



HAL
open science

Extraction et fusion de points d'intérêt et textures spectraux pour l'identification, le contrôle et la sécurité

Diarra Mamadou

► **To cite this version:**

Diarra Mamadou. Extraction et fusion de points d'intérêt et textures spectraux pour l'identification, le contrôle et la sécurité. Vision par ordinateur et reconnaissance de formes [cs.CV]. Université Bourgogne Franche-Comté; Université de Cocody (Abidjan, Côte d'Ivoire; 1995-2012), 2018. Français. NNT : 2018UBFCK031 . tel-02130804

HAL Id: tel-02130804

<https://theses.hal.science/tel-02130804v1>

Submitted on 16 May 2019

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

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

SPIM

Thèse de Doctorat



école doctorale **sciences pour l'ingénieur et microtechniques**
UNIVERSITÉ DE BOURGOGNE

N°

THÈSE présentée par

MAMADOU Diarra

Pour obtenir le

Grade de Docteur de

L'Université de Bourgogne

Spécialité : **Instrumentation et Informatique de l'Image**

Extraction et fusion de points d'intérêt et textures Spectraux pour l'identification, le contrôle et la sécurité

Unité de Recherche :

Laboratoire Electronique, Informatique et Image

Soutenue publiquement le 2 Août 2018 devant le jury composé de :

ORIA Vincent	Président	Professeur Titulaire, New Jersey Institute of Technology USA
GOUTON Pierre	Directeur	Professeur Titulaire, Université de Bourgogne DIJON
ADOU K. Jérôme	Co-Directeur	Professeur Titulaire, Université Félix Houphouët-Boigny ABIDJAN
EZIN C. Eugène	Rapporteur	Maître de Conférences, Université Abomey-Calavi BÉNIN
LOUM Georges	Rapporteur	Maître de Conférences, INP-HB YAMOUSSOUKRO

Sommaire

Dédicace	V
Remerciements	VI
Résumé	VII
Abstract	VIII
Liste des tableaux	IX
Liste des figures	X
Liste des abréviations	XII
Introduction générale	1
1- Contexte et motivation	1
2- Objectif	1
3- Contribution	2
4- Liste des publications	3
5- Organisation de la thèse	3
Chapitre 1. Etat de l'art de la reconnaissance de visage	5
Introduction	5
1.1. Fondamentaux de la biométrie	5
1.1.1. Architecture générale	6
1.1.2. Biométries monomodales	7
1.2.3. Enjeu de la biométrie	8
1.2. Reconnaissance de visage	10
1.2.1. Motivation de la biométrie de visage	10
1.2.2. Méthodes locales	11
1.2.3. Méthodes globales.....	13
1.2.4. Méthodes hybrides	13
1.3. Reconnaissance de visage en infrarouge	13
1.4. Reconnaissance de visage multispectrale	16
1.4.1. Notion d'images multispectrales	16
1.4.2. Notion de biométrie multispectrale	18
1.4.3. Applications de la biométrie multispectrale	19
1.4.4. Reconnaissance de visage multispectrale	19
1.5. Reconnaissance de visage et biométrie multimodale	24
1.6. Reconnaissance de visage 3D	27
1.7. Mesure de performance d'un système biométrique	28
Conclusion	30
Chapitre 2. Reconnaissance de visage multispectrale basée sur une extraction hybride de caractéristiques (MS-FRHF)	31

Introduction	31
2.1. Notion d'extraction de caractéristiques.....	31
2.1.1. Points d'intérêt	32
2.1.2. Texture	32
2.2. Méthodes d'extraction de caractéristiques	33
2.2.1. Extraction de points d'intérêt	33
2.2.2. Méthodes d'extraction de points d'intérêt.....	34
2.2.3. Extraction de texture	40
2.3. Méthode MS-FRHF proposée pour l'extraction de caractéristiques en reconnaissance de visage multispectrale.....	44
2.3.1. Objectifs	44
2.3.2. Reconnaissance de visage dans le visible et l'infrarouge thermique.....	45
2.3.3. Fusion d'images	46
2.3.4. Fusion d'images visible/infrarouge thermique	47
2.3.5. Description de la base de données multispectrale	51
2.3.6. Extraction de caractéristique par la méthode MS-FRHF.....	52
2.3.7. Résultats expérimentaux	54
Conclusion	56
Chapitre 3. Reconnaissance de visage multispectrale à l'aide de caractéristiques binaires	57
Introduction	57
3.1. Motivation	57
3.2. Présentation de détecteurs et descripteurs binaires	58
3.2.1. Méthodes basées sur les caractéristiques locales	58
3.2.2. Fonctionnement des détecteurs et descripteurs binaires	58
3.2.3. BRISK	60
3.2.4. BRIEF	62
3.2.5. ORB	63
3.2.6. FREAK.....	64
3.3. Méthode proposée pour la reconnaissance de visage multispectrale à l'aide de détecteurs et de descripteurs binaires.....	65
3.3.1. Présentation des méthodes d'extraction de caractéristiques utilisées dans la méthode proposée.....	67
3.3.2. Matériels et méthodes	75
3.3.3. Résultats expérimentaux	78
Conclusion	81
Chapitre 4. Reconnaissance de visage multispectrale à l'aide de score de similarité	82
Introduction	82

