 36.55 |
| | 10.33 | 13.19 | 0 | 13.37 | 13.93 | 35.25 | 39.33 | 32.59 | 31.57 | 40.75 | 38.65 | 36.55 |
| | | 16.30 | 13.37 | 0 | 10.41 | 31.42 | 34.90 | 27.70 | 29.99 | 36.44 | 35.63 | 31.12 |
| | 8.67 | 16.30 | 13.37 | 0 | 10.41 | 31.42 | 34.90 | 27.70 | 29.99 | 36.44 | 35.63 | 31.12 |
| | | 12.15 | 13.93 | 10.41 | 0 | 35.29 | 39.79 | 30.86 | 34.28 | 37.19 | 35.73 | 32.42 |
| | 11.18 | 12.15 | 13.93 | 10.41 | 0 | 35.29 | 39.79 | 30.86 | 34.28 | 37.19 | 35.73 | 32.42 |
| | | 33.16 | 35.25 | 31.42 | 35.29 | 0 | 14.01 | 5.93 | 7.89 | 22.25 | 21.08 | 14.54 |
| | 38.51 | 33.16 | 35.25 | 31.42 | 35.29 | 0 | 14.01 | 5.93 | 7.89 | 22.25 | 21.08 | 14.54 |
| | | 38.13 | 39.33 | 34.90 | 39.79 | 14.01 | 0 | 10.08 | 13.74 | 22.84 | 22.15 | 16.07 |
| | 42.97 | 38.13 | 39.33 | 34.90 | 39.79 | 14.01 | 0 | 10.08 | 13.74 | 22.84 | 22.15 | 16.07 |
| | | 29.96 | 32.59 | 27.70 | 30.86 | 5.93 | 10.08 | 0 | 8.49 | 22.57 | 21.33 | 14.74 |
| | 35.59 | 29.96 | 32.59 | 27.70 | 30.86 | 5.93 | 10.08 | 0 | 8.49 | 22.57 | 21.33 | 14.74 |
| | | 32.63 | 31.57 | 29.99 | 34.28 | 7.89 | 13.74 | 8.49 | 0 | 28.33 | 27.14 | 20.88 |
| | 33.13 | 32.63 | 31.57 | 29.99 | 34.28 | 7.89 | 13.74 | 8.49 | 0 | 28.33 | 27.14 | 20.88 |
| | | 37.76 | 40.75 | 36.44 | 37.19 | 22.25 | 22.84 | 22.57 | 28.33 | 0 | 3.07 | 16.29 |
| | 41.79 | 37.76 | 40.75 | 36.44 | 37.19 | 22.25 | 22.84 | 22.57 | 28.33 | 0 | 3.07 | 16.29 |
| | | 37.04 | 38.65 | 35.63 | 35.73 | 21.08 | 22.15 | 21.33 | 27.14 | 3.07 | 0 | 16.8 |
| | 39.69 | 37.04 | 38.65 | 35.63 | 35.73 | 21.08 | 22.15 | 21.33 | 27.14 | 3.07 | 0 | 16.8 |
| | | 32.78 | 36.55 | 31.12 | 32.42 | 14.54 | 16.07 | 14.74 | 20.88 | 16.29 | 16.78 | 0 |
| | 38.99 | 32.78 | 36.55 | 31.12 | 32.42 | 14.54 | 16.07 | 14.74 | 20.88 | 16.29 | 16.78 | 0 |

Tableau 5.1 Mesure de la similarité en utilisant la distance de Hausdorff pour différentes classes d'empreintes palmaires avec plusieurs acquisitions

Etude de la robustesse de la structure au bruit additif

Afin d'étudier la robustesse de notre approche vis-à-vis des bruits additifs, nous avons ajouté un bruit gaussien aux images des lignes principales extraites des empreintes palmaires acquises. A cet effet, nous avons considéré plusieurs SNRs (11.20 dB, 7.27 dB et 4.75 dB). La figure 5.9 illustre les résultats obtenus, et montre la robustesse de la transformation de Radon et la stabilité de la triangulation de Delaunay.

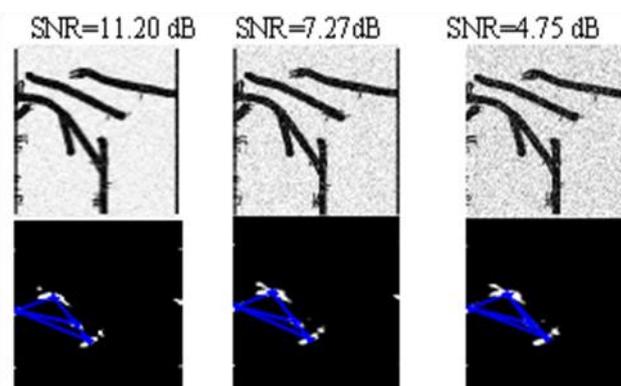


Figure 5.9. Triangulation de Delaunay des lignes principales d'images bruitées avec plusieurs SNRs

Par ailleurs, afin de confirmer la propriété de robustesse de l'approche proposée, nous avons calculé les distances de Hausdorff entre les triangulations de Delaunay des images originales

et celle des images bruitées. Les résultats du tableau 5.2 montrent que cette distance demeure négligeable.

| | | | | | |
|---------------------------|---|---|---|--|---|
| Image bruitée |  |  |  |  |  |
| Variance | 0.02 | 0.03 | 0.05 | 0.09 | 0.1 |
| SNR (dB) | 11.20 | 9.50 | 7.27 | 4.75 | 4.28 |
| Hausdorff distance | 1.78 | 1.83 | 1.95 | 4.08 | 4.98 |

Tableau 5.2 Mesure de similarité des images bruitées avec différents SNR (dB) en utilisant la distance de Hausdorff pour quelques images.

5.3.2. La transformée Ridgelets pour la caractérisation des images palmaires multispectrales

Dans cette deuxième contribution nous nous proposons de caractériser les images palmaires multispectrales en utilisant la transformée Ridgelets. Par ailleurs, nous nous intéressons à la fusion de données dont l'utilité a déjà été prouvée. En effet, la demande croissante pour le développement de systèmes biométriques avec des taux d'erreur moins significatifs a encouragé les scientifiques à s'intéresser de plus en plus aux images multispectrales, dans le but de surmonter les limites des techniques utilisant des images palmaires acquises en lumière blanche. Le présent travail s'insère dans une optique d'amélioration de la performance des systèmes biométriques reposant sur les empreintes palmaires multispectrales. A cet effet, nous explorerons deux niveaux de fusion: la fusion au niveau du pixel, et celle des caractéristiques. Le premier est basé sur la règle de décision maximale, cette règle permet de combiner plusieurs informations discriminatoires de différentes bandes spectrales de la transformée en ondelettes des images multispectrales. Le deuxième permet de fusionner les caractéristiques extraites des différentes images. Pour des raisons de simplicité et de réduction de la complexité computationnelle, nous proposons dans cette contribution d'extraire des caractéristiques en utilisant l'analyse statistique, et la distribution d'énergie des coefficients de la transformée Ridgelets finie (FRIT). Après cette phase, nous proposons une classification robuste pour identifier/vérifier les individus en considérant les deux niveaux de fusion. L'efficacité de notre approche a été évaluée en employant plusieurs classificateurs pour le cas binaire (vérification) et pour le cas multi-classes (identification). Une expérimentation rigoureuse sur les bases d'empreintes palmaires CASIA et PolyU a conduit à un taux de précision respectif de 100% et 99.79%. Par ailleurs, une étude comparative a révélé que l'approche proposée surpasse ou au moins égale les performances des méthodes de l'état de l'art

pour la reconnaissance des images multispectrales des empreintes palmaires. Le schéma global de l'approche proposée est illustré dans la figure 5.10.

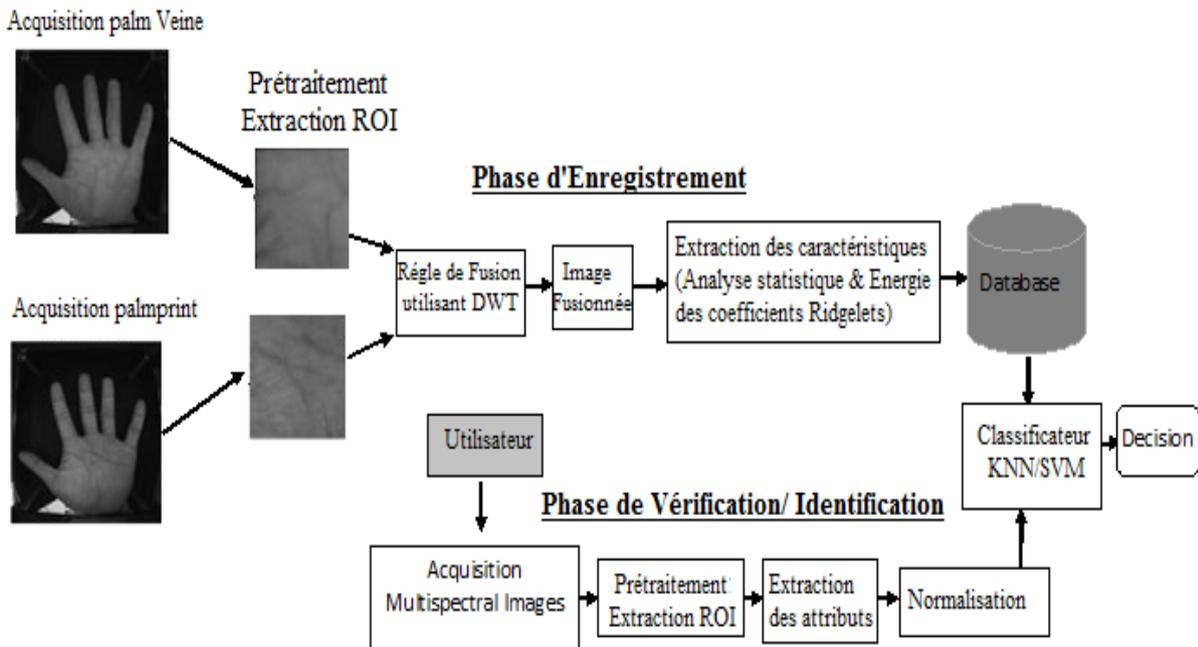


Figure 5.10. L'algorithme de l'approche proposée basée sur la fusion niveau pixel des images palmaires multispectrales en utilisant les classificateurs SVM /KNN

5.3.2.1. Fusion des images palmaires par les ondelettes

La fusion d'images a pour but l'intégration des informations complémentaires issues de plusieurs images acquises de sources différentes dans le but de l'obtention d'une meilleure description des images [16-18, 128-131, 146-148]. L'efficacité de ce concept requiert que les images obtenues par fusion possèdent toutes les informations pertinentes des images sources, sans introduction d'artefacts. Dans ce présent travail, nous nous focalisons sur deux aspects de la fusion d'images palmaires multispectrales. Dans un premier temps nous examinerons la fusion au niveau du pixel en utilisant une règle appropriée, à partir de laquelle chaque pixel est sélectionné parmi un ensemble de pixel, le but étant de combiner les coefficients d'ondelettes discrètes d'images de différentes sources spectrales. Par ailleurs, nous examinerons la fusion de caractéristiques dont le principe consiste en l'agrégation de plusieurs caractéristiques discriminatoires de plusieurs images en vue de parfaire la caractérisation.

La transformation en ondelette discrète (DWT pour Discrete Wavelets Transform) représente la transformée en ondelettes la plus utilisée pour la fusion d'images [16-18, 146-148]. Dans cette étude, l'approche de fusion d'images palmaires multispectrales proposée peut être implémentée en trois étapes. D'abord la fusion basée sur la DWT 2D décompose les images

sources à savoir : l'image de l'empreinte palmaire et l'image des veines de la paume en coefficients d'approximation et trois ensembles de détails représentant les directions horizontale, verticale et diagonale de l'image (voir figure 5.11 (a)). Pour un niveau de résolution J , toute image palmaire de chaque bande $f(t)$ peut être décomposée comme suit :

$$f(t) = \sum_k c_{Jk} \phi_{Jk}(t) + \sum_{j=1}^J \sum_n \sum_k d_{jk} \psi_{jk}^n(t) \quad (5.14)$$

avec: $\phi_{Jk}(t)$: la fonction échelle définie par $\phi_{Jk}(t) = 2^{J/2} \varphi(2^J t - k)$; $\psi_{jk}(t)$: fonction d'ondelettes définie par $\psi_{jk}^n(t) = 2^{j/2} \psi^n(2^j t - k)$; $n = 1, 2, 3$; c_{Jk} : les coefficients d'approximation et d_{jk} : les coefficients détails.

Ensuite, la méthode de fusion basée sur la DWT fusionne les coefficients complémentaires de la décomposition des deux bandes d'acquisition des images palmaires multispectrales, en utilisant la règle de sélection maximale (*maximum selection rule*) pour préserver des informations cruciales et pertinentes avec une réduction du bruit, et une meilleure représentation visuelle. La fusion par la règle maximale remplace les valeurs des petites magnitudes des coefficients d'ondelettes par de plus fortes valeurs, afin d'accentuer l'information importante comme bords. Soit, $WI_{i=1,\dots,6}$ la décomposition en ondelettes de chaque image palmaire enregistrée I_i . Les coefficients d'ondelettes de l'image fusionnée WI , sont calculés par :

$$WI_{fused}(x, y) = \begin{cases} WI_i(x, y) & \text{if } |WI_i|(x, y) \geq |WI_j|(x, y) \\ WI_j(x, y) & \text{if } |WI_j|(x, y) \geq |WI_i|(x, y) \end{cases} \quad i, j = 1, \dots, 6 \quad (5.15)$$

Finalement, l'image fusionnée I_{fused} est reconstruite en effectuant la transformée inverse en ondelettes discrète (DWT^{-1}) sur les coefficients d'ondelettes, on obtient alors :

$$I_{fused} = DWT^{-1}(WI_{fused}(x, y)) \quad (5.16)$$

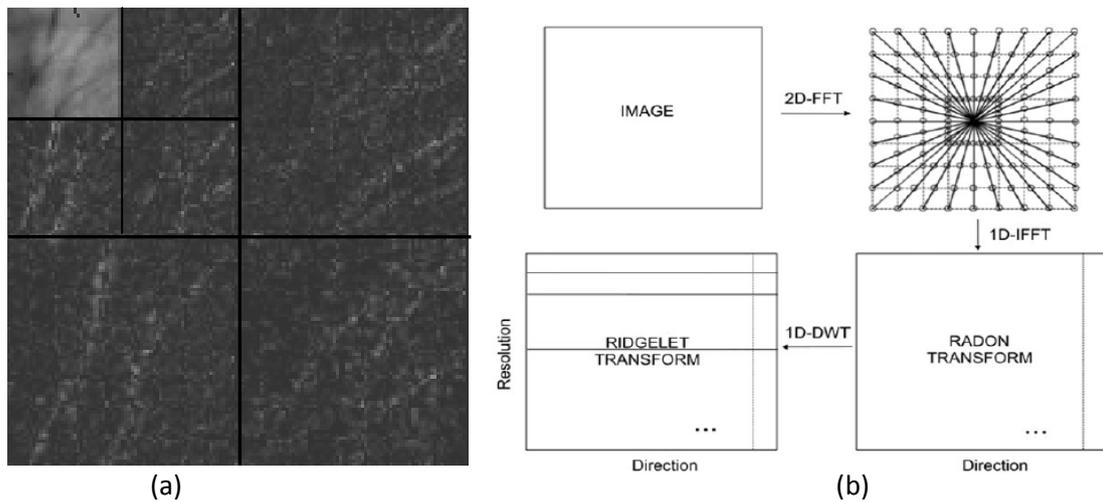


Figure 5.11. Représentation multiéchelle de l'empreinte palmaire, (a): décomposition DWT à un niveau égal à 2 et (b) : diagramme bloc de transformée en Ridgelets

Une fois l'image fusionnée obtenue, un pré-traitement des images de la main est effectué pour extraire la région d'intérêt (voir section 5.3.1.2). Une fois le ROI obtenu à partir de l'image palmaire multispectrale fusionnée, l'extraction des caractéristiques discriminantes est effectuée en utilisant une analyse statistique de la distribution de données de pixels et une analyse énergétique des coefficients de la transformée ridgelets finies. La transformée en Ridgelets discrète FRIT (voir figure 5.16 (b)) est obtenue par l'application de la transformée d'ondelette discrète (DWT) sur chaque projection de la transformée de Radon discrète.

5.3.2.2. Extraction des caractéristiques des images palmaires multispectrales

La représentation des caractéristiques doit contenir des informations cruciales, pertinentes et discriminantes qui peuvent être utilisées pour vérifier/identifier les individus. L'étape d'extraction est très importante pour la robustesse et la fiabilité d'un système biométrique. A cet effet, le vecteur de caractéristiques devrait être à peu près semblable dans la même classe (*c-à-d* la variance intra-classe devrait être faible pour les différents échantillons), et doit être sensiblement différente lorsque l'on considère des différentes classes (*i.e.* la variance interclasses doit être grande pour les échantillons de différentes classes) . Dans ce travail, afin de caractériser les images d'empreintes palmaires multispectrales, après la phase de fusion, des paramètres significatifs sont extraits. A cet effet, une analyse statistique utilisant les coefficients de Fisher est réalisée en considérant les énergies de coefficients ridgelets des images fusionnées dans chaque sous-bandes.

1. Extraction des caractéristiques basée sur l'analyse statistique

La fusion des empreintes palmaires et les images des veines, fournit une structure tout à fait aléatoire qui est souvent non homogène; une analyse statistique de la distribution des niveaux de gris sera donc nécessaire. Les paramètres descriptifs de chaque ROI sont calculés à partir de l'histogramme après la phase de fusion. Une tâche fondamentale dans de nombreuses analyses statistiques est la caractérisation de l'emplacement et de la variabilité de la distribution des données $x = [x_1, x_2, \dots, x_n]$. Les caractéristiques statistiques utilisées dans notre étude sont les coefficients de Fisher [149,150], ils représentent le 4th moment normalisé (le kurtosis ou l'aplatissement) et le 3^{ème} moment normalisé (skewness ou l'asymétrie). En outre, ces mesures sont calculées sur un bloc I d'image de dimension $N \times N$. Le Kurtosis est défini comme suit:

$$k = E\left(\left(\frac{x - \mu}{\sigma}\right)^4\right) = \frac{\frac{1}{N^2} \left(\sum_{i,j} (I(i,j) - \mu)^4 \right)}{\sigma^4} \quad (5.17)$$

où μ, σ représentent respectivement la moyenne et l'écart-type du bloc I de l'image.

Le coefficient d'asymétrie de Fisher (*skewness*) est une mesure de l'asymétrie de la distribution de probabilité d'une variable aléatoire à valeurs réelles, et il est défini par :

$$sk = E\left(\left(\frac{x - \mu}{\sigma}\right)^3\right) = \frac{\frac{1}{N^2} \left(\sum_{i,j} (I(i,j) - \mu)^3 \right)}{\sigma^3} \quad (5.18)$$

2. Extraction des caractéristiques basée sur l'analyse de l'énergie de FRIT

Afin d'extraire des caractéristiques discriminantes à partir de la représentation FRIT des images palmaires multispectrales, nous calculons le premier et le deuxième ordre de l'énergie (moments de premier et deuxième ordre) pour chaque carré des coefficients ridgelets de taille S_i , Les valeurs calculées seront contenues dans deux tableaux distincts E_1 et E_2 , respectivement.

$$E_1(i) = \frac{1}{S_i} \sum_{S_i} |frit(i)| \quad (5.19)$$

$$E_2(i) = \left(\frac{1}{S_i} \sum_{S_i} (frit(i))^2 \right)^{1/2} \quad (5.20)$$

Ces deux paramètres sont des paramètres discriminants [16-18, 151-158] pour chaque sous-bande de la ROI dans la même classe (intra-classe) et dans différentes classes (extra-classe). L'extraction de la signature de l'énergie des coefficients de ridgelets est résumé dans l'algorithme 1. La signature énergie des images multispectrales est définie comme suit:

$$E_s = [E_1(1), E_1(2), \dots, E_1(b), E_2(1), E_2(2), \dots, E_2(b)] \quad (5.21)$$

Le vecteur global de caractéristiques est définie par la combinaison des caractéristiques issues de l'analyse statistique, et ceux basées sur l'analyse de l'énergie des sous-bandes des coefficients FRIT. La combinaison des paramètres discriminants obtenus à partir des deux modalités est réalisée par la concaténation des paramètres de telle sorte qu'ils peuvent être tous les deux pris en compte dans l'étape de décision du système biométrique. Les techniques de réduction des caractéristiques, par exemple l'Analyse en Composantes Principales (ACP), peuvent être utilisées pour réduire le plus grand ensemble de fonctionnalités. L'algorithme suivant décrit l'estimation de la signature de l'énergie des coefficients de ridgelets.

Algorithme 1 : Estimation de la signature de l'énergie des coefficients de ridgelets pour des blocs non chevauchés (non overlapping blocks)

1. Initialization
 - Ω : base des images fusionnées;
 - P : taille du bloc (p est choisi comme étant un nombre premier);
 - b : nombre de blocs non chevauchés;
 - $E_s = []$;
2. Pour chaque image $I(N \times N) \in \Omega$
 - 2.1 $b_blocks \leftarrow$ Partition I en b blocs non chevauchés de taille (pxp);
 - 2.2 Pour chaque block $B_i \in b_blocks$ $i \in \{1 \dots b\}$
 - 2.2.1 calculer de la transformée ridgelet finie (FRIT);
 - 2.2.2 $b_Ridgelet \leftarrow$ Finite ridgelet de b blocs non chevauchés ;
 - fin
 - 2.3 Pour $i = 1: b$
 - 2.3.1 Calculer $E_1(i)$ du i^{th} block;
 - 2.3.2 Calculer $E_2(i)$ du i^{th} block;
 - 2.3.3 $E_s = [E_s, E_1(i), E_2(i)]$
 - fin

Fin

5.3.2.3. Résultats expérimentaux

Afin de valider les approches proposées, les expérimentations ont été conduites sur deux bases d'images palmaires multispectrales disponibles au public; et faisant autorité dans ce domaine à savoir : la base CASIA [157] (pour "Chinese Academy of Sciences Institute of Automation") et la base PoyU de l'institut technique de Hong Kong [158]. Les deux bases de données sont capturées dans différentes bandes spectrales; chaque bande spectrale présente les fonctionnalités spécifiques de la paume. Chaque échantillon de base de données est acquis pour chaque utilisateur lors deux séances d'acquisition séparée. La base de données CASIA, contient 7 200 images d'empreintes palmaires, capturées à partir de 100 personnes différentes en utilisant un dispositif d'imagerie spectrale multiples auto-conçu (*a self-designed multiple spectral imaging device*). Six images ont été acquises auprès de chaque utilisateur (trois images à chaque session) avec un intervalle minimum d'un mois. Les images palmaires sont codées sur 8 bits de niveaux de gris selon la norme JPEG. La base de données PolyU est collectée avec le dispositif sensible à la lumière à partir de 250 sujets, dont 55 femmes et 195 hommes. La répartition par âge est de 20 à 60 ans. Six images ont été acquises de son / sa paume gauche et droite dans chaque acquisition, avec un intervalle minimum de 9 jours, avec une résolution de l'image de 352x288 (<100dpi). Etant donné, que les ROIs de PolyU sont disponibles pour une utilisation simple, on s'est intéressé uniquement à l'extraction des ROIs pour les échantillons de la base de données CASIA, en utilisant la technique proposée dans la section 5.3.1.2.

1. Extraction des ROIs de la base CASIA et fusion des images palmaires

Après extraction des régions d'intérêt (ROIs) des images d'empreintes palmaires et celles des veines de la paume, vient la phase de la fusion basée sur la transformation en ondelette dont découle la région d'intérêt (ROI) de l'image fusionnée. Il est à noter que pour les besoins de l'extraction des régions d'intérêt (ROIs), nous avons sélectionné une zone centrale de la paume de la main qui satisfait les conditions d'invariance à la rotation et la translation de l'image palmaire. La figure 5.12 ci-dessous décrit quelques exemples de zones d'intérêt extraites sans rehaussement des échantillons de la base d'images palmaires CASIA. Par ailleurs, afin de proposer un système de reconnaissance d'images palmaires multispectrales performant, deux algorithmes de fusion ont été testés pour l'extraction du vecteur de caractéristique globale à savoir :

a. Algorithme de fusion au niveau pixel

Afin d'explorer les résultats de cet algorithme, nous avons utilisé la transformée en ondelettes discrète pour combiner les informations visuelles issues des images sources en

vu de l'obtention d'une image fusionnée plus informative, sans distorsion ni de perte d'information.

b. Algorithme de fusion des caractéristiques

Après extraction du vecteur de caractéristiques de chaque modalité séparément, nous avons combiné les deux vecteurs pour l'obtention d'un seul vecteur global.

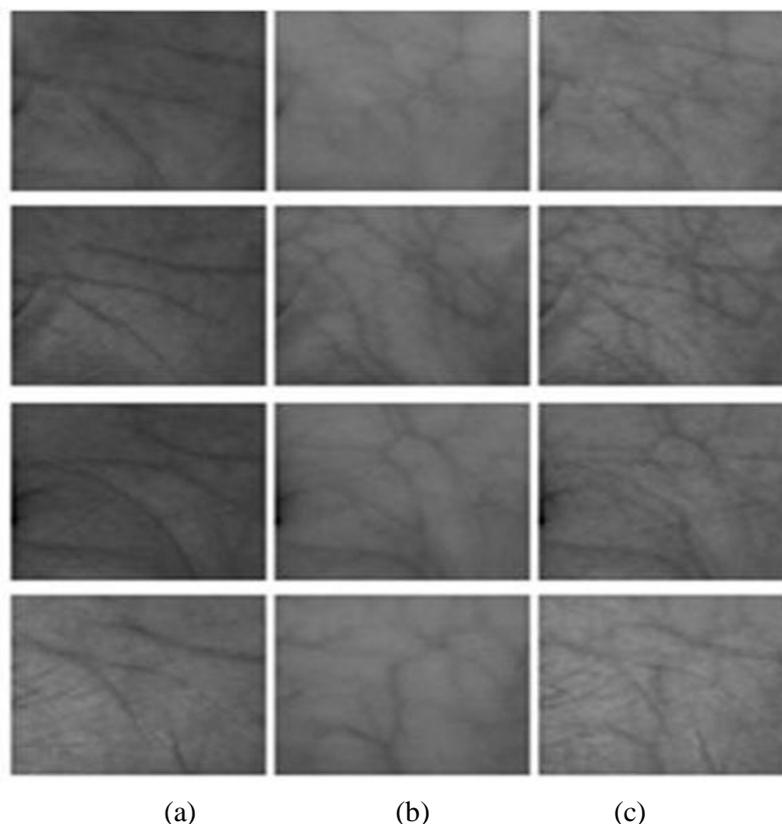


Figure: 5.12. Exemple des ROIs extraits dans la base de données, (a): ROI de l'empreinte palmaire, (b) :ROI de l'image palmaire des veines et (c): ROI des images fusionnées.

2. Extraction des caractéristiques de l'image palmaire multispectrale fusionnée

Avant d'appliquer la transformation FRIT à chaque bloc de l'image palmaire divisée, nous proposons de déterminer la distribution de niveau de gris des zones d'intérêt (ROIs) des images fusionnées intra-classe et extra-classes. Sachant que chaque individu représente une classe contenant six acquisitions différentes de son empreinte palmaire. Les zones d'intérêt de la base ont toutes la même taille et la même résolution. De l'histogramme de la distribution de niveaux de gris ; il en découle que plusieurs caractéristiques discriminatoires peuvent être extraites telles que le kurtosis et la skewness pour la caractérisation de l'empreinte palmaire. La figure 5.13 ci-dessous illustre la distribution et les valeurs obtenues pour trois classes différentes. Sachant que les zones d'intérêt des images palmaires multispectrales fusionnées sont divisées en 36 blocs qui ne se recouvrent pas , on applique la transformation FRIT à chaque bloc pour obtenir une matrice de

coefficients des 36 bandes. Après détermination de l'énergie basée sur les Ridgelets pour chaque bloc, 36 paramètres sont obtenus pour E_1 (ordre 1) et 36 paramètres pour E_2 (ordre 2). Il est à noter que ces caractéristiques présentent une faible variance pour les acquisitions de la même classe, spécifiquement pour les acquisitions effectuées dans une même session. Par contre, les caractéristiques extra-classes sont significativement éloignées.

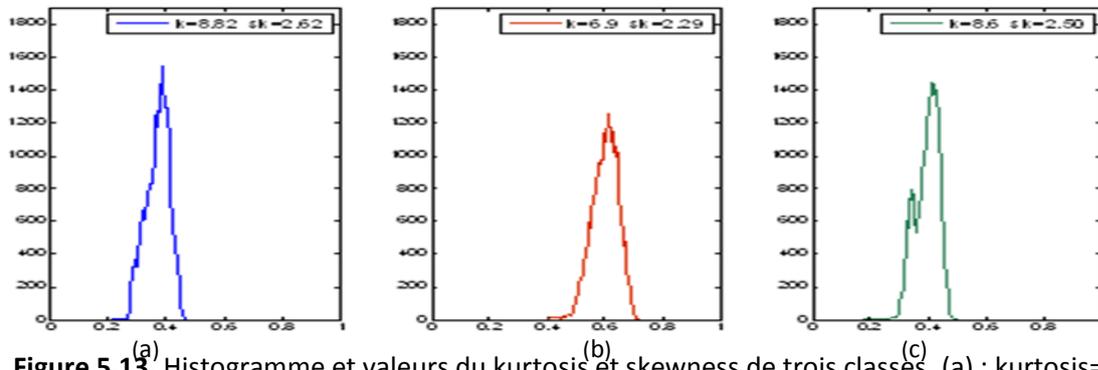


Figure 5.13. Histogramme et valeurs du kurtosis et skewness de trois classes, (a) : kurtosis=8.82, skewness =2.62; (b): kurtosis=6.9, skewness =2.29 et (c): kurtosis=8.6, skewness =2.50.

De ce qui précède il en découle que les caractéristiques sélectionnées ont un vrai pouvoir discriminatoire. La figure 5.14 (a) représente l'énergie par bloc de trois acquisitions correspondant à la même personne. Nous remarquons que quelque soit l'acquisition, l'écart entre les différents vecteurs caractéristiques est faible, ce qui confirme que les caractéristiques extraites en se basant sur la transformées en Ridgelets par bloc, permettent de caractériser efficacement un individu quelque soit la position et l'instant de la prise de l'acquisition. Par ailleurs, la figure 5.14 (b) met en exergue les coefficients d'énergies de Ridgelets de classes différentes (d'individus distincts). Dans ce cas, nous remarquons que l'écart entre les différents vecteurs caractéristiques est plus au moins important.

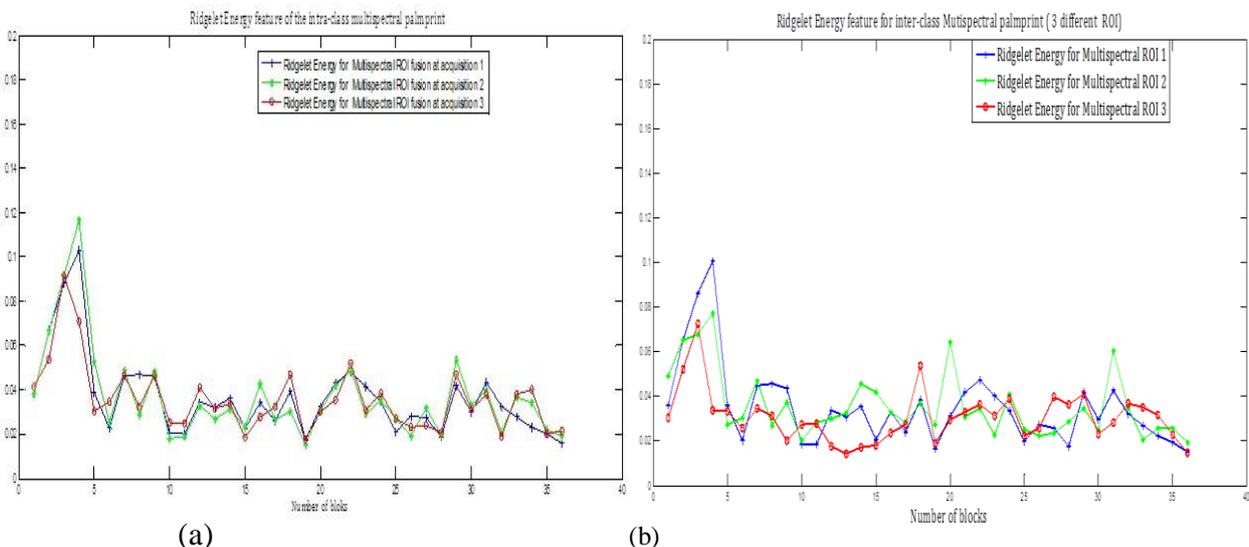


Figure.5.14. Coefficients des énergies des blocs pour ; (a): intra-classes (3 acquisitions de la même Multispectral ROI) et (b) : Inter-class (3 différentes multispectral ROI)

Afin d'évaluer le pouvoir discriminatoire des vecteurs de caractéristiques extraits, nous avons effectué une vérification (*matching*) en utilisant la distance euclidienne normalisée. La figure 5.15 ci-dessous illustre la courbe ROC (Receiver Operating Characteristic Curve) qui mesure la performance de l'algorithme proposé et la fusion des caractéristiques pour toutes les bandes validés sur la base d'images palmaires multispectrales PolyU. La courbe ROC illustre le compromis entre les mesures FAR (False Acceptance Rate) et FRR (False Rejection Rate). L'EER (Equal Error Rate) est la mesure prélevée lorsque les mesures FAR et FRR sont égales.

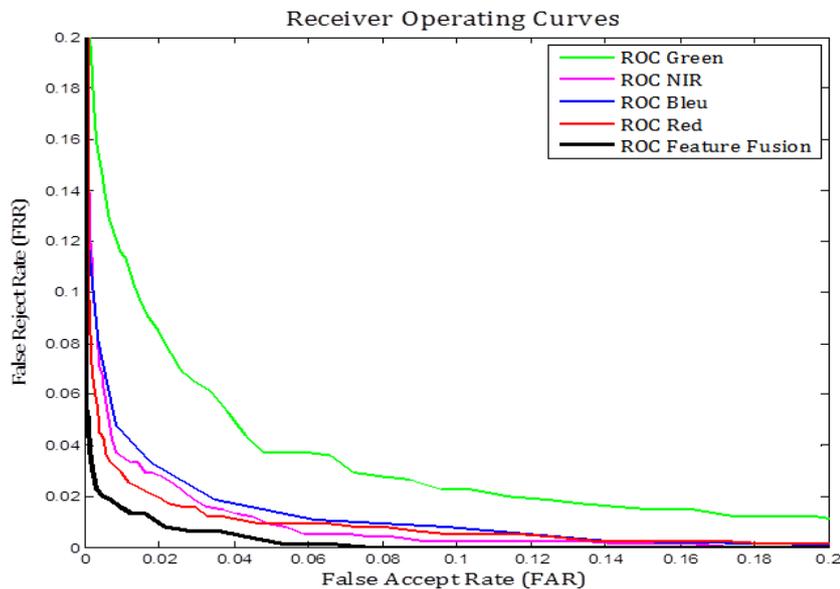


Figure 5.15. Courbe ROC de l'approche proposée

Les valeurs de EER obtenues en utilisant l'algorithme proposé pour les caractéristiques fusionnées et pour les différentes bandes spectrales sont: $EER(\text{fusion})=0.0167\%$, $EER(\text{Red})=0.0219\%$, $EER(\text{NIR})=0.0314\%$; $EER(\text{Bleu})=0.0372\%$ et $EER(\text{Green})=0.0575\%$. D'après les résultats obtenus, nous pouvons confirmer que l'approche proposée est robuste et performante. Ceci est dû, du fait que plus la valeur EER est faible et plus le système biométrique de vérification est performant.

3. Evaluation de la fusion

Dans ce travail le choix de la DWT comme méthode de fusion n'est pas fortuit. Dans le but de valider ce choix, en plus de la représentation visuelle; plusieurs critères d'évaluation de la qualité de la fusion ont été utilisés; à savoir : l'entropie, l'écart type, le facteur de fusion et la symétrie de la fusion. En effet, plus la valeur de l'entropie est grande plus l'image contient de l'information. D'où l'utilisation de cette métrique pour mesurer le contenu en information de l'image fusionnée. Les critères de mesure des performances de la fusion utilisés, sont définis comme suit [147]:

3.1. L'entropie

L'entropie mesure la quantité d'information contenue dans l'image. Plus la valeur est grande plus la quantité d'information dans l'image est grande. Formellement l'entropie est définie par l'équation suivante :

$$H = - \sum_{i=0}^{L-1} p_i \log_2(p_i) \quad (5.22)$$

où L représente le nombre de niveaux de gris de l'image fusionnée, pour une image type codée sur 8 bits, et p_i est le rapport du nombre de pixels de niveau de gris i et le nombre total de pixels (la fréquence normalisée de l'occurrence de chaque niveau de gris).

3.2. L'écart-type

Par ailleurs, l'écart type mesure le contraste de l'image fusionnée, plus la valeur de cette métrique est grande plus l'image contient du contraste. Cependant, cette mesure présente quelques limites en présence de bruits. Etant donnée une image d'une taille $M \times N$ l'écart type est calculé par la formule ci-après :

$$\sigma = \left(\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (F(m,n) - \mu)^2 \right)^{1/2} \quad (5.23)$$

où $F(m,n)$ représente la valeur du pixel de l'image fusionnée et μ la valeur moyenne de l'image.

3.3. Facteur de fusion

Etant donné deux images sources A et B, le Facteur de Fusion (FF) est donné par la formule :

$$FF = I_{AF} + I_{BF} \quad (5.24)$$

Où I_{AF} et I_{BF} représentent l'information mutuelle entre les images sources et l'image fusionnée. Une valeur élevée de FF indique que l'image fusionnée contient plus d'informations que les images sources. Cependant, ceci n'implique pas que l'information obtenue est fusionnée symétriquement. A cet effet, il est nécessaire d'examiner la métrique : facteur de fusion symétrique (FS).

3.4. Facteur de fusion symétrique

En effet, cette mesure renseigne sur la symétrie de l'image fusionnée par rapport aux images sources, plus petite est sa valeur plus efficace est l'algorithme de fusion. Le FS est évalué par la formule :

$$FS = abs\left(\frac{I_{AF}}{I_{AF} + I_{BF}} - 0.5\right) \quad (5.25)$$

Pour l'évaluation de la technique de fusion proposée, nous avons confronté la fusion par la DWT à plusieurs autres méthodes à savoir : La transformation en ondelettes stationnaires (SWT pour Stationary Wavelet Transform), la transformée de pyramide laplacienne (LPT pour Laplacian Pyramid Transform) et l'analyse en composantes principales (ACP). Le tableau 5.3 ci-dessous résume les résultats de la comparaison et montre la prévalence de la DWT.

| Méthode de fusion | Entropie (H) | Ecart type (σ) | Facteur de fusion (FF) | Facteur de Fusion Symétrique (FS) |
|--------------------------|---------------------|---|-------------------------------|--|
| DWT | 5.1671 | 0.0095 | 0.4197 | 1.1102e-016 |
| ACP | 0.1041 | 0.0340 | 0.4223 | 0.0024 |
| LPT | 0.4099 | 0.0355 | 0.2415 | 0.3961 |
| SWT | 0.0397 | 0.0282 | 0.4297 | 0.0113 |

Tableau 5.3 Comparaison et évaluation de plusieurs méthodes de fusion

En examinant les valeurs métriques obtenues dans le tableau 5.3, nous pouvons facilement observer que la transformée en ondelettes discrète permet une fusion efficace des informations discriminantes contenues dans les différentes bandes spectrales. Pour un niveau de décomposition optimal $J = 3$, l'approche proposée donne un entropie élevé ($H=5.1671$), un facteur de fusion symétrique faible ($FS=1.1102e-016$) et un facteur de fusion ($FS=0.4197$). Donc, la méthode de fusion DWT possède les meilleures valeurs pour les métriques considérées (dans la plupart des cas) par rapport aux autres approches testées.

4. Résultats de la classification de l'authentification et l'identification

Dans notre approche, deux types de classificateurs à savoir K-NN et SVM [160-164] ont été utilisés et comparés pour les images palmaires multispectrales en se basant sur l'analyse de la distribution de l'énergie des coefficients FRIT et l'analyse statistique. Pour le cas multi-classes, nous avons choisi la stratégie OAO introduite par Krebel [100]. Cette stratégie transforme le problème de classification à n classes en $n(n - 1)/2$ classifications binaires. Ensuite la classe votée est déterminée conformément à la stratégie de vote "Max-Win". Etant donné, un vecteur de caractéristiques x , la fonction de décision de la classe i contre la classe j est exprimée par :

$$D_i(x) = \sum_{j=1}^n sign(D_{ij}(x)) \quad (3.26)$$

avec $signD_{ij}(x) = \begin{cases} 1 & \text{if } D_{ij} > 0 \\ -1 & \text{else} \end{cases}$; le vecteur x est assigné à la classe k qui satisfait l'équation (4.73).

Pour comparer les classificateurs, nous avons choisi diverses statistiques basées sur la matrice de confusion à savoir: la spécificité, la sensibilité et la précision (accuracy). Dans le cas global, ils sont donnés par :

$$sensibilité\ globale = \frac{\sum_1^n TP_i}{\sum_{i=1}^n TP_i + \sum_{i=1}^n FN_i} \quad (3.27)$$

$$spécificité\ globale = \frac{\sum_1^n TN_i}{\sum_{i=1}^n TN_i + \sum_{i=1}^n FP_i} \quad (3.28)$$

$$précision\ globale = \frac{\sum_1^n TP_i + \sum_1^n TN_i}{\sum_{i=1}^n TP_i + \sum_{i=1}^n TN_i + \sum_{i=1}^n FP_i + \sum_{i=1}^n FN_i} \quad (3.29)$$

où n est le nombre de classe, TP_i , TN_i , FP_i , and FN_i représentent respectivement le nombre de vrai positif, vrai négatif, faux positif et faux négatif de la $i^{ème}$ classe.

Nos expérimentations ont été conduites sur les bases d'images d'empreintes palmaires multispectrales CASIA et PolyU qui renferment pour chaque individu six acquisitions ou représentations différentes de la palme de la main. Ce nombre d'acquisitions différentes nous permettra de conforter nos résultats quand à notre souhait d'identifier les individus quelque soit la position de la main. Afin de mener à bien l'expérimentation d'une manière rigoureuse, il est primordial d'éviter les choix aléatoires, que ce soit pour l'affectation d'une classe (label) à un individu de la base, ou le choix d'une acquisition pour le représenter. A cet effet, nous avons eu recours à la validation croisée dont la moyenne des estimations correspond aux taux réels d'imprécision ou de précision des méthodes évaluées, aussi celle-ci permet d'éviter le phénomène de sur-apprentissage, et donc par la même occasion permet d'estimer le vrai pouvoir de généralisation des classificateurs. Par ailleurs, nous avons procédé à deux validations croisées imbriquées dans notre algorithme, celle située à l'extérieur a été utilisée afin de libeller les individus, conséquemment, chaque individu se verra alternativement considéré comme authentique et imposteur. Par contre, Celle interne permettra de considérer alternativement chaque individu comme échantillon de la base de test et de la base d'apprentissage. Chaque base choisie a été divisée en quatre parties contenant plus ou moins le même nombre d'échantillons successifs, dans ce qui suit chaque partie sera appelée répertoire. Deux répertoires constitueront la classe authentique, tandis que les deux autres formeront la classe imposteur. Afin que notre algorithme puisse avoir une bonne aptitude de généralisation, nous avons effectué toutes les combinaisons possibles des répertoires et réitéré les calculs pour chaque combinaison, pour retenir enfin la moyenne de tous les résultats obtenus. Sachant que chaque individu est

représenté par six acquittions différentes, pour chaque classe nous avons effectué une validation croisée à 6-répertoires (celle-ci représente la validation croisée interne). En d'autres termes, chaque classe est divisée en six répertoires contenant un nombre d'échantillons approximativement égale, puis on procède à toutes les combinaisons possibles pour extraire les bases d'apprentissage et de test. Dans cette expérimentation nous avons choisi d'effectuer la vérification en utilisant les SVMs dans le cas binaire (deux classes) et l'authentification par la méthode OAO (voir chapitre technique de classification) pour le cas multi-classes. En se basant sur les travaux de Hsu et al. [159,164], nous avons choisi la fonction noyau radiale (RBF pour Radial Basis Function) représentée par l'équation $K(\vec{x}_i, \vec{x}_j) = e^{-\frac{\|\vec{x}_i - \vec{x}_j\|^2}{2\sigma}}$. S'il existe des ajustements qui ne sont pas des moindres dans la techniques des SVMs, c'est bien le choix des paramètres définies par l'utilisateur C et σ représentant respectivement le compromis entre la maximisation de la marge et l'erreur de classification, et la largeur du noyau (§ chapitre classification). A cet effet, nous avons entrepris une recherche en grille (*grid-search*) sur C et σ pour les valeurs respectives suivantes $C = \{2^{-5}, 2^{-3}, 2^{-1}, \dots, 2^{13}, 2^{15}\}$ et $\sigma = \{2^{-15}, 2^{-13}, 2^{-11}, \dots, 2^1, 2^3\}$. Dans un premier temps nous avons construit une grille grossière (*coarse-grid*) qui contient les performances du classificateur données par les valeurs des couples (C, σ) , puis nous avons raffinée cette grille aux voisinages du couple de valeurs (C^*, σ^*) ayant donné le meilleur résultat de classification. Le tableau 5.4 montre les résultats obtenus pour la base CASIA pour plusieurs niveaux de fusion des images palmaires multispectrales en utilisant les K-plus proches voisins (KNN) et les SVMs. Nous avons considéré plusieurs méthodes d'optimisation pour ces derniers (SVMs) à savoir : LS (Least squares) et SMO (Sequential Minimal Optimization). La meilleure précision moyenne obtenue est celle de la fusion des caractéristiques dont la valeur est de 98,15%, ce qui correspond à FAR=0.023 et FRR=0.014.

| Niveau de fusion des pixels | | | | | |
|---------------------------------------|-------|-------|-------------|-------------|-----------|
| Classificateur | FAR | FRR | Spécificité | Sensibilité | Précision |
| KNN | 0.038 | 0.026 | 96.15 % | 97.44% | 96.79 % |
| SVM(SMO) $C=2^3, \sigma=2^{2.25}$ | 0.044 | 0.031 | 95.59% | 96.93% | 96.50% |
| SVM(LS) $C=2^{6.5}, \sigma=2^{2.25}$ | 0.044 | 0.023 | 95.59% | 97.73% | 96.71% |
| Niveau de fusion des caractéristiques | | | | | |
| KNN | 0.028 | 0.022 | 97,23% | 97,85% | 97.43% |
| SVM(SMO) $C=2^4, \sigma=2^3$ | 0.023 | 0.018 | 97,69% | 98,17% | 97.94% |
| SVM(LS) $C=2^{8.5}, \sigma=2^{3.2}$ | 0.023 | 0.014 | 97,69% | 98,57% | 98.15% |

Tableau 5.4 Résultats expérimentaux pour KNN et SVM binaires sur la base CASIA

Par ailleurs, le tableau 5.5 montre les résultats obtenus sur la base de PolyU d'empreintes palmaires multispectrales (sur les zones d'intérêt). Afin de mettre en exergue la robustesse de la méthode proposée, nous avons évalué la précision de la vérification pour chaque bande considérée séparément et celle du niveau de fusion des caractéristiques. La précision moyenne de la fusion des caractéristiques est de 99.89% , ce qui correspond à FAR=0.0018 et FRR = 0.0022.

| Niveau de fusion de caractéristiques SVM $C = 2^1, \sigma = 2^{1.75}$ | | | | | |
|---|--------|--------|-------------|-------------|---------------|
| Modalité | FAR | FRR | Spécificité | Sensibilité | Précision |
| Bande rouge | 0.0018 | 0.0022 | 99.78% | 99.82% | 99.89% |
| Bande verte | 0.0045 | 0.0012 | 99.88% | 99.55% | 99.77% |
| Bande bleu | 0.000 | 0.0021 | 99.79% | 100 % | 99.98% |
| Bande NIR | 0.000 | 0.000 | 100 % | 100 % | 100 % |
| Fusion des caractéristiques pour toutes les bandes | 0.0018 | 0.0022 | 99.78% | 99.82% | 99.89% à 100% |

Tableau 5.5. Résultats expérimentaux avec les SVMs binaires validés sur la base PolyU

Les résultats du processus d'identification mené par la méthode OAO pour les différents types d'optimisation (LS, SMO et QP pour Quadratic Programming) et validés sur les deux bases , sont illustrés dans le tableau 5.6 ci-dessous.

| Base CASIA | | | |
|-----------------------------------|---------------------------|---------------------------|-------------------------|
| Méthode d'optimisation | Moyenne de la sensibilité | Moyenne de la spécificité | Moyenne de la précision |
| SMO ($C=2^5 \sigma = 2^{2.75}$) | 94.14% | 99.89% | 99.78% |
| QP ($C=2^5 \sigma = 2^{2.75}$) | 94.14% | 99.89% | 99.78% |
| LS ($C=2^6 \sigma = 2^4$) | 94.44% | 99.89% | 99.79% |
| Base PolyU | | | |
| SMO, QP et LS | 100% | 100% | 100% |

Tableau 5.6. Résultats expérimentaux avec la méthode OAO validés sur les deux bases

Finalement, afin de montrer l'efficacité et la robustesse de notre approche proposée, nous l'avons confronté aux méthodes de l'état de l'art les plus connues. Le tableau 5.7 résume le résultat de la comparaison en terme de précision.

| Méthodes | Auteurs | Précision |
|---|---------------------------------|--|
| D_1, D_2, D_3 (PCA, LDA) | J. Cui [130] | 99.57% (PolyU) |
| Image fusion scheme based on dependency measure | R. Raghavendra & C. Busch [131] | Protocole 1 EER= 1.64% Protocole2 EER= 7.13% (CASIA) |
| Representation in feature Space | S. Zhang & X. Gu [165] | 98.20 % sur le Spectre Rouge (PolyU) 96.50 % sur le Spectre Vert 96.70 % sur le Spectre Bleu 98.40 % sur le Spectre NIR |
| Online joint palmprint and palm vein verification | D. Zhang et al. [127] | EER = 0.021% à 0.0158% (PolyU) |
| The fused BDOC-BHOG method | D. Hong et al. [166] | EER = 0.0074% (PolyU) |
| Notre Approche | | 99.89 % à 100% (PolyU) 98.15 % (CASIA) |

Tableau 5.7. Confrontation de l'approche proposée à celles de l'état de l'art.

5. Discussion des résultats

La transformation en Ridgelets est plus appropriée pour l'extraction des caractéristiques multidirectionnelles des images présentant des structures radiales linéaires. En effet, c'est le cas pour la plupart des structures dominantes de l'empreinte palmaire, mais il demeure des structures qui ne le sont pas. Afin de remédier à ce problème est de traiter aussi les structures non linéaires nous avons divisé les images des empreintes palmaires en blocs qui ne se recouvrent pas. Par conséquent, les structures de chaque bloc peuvent ainsi être considérées comme linéaires, et l'efficacité se trouve ainsi booster. Une limitation de cette technique concerne le nombre de blocs choisi. Effectivement, dans notre expérimentation ce nombre a été choisi d'une manière empirique est doit être un nombre premier. Par ailleurs, la méthode de classification choisie génère une grande complexité dans sa phase d'apprentissage. Heureusement, elle n'est effectuée qu'une seule fois, d'autant plus que la phase de test est rapide et directe, celle-ci n'affecte pas les performances du système proposé. Par ailleurs, de nos jours il y a un réel besoin d'adapter les systèmes de reconnaissance aux besoins des données massives générées par les différents secteurs. L'utilisation de la programmation parallèle est nécessaire pour ce genre d'application, surtout qu'avec l'avènement d'outils appropriés tels que (les langages de programmation des unités graphiques, la programmation distribuée et la programmation multicores). Le système que nous avons proposé est approprié pour une utilisation parallèle, ceci peut être constaté à travers le fait que nous appliquons le même traitement sur chaque bloc de l'image palmaire que nous avons scindée en plusieurs blocs dans la phase d'extractions des caractéristiques. Par ailleurs, la

partie classification peut elle aussi bénéficier de ces techniques, il suffit d'effectuer chaque classification binaire de la méthode *OAO* sur des CPUs, cores ou nœuds différents.

5.4. Conclusion

Dans ce chapitre nous avons présenté deux travaux validés et publiés sur l'utilisation des transformations géométriques dans le domaine de la biométrie. Dans la première partie de nos travaux nous avons proposé une nouvelle approche pour caractériser les empreintes palmaires basée sur la transformation de Radon et la triangulation de Delaunay. Les résultats obtenus ont montrés que cette approche présente une grande stabilité et une robustesse effective. Au cours de notre travail, nous avons expérimenté et extraits plusieurs caractéristiques statistiques et attributs visuels de plusieurs images d'empreintes palmaires. Nous avons constaté que le nombre de centroïds de la transformation de Radon est plus petit comparé à ceux des minuties, ce qui a permis de réduire la complexité des algorithmes. Nous envisageons comme perspective d'améliorer l'approche proposée en combinant les signatures obtenues par celle-ci; à de l'information utile, telle que la texture de la main.

Par ailleurs, dans la seconde partie nous avons proposé une approche qui contribue à la vérification et l'authentification des individus en utilisant des images d'empreintes palmaires multispectrales. Nous avons expérimenté la fusion de l'information à travers deux niveaux de fusion, la fusion au niveau du pixel et la fusion des caractéristiques. La phase d'extraction des caractéristiques proposée basée sur l'analyse statistique, et la distribution d'énergies des coefficients de la transformée Ridgelets est simple, mais efficace et d'une complexité moindre comparée à plusieurs autres méthodes. L'étude comparative menée a montré que nos résultats en terme de précision sont meilleurs et rivalisent avec les approches de l'état de l'art.

Conclusion Générale et Perspectives

Les bases d'ondelettes séparables utilisées pour décomposer les images possèdent un support carré indéformable et elles sont isotropes. Pour ces raisons, elles ne peuvent pas représenter de manière optimale des régions comportant des contours ou des singularités locales. De nombreuses transformées en ondelettes anisotropes et aptes à la représentation optimale d'images comportant des contours, ont été proposées dans cette thèse afin de pallier à cet inconvénient. Ces transformées sont les transformées en ondelettes géométriques.

Nous avons tout d'abord présenté les approches non adaptatives pour lesquelles la décomposition ne dépend pas du signal à analyser. Ces approches ont le mérite de ne pas nécessiter de surcoût pour les spécifier, lors de l'analyse l'image. En revanche, la plupart de ces transformées ont l'inconvénient d'être redondantes, ce qui est désavantageux pour leur application en compression d'image. Parmi, ces transformées, nous avons cité la transformée en Ridgelets, la transformée en Curvelets, etc.

Ensuite, nous avons présenté les approches dépendantes du signal, qui offrent en général une plus grande flexibilité que les approches non adaptatives. Ces transformées définissent une base dont les fonctions sont choisies au mieux pour s'adapter à une image donnée, de sorte à limiter la redondance. Pour cela, une étape préalable d'estimation de la géométrie de l'image est nécessaire avant de procéder à la décomposition. Pour ce type de transformées, nous nous sommes focalisés principalement sur les transformées en Bandelettes introduites par Lepennec et Mallat, ensuite développées pour des applications de compression par Peyré et Mallat.

Deux axes de recherche ont été étudiés et développés dans le cadre de cette thèse: Le premier axe consiste à l'étude et l'application des transformées directionnelles pour la compression progressive des images hautes résolutions (images satellitaires et images médicales). Alors que le deuxième axe est porté sur la reconnaissance biométrique des images palmaires et des images palmaires multispectrales.

- ***Compression des images par la transformée en bandelette d'ordre 0.***

Dans cette partie de travail, nous avons étudié et implémenté la transformée en bandelette d'ordre 0 pour la compression des images SAR et médicales. La qualité des images reconstruites montre que cette méthode est capable de préserver les caractéristiques géométriques de l'image originale. Nous avons aussi proposée une approche de compression basée sur la transformée en bandelette combinée avec le codage des coefficients transformés par l'algorithme SPIHT, qui permet une compression progressive de l'image avec un contrôle précis du

débit de compression. Une étude comparative entre plusieurs transformées directionnelles (ondelettes, bandelettes et Ridgelets) a été aussi présentée dans cette contribution. Les résultats obtenus sont encourageants. Ainsi, diverses directions peuvent être envisagées en ce qui concerne l'amélioration de l'approche proposée pour la compression, nous citons :

- L'amélioration des résultats de compression en combinant la transformée en Bandelette d'ordre 0 avec le codeur EBCOT et comparer avec l'encodeur SPHIT.
 - L'Optimisation du seuil de quantification, choisi généralement empiriquement par l'utilisateur. Pour cela, il est intéressant de trouver une relation directe entre le débit et le seuil de quantification, ensuite minimiser cette relation pour un débit désiré.
 - Il est souhaitable que cette étude soit exploitée aux images en 3D, aux images couleurs et à la vidéo.
- ***Authentification biométrique de l'identité d'un individu en utilisant les empreintes palmaires et les empreintes palmaires multispectrales.***

Dans cette partie, nous avons proposé deux systèmes biométriques. Le premier système est basé sur la transformée de Radon et la triangulation de Delaunay pour caractériser les lignes principales de l'empreinte palmaire. A cet effet nous nous sommes basés sur l'encodage de leurs positions dans le domaine de Radon. Ceci a permis de construire une structure unique, stable et robuste au bruit propre à chaque individu. Afin de comparer deux structures obtenues par triangulation de Delaunay de deux images palmaires différentes, nous avons utilisé dans notre approche; la distance de Hausdorff comme une métrique de mesure de similarité.

Dans le deuxième système, nous nous sommes orientés dans nos recherches vers la multibiométrie en utilisant deux niveaux de fusion à savoir : le niveau pixel basé sur la transformée d'ondelette et le niveau caractéristique par concaténation. Différentes modalités des empreintes palmaires multispectrales (telle que les veines de la paume), ont été utilisées dans cette approche, et qui sont riches, robustes et difficile à usurper. L'algorithme multibiométrique proposé est basé sur l'analyse statistique et la transformée Ridgelets par bloc pour caractériser les modalités biométriques. Afin de valider les approches proposées, les expérimentations ont été conduites sur deux bases d'images palmaires multispectrales disponibles au public; faisant autorité dans ce domaine à savoir : la base CASIA et la base PoyU de l'institut technique de Hong Kong. Les deux bases de données sont capturées dans différentes bandes spectrales; chaque bande spectrale présente les fonctionnalités spécifiques de la paume. Sachant que chaque individu est représenté par six acquittions différentes, pour chaque classe nous avons effectué une

validation croisée à 6-répertoires. Dans cette expérimentation nous avons choisi d'effectuer la vérification en utilisant les K-NN et les SVMs dans le cas binaire (deux classes) et l'authentification par la méthodes *OAO*. L'étude comparative menée a montré que nos résultats en terme de précision sont meilleurs et rivalisent avec les approches de l'état de l'art.

Enfin, les différents travaux de recherches menés pour l'authentification biométrique, ouvrent différentes perspectives pour des travaux futurs:

- Valider les performances des systèmes biométriques développés, en appliquant les autres niveaux de fusion à savoir le niveau décision et le niveau score.
- Exploiter ces systèmes aux images palmaires 3D et à d'autres modalités biométriques.
- Pour le premier système biométrique proposé, les résultats obtenus ont été validés sur une base locale au laboratoire LiSSi, il est plus intéressant d'élargir la base en utilisant les bases CASIA ou PolyU.

Références

- [1] J.G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two dimensional visual cortical filters," J. Opt. Soc. Am., Vol.2(7), pp. 1160-1169, 1985.
- [2] D. Hubel and T. Wiesel, "Receptive fields and functional architecture of monkey striate cortex", Journal of Physiology, vol. 195, pp. 215–243, 1968.
- [3] R.L. De Valois, D.G. Albrecht, L.G. Thorell, Spatial frequency selectivity of cells in Macaque visual cortex, Vis. Res., Vol. 22 (5), pp. 545–559, 1982.
- [4] L. Jacques, L. Duval, C. Chaux et G. Peyré, "A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity," Signal Processing, Vol. 91, pp. 2699-2730, 2011.
- [5] S.G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 11 (7), pp. 674-693, 1989.
- [6] E. Le Pennec, "Bandelettes et représentation géométrique des images," thèse de Doctorat, Ecole Polytechnique, 2002
- [7] E. Le Pennec and S. Mallat, "Sparse geometric image representation with bandelets," IEEE Trans. Image Proc., vol. 14 (4), pp. 423–438, 2004.
- [8] G. Peyré, "Géométrie multi-échelles pour les images et les textures", Thèse de Doctorat, Ecole polytechnique, 2005.
- [9] S. Mallat et W-L. Hwang, "Singularity detection and processing with wavelets", IEEE Trans. on information theory, Vol. 38 (2), pp. 617-643, 1992.
- [10] S. Mallat et S. Zhong, "Characterization of Signals from Multiscale Edges", IEEE. Trans. on Pattern Analysis and Machine Intelligence, Vol. 14 (07), pp. 710-732, 1992.
- [11] CS Burrus, R.A. Gopinath et H. Guo, "Introduction to wavelet and wavelet transform", Ed. Prentice Hall, New jersey, 1998.
- [12] E. J. Candes et D. L. Donoho, "Curvelets, multiresolution representation, and scaling laws", Proc. SPIE 4119, Wavelet Applications in Signal and Image Processing, 2000.
- [13] M. N. Do et M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation", IEEE Transactions on Image Processing, 14(12), pp. 2091–2106, 2005.
- [14] A. Bouchemha, N. Doghmane et A. Nait-ali, "Multiscal geomtric Analysis for SAR image compression", 3rd International Conference on Information Processing and Electrical Engineering, Tebessa, Algeria, 2014.
- [15] A. Bouchemha, N. Doghmane et A. Nait-ali, "A robust technique to characterize the palmprint using radon transform and Delaunay triangulation", International Journal of Computer Applications, Vol.10(10), pp.35-42, 2010.
- [16] Amel Bouchemha, "Classification of multispectral palmprint features by pixel level fusion in Ridgelet domain: A robust biometric technique," 3rd European Workshop on visual information processing, 2011.
- [17] A. Bouchemha, N. Doghmane, A. Nait-ali, "Level feature fusion of multispectral palmprint recognition using the ridgelet transform and OAO multi-class classifier," 10th international conference on networking, sensing and control , pp. 771-774, 2013.
- [18] A. Bouchemha, N. Doghmane, A. Nait-ali et M.C. Nait-hamoud, "Multispectral palmprint recognition methodology based on multiscale representation," J. Electron. Imaging., Vol. 24(4), 2015.
- [19] P. C. Lemarie and Y. Meyer, "Ondelettes et bases hilbertiennes," Rev. Mat. Ibero-Amer, Vol. 2 (1), pp. 1–18, 1986.

- [20] Y. Meyer, " *Ondelettes, fonctions splines et analyses graduées*," Rend. Sem. Mat., Université de Torino, Vol. 45(1), 1987.
- [21] Y. Meyer, " *Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs*," Séminaire N. Bourbaki, n° 662, pp. 209-223, 1985.
- [22] I. Daubechies, " *Orthonormal bases of compactly supported wavelets*," Communications on Pure and Applied Mathematics, Vol. 41, pp. 909–996, 1988.
- [23] I. Daubechies, " *Ten Lectures on Wavelets*," Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
- [24] A. Grossman et J. Morlet, " *Decomposition of hardy functions into square integrable wavelets of constant shape*," SIAM Journal on Mathematical Analysis, vol. 15 (4), pp. 723–736, 1984.
- [25] M.N. Do, " *Directionnal Multiresolution Image Representation*," Ph.D. thesis, Department of Communication Systems, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, 2001.
- [26] M.N. Do and M. Vetterli, " *Contourlets: A directional multiresolution image representation*," in Proceedings of ICIP'02, Vol.1, pp. 357–360, 2002.
- [27] E. Candès, " *Ridgelet : Theory and Applications*," Ph.D. thesis, Department of statistics, Stanford university, 1998.
- [28] E. J. Candès et D. L. Donoho, " *Ridgelets: A key to higher-dimensional intermittency?*," Phil. Trans. R. Soc. Lond. A., Vol. 357, pp. 2495–2509, 1999.
- [29] E. J. Candès, " *Monoscale ridgelets for the representation of images with edges*," Technical Report, Department of Statistics, Stanford University, 1999.
- [30] J. M. Fadili et J-L. Starck, " *Curvelets and Ridgelets*," R.A. Meyers, ed. Encyclopedia of Complexity and Systems Science, 14, Springer New York, pp.1718-1738, 2009.
- [31] M.N. Do and M. Vetterli, " *The Finite Ridgelet Transform for Image Representation*," IEEE Trans. Image Processing, Vol.12, pp. 16-28, 2003.
- [32] E. J. Candès and D. L. Donoho, " *New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities*," Communications on Pure and Applied Mathematics, Vol. 57 (2), pp. 219–266, 2004.
- [33] L. Demanet, " *Curvelets, Wave Atoms and Wave Equations*," Ph. D. thesis, California Institute of Technology, 2006.
- [34] E. Candès, L. Demanet, D. Donoho et L. Ying, " *Fast discrete curvelet transforms. Multiscale Modeling and Simulation*," Vol. 5(3), p. 861–899, 2006.
- [35] J-Luc Starck, E. J. Candès and D. L. Donoho, " *The Curvelet Transform for Image Denoising*," IEEE Trans. on image processing, Vol. 11 (6), pp. 670-684, 2002.
- [36] M.N. Do and M. Vetterli, " *Contourlets*," in Beyond Wavelets, G.V. Welland, Ed. New York: Academic, 2003.
- [37] D. D.-Y. Po and M. N. Do, " *Directional multiscale modeling of images using the contourlet transform*," IEEE Trans. on Image Processing, Vol. 15 (6), pp. 1610–1620, 2006.
- [38] P. J. Burt and E. H. Adelson, " *The Laplacian pyramid as a compact image code*," IEEE Trans. Commun., Vol. 31 (4), pp. 532–540, 1983.
- [39] R. H. Bamberger and M. J. T. Smith, " *A filter bank for the directional decomposition of images: theory and design*," IEEE Trans. Signal Process., Vol. 40 (4), pp. 882–893, 1992.
- [40] L. Demanet, and L. Ying, " *Wave atom and sparsity of oscillatory patterns*," Appl. Comput, vol. 23 (3), pp. 368-387, 2007.
- [41] G. Peyré, " *A Review of Adaptive Image Representations*," IEEE journal of selected topics in signal processing, Vol. 5 (5), pp. 896-911, 2011.

- [42] G. Peyre, S. Mallat, "Discrete bandelets with geometric orthogonal filters", in: Proceedings of ICIP'05, Vol. 1, pp. 65-68, 2005.
- [43] S. Yang, Y. Lu, M. Wang et L. Jia, "Low bit rate SAR image coding based on adaptive multiscale Bandelets and cooperative decision," Signal Processing, Vol.89, pp. 1910-1920, 2009.
- [44] D. L. Donoho, "Wedgelets: Nearly minimax estimation of edges," in *Annals of Statistics*, Vol. 27(3), pp. 859-897, 1999.
- [45] V. Velisavljević, B-L. Baltasar et V. Martin, "Directionlets: Anisotropic multidirectional representation with separable filtering," IEEE Transactions on Image Processing, 15(7), pp. 1916-1933, 2006.
- [46] V. Velisavljević, M. Vetterli, B-B. Lozano, "Sparse Image Representation by Directionlet," Advances in Imaging and Electron Physics., Vol. 161(4), pp. 147-209, 2010.
- [47] W. Ding, F. Wu, and S. Li, "Lifting-based wavelet transform with directionally spatial prediction," in Proc. Picture Coding Symposium, San Francisco, USA, 2004.
- [48] C-Ling Chang and B. Girod, "Direction-Adaptive Discrete Wavelet Transform for Image Compression", IEEE Trans. on image processing, Vol. 16 (5), pp. 1289-1302, 2007.
- [49] V. Chappelier, "Codage progressif d'images par ondelettes orientées," Thèse de Doctorat, IRISA/TEMICS, Université de Rennes 1, 2005.
- [50] V. Chappelier and C. Guillemot, "Oriented Wavelet Transform for Image Compression and Denoising", IEEE Trans. on image processing, Vol. 15, pp. 2892-2903, 2006.
- [51] G. Lebrun, "Ondelettes géométriques adaptatives : vers une utilisation de la distance géodésique," Thèse de doctorat, Université de Poitiers, France 2006.
- [52] X. Delaunay, "Compression d'images satellite par post-transformées dans le domaine ondelettes", Thèse Doctorat de l'université de Toulouse, 2008.
- [53] R. Baxter et M. Seibert, "Synthetic Aperture Radar Image Coding", Lincoln laboratory journal, Vol. 11 (2), pp.121-158, 1998.
- [54] K. Sayood, "Introduction to Data Compression," 3rd Ed. Morgan Kaufmann Elsevier, 2006.
- [55] D.A. Huffman, "A method for the construction of minimum-redundancy codes," Proc. of the IRE, Vol.40, pp. 1098-1101, 1952.
- [56] I. H. Witten, R. M. Neal, et J. G. Cleary, "Arithmetic Coding for Data Compression," Communications of the ACM, Vol. 30 (6), pp. 520-540, 1987.
- [57] T. Acharya et P. S. Tsai, "JPEG2000 standard for image compression concepts, algorithms and VLSI architectures," John Wiley & Sons, New Jersey 2005.
- [58] M. Vetterli, "Wavelets, approximation and compression," in IEEE Signal Processing Magazine, pp. 59-73, Sept. 2001.
- [59] A. K Jain, "Image data compression: a review," Proceedings of IEEE, Vol. 69(3), pp. 349-389, 1981.
- [60] S. Grgic, M. Grgic et Z-C. Branka, "Performance Analysis of Image Compression Using Wavelets," IEEE Trans. on industrial electronics, Vol. 48 (3), pp. 682-695, 2001.
- [61] E. Le Pennec et S. Mallat, "Image compression with geometrical wavelets", IEEE International Conference on Image Processing, Vol. 1, pp. 661-664, 2000.
- [62] G. K. Wallace, "The JPEG still Picture compression Standard," Communication of the ACM, Vol. 34 (4), pp. 30-44, 1991.
- [63] J.W. Mark, "Adaptive predictive run-length encoding for analogue sources," Proc. IEE, Vol. 123 (11), pp. 1189-1196, 1976.
- [64] J. Taquet, "Techniques avancées pour la compression d'images médicales," Thèse de Doctorat, Université de Rennes 1, 2011.
- [65] Y. Gaudeau, "Contributions en compression d'images médicales 3D et d'images naturelles 2D," Thèse de Doctorat, Université Henri Poincaré, Nancy 1, 2006.
- [66] J.W. Woods et S.D. O'Neil, "Subband coding of images," IEEE Trans. on Acoustics, Speech and Signal processing, Vol. 34 (5), pp. 1278-1288, 1986.

- [67] A. Islam et W. A. Pearlman, "An embedded and efficient low-complexity hierarchical image coder," in SPIE Visual Communications and Image Processing, vol. 3653, pp. 294-305, 1999.
- [68] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," IEEE Transactions Signal Processing, Vol. 41, pp. 3445-3462, 1993.
- [69] A. Said et W. A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Trans. On Circuits and Systems for Video Technology, Vol. 6 (3), pp. 243-250, 1996.
- [70] D. Taubman, "High performance scalable image compression with EBCOT", IEEE Trans. on image processing, vol. 9(7), pp. 1158-1170, 2000.
- [71] X. Li, "On Exploiting Geometric Constraint of Image Wavelet Coefficients", IEEE Trans. on image processing, Vol. 12 (11), pp. 1378-1387, 2003.
- [72] J. Liu and P. Moulin, "Information-Theoretic Analysis of interscale and intrascale Dependencies between image wavelet coefficients", IEEE Trans. on image processing, Vol. 10 (11), pp. 1647-1658, 2001.
- [73] X. Delaunay, M. Chabert, V. Charvillat et G. Morin, " Satellite image compression by post-transforms in the wavelet domain," Signal Processing, Vol. 90, pp. 599–610, 2010.
- [74] A. K. Jain, A. Ross, et S. Prabhakar, " An Introduction to Biometric Recognition," IEEE Trans. on circuits and systems for video technology, Vol. 14 (1), pp. 4-20, 2004.
- [75] A. Naït-Ali et R. Fournier, "Signal and image processing for biometrics", Ed. ISTE Willy, 2012.
- [76] A. Jain, K. Nandakumar, A. Ross, " Score normalization in multimodal biometric systems," Pattern Recognition, Vol. 38, pp. 2270-2285, 2005.
- [77] A- K. Jain, A. Ross and S. Pankanti, "Biometrics: A Tool for Information Security", IEEE Trans. on information forensics and security, Vol. 1,(2), pp. 125-143, 2006.
- [78] A. Ross, A. K. Jain, "Human recognition using biometrics: an overview", Ann. Telecommunication, 62, (n°1-2), pp. 11-35, 2007.
- [79] M. El Abed, "Evaluation de système biométrique", Thèse de Doctorat de l'Université de Caen Basse-Normandie, 2006.
- [80] L. Hong, A. K. Jain, et S. Pankanti, "Can multibiometrics improve performance ?," in Proc. AutoID'99 Summit. NJ, pp. 59–64, 1999.
- [81] P. Grother et E. Tabassi, "Performance of biometric quality measures," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 29, pp. 531–543, 2007.
- [82] ISO/IEC 19795-1: 2006 , Information technology : Biometric performance testing and reporting, Part1 : Principles and framework, 2006.
- [83] J. Daugman, "Biometric decision landscapes," Technical Report, N° 482, Computer Laboratory, University of Cambridge, 2000.
- [84] F. Perronnin, J-L. Dugelay, "Introduction à la Biométrie Authentification des Individus par Traitement Audio-Vidéo", Traitement du Signal, Vol. 19 (4), pp. 253-265, 2002.
- [85] K. Nandakumar, "Multibiometric Systems: Fusion Strategies and Template Security," Ph.D. Thesis of Michigan State University, 2008.
- [86] M. He et al., "Performance evaluation of score level fusion in multimodal biometric systems," Pattern Recognition Vol. 43, pp. 1789–1800, 2010.
- [87] J. Kittler, M. Hatef, R-P.W. Duin et J. Matas, "On Combining Classifiers", IEEE Trans. On Pattern analysis and machine intelligence, Vol. 20, (3), pp. 226-239, 1998.

- [88] L. Lam and C- Y. Suen, "*Application of Majority Voting to Pattern Recognition: An Analysis of Its Behavior and Performance*", IEEE Trans. on systems, man and cybernetics-Part A: Syst. and humans, Vol. 27,(5), pp. 553-568, 1997.
- [89] T. K. Ho, J- J. Hull and S-N. Srihari, "*Decision Combination in Multiple Classifier Systems*", IEEE Trans. on Pattern Analysis and machine intelligence, Vol. 16 (1), pp. 66-75, 1994.
- [90] L. Xu, A. Krzyzak et C.Y. Suen, "*Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition*," IEEE Trans. on systems, man, and cybernetics, Vol. 22(3), pp. 418-435, 1992.
- [91] A. Saranli, M. Demirekler, "*A statistical unified framework for rank-based multiple classifier decision combination*," Pattern Recognition, Vol. 34, pp. 865-884, 2001.
- [92] V. Vapnik, "*The Nature of Statistical Learning Theory*", Springer, New York 1995.
- [93] C.M. Bishop, "*Pattern recognition and machine learning*," Springer, 2006.
- [94] X. Wu, V. Kumer, "*The top ten algorithms in data mining*," CRC Press, 2009.
- [95] T. Mitchell, "*Machine Learning*," McGraw Hill, 1997.
- [96] P-N. Belhumeur, J- P. Hespanha, D-J. Kriegman, "*Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*," IEEE Trans. on pattern analysis and machine intelligence, Vol.19,pp. 711-720, 1997.
- [97] W. S. McCulloch et W. Pitts, "*A Logical calculus of the ideas immanent in nervous activity*," Bulletin of Mathematical Biophysics, Vol. 5, pp. 115-133, 1943.
- [98] M. Parizerau, "*Réseaux de neurones*", Université Laval, 2004.
- [99] C. Cortes, et V. Vapnik, "*Support vector networks*," Machine Learning, Vol. 20, pp. 273-297, 1995.
- [100] U. Krebel, "*Pairwise classification and support vector machines. Advances in Kernel Methods Support Vector Learning*," Eds. Cambridge, MA: MIT Press, pp.255-268, 1999.
- [101] C-W. Hsu, C-C. Chang, C-I. Lin,"*A practical guide to support vector classification*," Technical report, Departement of computer science and information engineering, National Taiwan university, 2003.
- [102] J. Cervantes, X. Li, W. Yu, "*Support vector machine classification based on fuzzy clustering for large data sets, Advances in Artificial Intelligence*," LNCS. Springer, Vol. 4293, pp. 572-582, 2006.
- [103] G.Lebrun, C. Charrier, H. cardot, "*SVM training time reduction using vector quantization*," 17th conference on patter recognition, pp.160-163, 2004.
- [104] C.-F. Lin, and S.-D. Wang, "*Fuzzy support vector machines*", IEEE Trans. on Neural Networks, Vol. 13, pp. 464-471, 2002.
- [105] D. Zhang, A. Kong, J. You and M. Wong, "Online palmprint identification", IEEE Trans. Pattern Anal. Mach. Intel, vol. 25, n°9, 1041–1050, 2003.
- [106] W. Xuan, L. Li and W. Mingzhe, «Palmprint verification based on 2D – Gabor wavelet and pulse-coupled neural network", *Knowledge-Based Systems*, Vol. 27, pp. 451–455, March 2012.
- [107] D. S Huang, W. Jia and D. Zhang, "Palmprint verification on principal lines," Pattern Recognition, Vol. 41, 1316-1328, 2008.
- [108] L. Zhang, D. Zhang, "*Characterization of palmprints by wavelet signatures via directional context modeling*," IEEE trans. Syst. Man, and cybernetics, Part B: Cybernetics, Vol. 34 (3), pp. 1335-1347, 2004.
- [109] C-C Han, H.L Cheng, C.L Lin et K.C. Fan, "*Personal authentication using palm-print features*," Pattern Recognition, Vol. 36, pp. 371 – 381, 2003.
- [110] A. Kumar, D. Zhang, "*Personal authentication using multiple palmprint representation*", Pattern Recognition, Vol. 38, pp. 1695–1704, 2005.

- [111] A. Kong, D. Zhang, "Palmprint identification using feature-level fusion", *Pattern Recognition*, Vol.39, pp. 478–487, 2006.
- [112] W. Jia, B.Ling, K-W. Cha et L. Heutte, "Palmprint identification using restricted fusion," *Applied Mathematics and Computation*, Vol. 205, pp. 927–934, 2008.
- [113] G.S. Badrinath, P. Gupta, "Palmprint based recognition system using phase-difference information", *Future Generation Computer Systems*, Vol. 28, pp.287–305, 2012.
- [114] D.Hong, Z. Pan and X. Wu, "Improved differential box counting with multi-scale and multi-direction: A new palmprint recognition method", *Optik*, Vol.125, pp. 4154–4160, 2014.
- [115] B. Zhang, W. Li, P. Qing and D. Zhang, "Palm-Print Classification by Global Features", *IEEE. Trans. Syst. Man. Cyber. Syst.*, Vol. 43(2), pp. 370-378, 2013.
- [116] G. Lu, D. Zhang and K. Wang, "Palmprint recognition using eigenpalms features," *Pattern Recognition Letters*, Vol. 24, pp.1463–1467, 2003.
- [117] T. Connie, A.T-B. Jin, M. G-K. Ong et D. N-C. Ling, "An automated palmprint recognition system", *Image and Vision Computing*, Vol. 23, pp. 501–515, 2005.
- [118] W. Li, J. You, "Texture-based palmprint retrieval using a layered search scheme for personal identification," *IEEE Trans. Multimedia*, Vol. 7(5), pp. 891–898, 2005.
- [119] G. Kah Ong Michael, T. Connie and A. Beng Jin Teoh, "Touch-less palm print biometrics: Novel design and implementation," *Image and Vision Computing*, Vol. 26, pp. 1551–1560, 2008.
- [120] F. Li, M. K.H. Leung et C.S. Chian, "Making Palm Print Matching Mobile," *Inter. Journal of Computer Science and Information Security*, vol.6 (2),pp.1-10, 2009.
- [121] D. Zhang, W. Zuo et F. Yue, "A Comparative Study of Palmprint Recognition Algorithms," *ACM Comput. Surv.*, Vol. 44 (1),pp.1-37, 2012.
- [122] Y. Hao, Z.Sun and T. Tan, "Comparative Studies on Multispectral Palm Image Fusion for Biometrics," *Proc. of 8th Asian Conference on Computer Vision, Part II*, pp.12-21, 2007.
- [123] J-G. Wang, W.Y. Yau, A. Suwandy et E. Sung, "Person recognition by fusing palmprint and palm vein images based on Laplacianpalm representation," *Pattern Recognition*, Vol.41(5), pp. 1514–1527 (2008).
- [124] Y. Zhou et A. Kumar, "Contactless Palm Vein Identification using Multiple Representations," 4th IEEE International Conference on Biometrics: Theory Applications and Systems, pp. 1-6, 2010.
- [125] Z. Guo, D. Zhang et D. Zhang, "Feature Band Selection for Multispectral Palmprint Recognition," 20th International Conference on Pattern Recognition, pp. 1136- 1139, 2010.
- [126] D. Zhang, Z. Guo G. Lu, L. Zhang et W. Zuo, "An Online System of Multispectral Palmprint Verification," *IEEE Trans. Inst. Measurement*, Vol.59(2), pp. 480-490, 2010.
- [127] D. Zhang, Z. Guo, G. Lu, L. Zhang, Y. Liu et W. Zuo, "Online joint palmprint and palm vein verification," *Expert Systems with Applications*, Vol. 38, pp. 2621-2631, 2011.
- [128] Y. Hao, Z. Sun, T. Tan, et C. Ren, "Multispectral palm image fusion for accurate free palmprint recognition," *Proc. of IEEE international conference on Image Processing*, pp. 281-284, 2008.
- [129] D. Han, Z. Guo et D. Zhang, "Multispectral palmprint recognition using wavelet-based image fusion," *Proc. of IEEE international Conference on Signal Processing*, pp. 2074-2077, 2008.
- [130] J. Cui, "Multispectral fusion for palmprint recognition," *Optik*, Vol.124, pp. 3067– 3071, 2013.
- [131] R. Raghavendra et C. Busch, "Novel image fusion scheme based on dependency measure for robust multispectral palmprint recognition," *Pattern Recognition*, Vol.47, pp. 2205–2221, 2014.
- [132] D.V. Jadhav et R. Holambe, "Feature extraction using Radon and wavelet transforms with application to face recognition," *Neuro computing*, Vol.72, pp. 1951–1959, 2009.
- [133] Q. Zhang et I. Couloiger, "Accurate Centerline Detection and Line width estimation of thick Lines Using the Radon Transform," *IEEE Trans. On Image processing*, Vol. 16 (2),pp. 310-316, 2007.

- [134] S. Tabbone, L. Wending et J.P Salmon, "A new shape descriptor defined on the Radon transform," Computer Vision and Image Understanding, Vol.102,pp. 42-51, 2006.
- [135] X. Wang, B. Xian, J-F. Ma et X-L. Bi, "Scaling and rotation invariant analysis approach to object recognition based on Radon and Fourier–Mellin transforms," Pattern Recognition, Vol. 40, pp.3503-3508, 2007.
- [136] X.Wang, F.X Guo, B. Xiao et J-F. Ma, "Rotation analysis and orientation estimation method for texture classification based on Radon transform analysis," Journal Vision Communication Image, Vol. 21, pp. 29-32, 2010.
- [137] Y. W. Chen et Y.Q Chen, "Invariant descriptor and retrieval for planar shapes using Radon composite feature," IEEE transaction on signal processing, Vol. 56(10),pp. 4762-4771, 2008.
- [138] G.Y. Chen et W.F. Xie, "Pattern recognition with SVM and dual-tree complex wavelets," Image Vis. Computing, Vol. 6,pp. 960–966, 2007.
- [139] A. Kumar et K. V. Prathysha, "Personal Authentication using hand vein triangulation," IEEE Trans. on Image Processing, Vol. 18,pp. 2127-2136, 2009.
- [140] T. UZ, G. Bebis, A. Erol and S. Prabhekar, "Minutiae- based template synthesis and matching for fingerprint authentication", Computer vision and Image Understanding, Vol. 113, pp. 979-992, 2009.
- [141] D.P. Huttenlocher, G.A. Klanderman et W.J. Rucklidge, "Comparing images using the Hausdorff distance," IEEE Trans Pattern Anal Mach Intel, Vol. 15 (9),pp. 850–863, 1993.
- [142] C. Zhao, W. Shi and Y. Deng, "A new Hausdorff distance for image matching", Pattern Recognition Letters, vol. 26, 2005, 581– 586.
- [143] E.P. Vivek, N. Sudha, "Gray Hausdorff distance measure for comparing face images," IEEE Transa. on Information, Forensics and Security, Vol. 1, pp. 342–349, 2006.
- [144] C-H.T. Yanga, S-H. Lai et L-W. Chang, "Hybrid image matching combining Hausdorff distance with normalized gradient matching," Pattern Recognition, Vol. 40, pp.1173 – 1181, 2007.
- [145] C-B. Yu, H-F. Qin, L. Zhang et Y-Z. Cui, "Finger-Vein Image Recognition Combining Modified Hausdorff Distance with Minutiae Feature Matching," J. Biomedical Science and Engineering, Vol. 2, pp. 261-272, 2009.
- [146] R. Singh, M. Vatsa et A. Noore, "Hierarchical fusion of multi-spectral face images for improved recognition performance," Information Fusion, Vol.9, pp. 200-210, 2008.
- [147] R. Singh et A. Khare, "Fusion of multimodal medical images using Daubechies complex wavelet transform– A multiresolution approach," Information fusion, Vol.19, pp. 49-60, 2014.
- [148] K. Amolins, Y. Zhang et P. Dare, "Wavelet based image fusion techniques —An introduction, review and comparison," ISPRS Journal of Photogrammetry & Remote Sensing, Vol.62, pp. 249–263, 2007.
- [149] R.A. Groeneveld et G. Meeden, " Measuring Skewness and Kurtosis," Journal of the Royal Statistical Society. Series D (The Statistician), Vol. 33 (4), pp. 391-399, 1984.
- [150] D. P. Doane, L. E. Seward, " Measuring Skewness: A Forgotten Statistic?," Journal of Statistics Education, Vol. 19 (2), pp. 1-18, 2011.
- [151] Chen, T-D. Bui et A. Krzyzak, "Rotation invariant pattern recognition using ridgelets, wavelet cycle-spinning and Fourier features," Pattern Recognition, Vol.38,pp. 2314-2322, 2005.
- [152] S. Arivazhagan, L. Ganesan et T.G. Subash Kumar, "Texture classification using ridgelet transform," Pattern Recognition Letters, Vol.27,pp. 1875-1883, 2006.
- [153] W. Pan, T-D. Bui et C.Y. Suen, "Rotation invariant texture classification by ridgelet transform and frequency–orientation space decomposition," Signal Processing, Vol.88, pp. 189-199, 2008.

- [154] S-B. Nikam et S. Agarwal, "*Ridgelet-based fake fingerprint detection*," Neurocomputing, Vol.72, pp. 2491-2506, 2009.
- [155] Y-X. Liu, Y-H. Peng et W.C Siu, "*Energy-based adaptive transform scheme in the DPRT domain and its application to image denoising*," Signal Processing, Vol.89,pp.31-44, 2009.
- [156] Y-L. Qiao, C-Y. Song et C-H. Zhao, "*M-band ridgelet transform based texture classification*," Pattern Recognition Letters, Vol.31, pp. 244-249, 2010.
- [157] CASIA Multispectral Palmprint Database [Online]. Available: <http://www.cbsr.ia.ac.cn/MSPalmprintDatabase.asp>
- [158] PolyU Multispectral Palmprint Database. [Online]. available: <http://www4.comp.polyu.edu.hk/~biometrics/MultispectralPalmprint/MSP.htm>
- [159] C-W. Hsu et C-J Lin, "*A Comparison of Methods for Multi-class Support Vector Machines*," IEEE Trans. Neural Networks, Vol.13(02), pp. 415-425, 2002.
- [160] F. Melgani et L. Bruzzone, "*Classification of Hyperspectral Remote Sensing Images with support vector machines*," IEEE Trans. Geosc. Remote Sensing, Vol.42(08),pp. 1778-1790 (2004).
- [161] G-Y. Chen et W-F. Xie, "*Pattern recognition with SVM and dual-tree complex wavelets*," Image and Vision Computing, Vol.25, pp. 960-966, 2007.
- [162] S. Lin-Lin et J. Zhen, "*Gabor Wavelet Selection and SVM classification for object recognition*," Acta Automatica Sinica, Vol. 35(04), pp. 350-355, 2009.
- [163] V. Ranaee, A. Ebrahimzadeh et R. Ghaderi, "*Application of the PSOSVM model for recognition of control chart patterns*," ISA Transactions, Vol.49, pp. 577-586, 2010.
- [164] C.W. Hsu, C.C. Chang et C.J. Lin, "*A Practical Guide to Support Vector Classification*," Technical report, Department of Computer Science, National Taiwan University, pp. 1-16, 2003.
- [165] S. Zhang and X. Gu, "*Palmprint recognition based on the representation in the feature space*," Optik, Vol.124,pp. 5434– 5439, 2013.
- [166] D. Hong, W. Liu, J. Su, Z. Pan et G. Wang, "*A novel hierarchical approach for multispectral palmprint recognition*," Neurocomputing, Vol.151, pp.511–521, 2015.