4.1. Approches de comparaison d'objets.....	83
4.2. Motivation de l'utilisation des scores de similarité	83
4.3. Mesures de similarité appliquées aux images.....	85
4.3.1. Définitions.....	85
4.3.2. Applications des mesures de similarité.....	86
4.4. Approche proposée pour la reconnaissance de visage multispectrale à l'aide de score de similarité.....	87
4.4.1. Introduction à la notion de similarité	88
4.4.2. Approche théorique.....	89
4.4.3. Méthode proposée pour l'étude de similarité.....	94
4.4.4. Résultats expérimentaux	99
Conclusion	102
Conclusion générale et perspectives.....	103
Perspectives	104
Bibliographie.....	105
Table des matières	117

Dédicace

A mes parents

A mon épouse

A mes enfants

Remerciements

S'il existe sur cette "terre des hommes", des travaux menés seuls, celui-ci n'en est pas un. En effet, son aboutissement résulte de la contribution conjuguée de plusieurs personnes qu'il convient de remercier. Au rang de ces personnes :

- les Présidents des Universités Félix Houphouët-Boigny et de Bourgogne pour la convention de cotutelle de thèse entre les dites universités ;
- le Directeur de l'UFR Mathématiques et Informatique, Prof. ADJE Assohoun pour sa sollicitude ;
- les Professeurs GOUTON Pierre et ADOU Kablan Jérôme, mes Directeurs de thèse pour leurs encadrements et conseils avisés qui ont conduit à l'aboutissement de cette thèse ;
- le Doyen honoraire, le Professeur KOUA Konin pour ses conseils éclairés ;
- les Sous-Directeurs, les Responsables de laboratoires, les Responsables de filières pour leur contribution ;
- les collègues enseignants-chercheurs, les doctorants, le personnel administratif et technique des Universités de Bourgogne et Félix Houphouët-Boigny pour leur appui ;
- les Docteurs SANDA Amadou T. et KEIVAN Ansari pour leurs apports en imagerie ;
- tous ceux qui n'ont pas été nommément cités et qui m'ont aidé à un moment ou à un autre pour l'aboutissement de ce travail ;
- enfin, les membres du jury qui me font l'honneur de juger ce travail.

Résumé

La biométrie, domaine de la vision par ordinateur est une technologie émergente qui propose de nouvelles méthodes de contrôle, d'identification et de sécurité. Les applications biométriques sont souvent victimes de menaces. Les systèmes actuels sont basés sur de multiples modalités telles que la reconnaissance de visage, d'empreintes digitales, de l'iris, de la main, l'oreille, la démarche, la frappe au clavier, le pied, etc. Mais, la plus connue de ces techniques biométriques reste la reconnaissance de visage qui ne cesse de connaître des progrès. Dans le domaine de la reconnaissance faciale, plusieurs approches existantes utilisent des images dans le spectre visible. Ces systèmes traditionnels opérant dans le spectre visible souffrent de plusieurs limitations dues aux changements d'éclairage, de poses et d'expressions faciales.

Comme contribution, le présent travail utilise l'imagerie multispectrale, plus particulièrement l'imagerie infrarouge et visible, pour améliorer la performance de la reconnaissance faciale et pallier les insuffisances du spectre visible. Egalement, les spectres infrarouges présentent quelques limitations dont le port de lunettes, les modifications de la température ambiante et des processus métaboliques des sujets dans des conditions de peur, de sport, de maladie, etc. Les solutions proposées visent à apporter des contributions à la résolution de ces problématiques.

La méthodologie présentée dans cette thèse est basée sur la reconnaissance faciale multispectrale. Les images multispectrales utilisées dans l'étude sont obtenues par fusion d'images visibles et infrarouges thermiques extraites de la base de données multispectrale IRIS. Les différentes techniques de reconnaissance ont utilisé l'extraction de caractéristiques qui sont la texture et les points d'intérêt. D'abord, une première méthode est basée sur une extraction hybride de ces caractéristiques. Ensuite, une seconde méthode utilise les approches binaires d'extraction de caractéristiques. Enfin, une méthode de reconnaissance faciale est construite à l'aide de mesures de similarité basée sur une distance tenant compte des caractéristiques extraites.

Les résultats des expériences démontrent l'efficacité des méthodes proposées pour construire un système de reconnaissance des visages plus performant.

Mots-clés: reconnaissance de visage, multispectral, visible, infrarouge, points d'intérêt, texture.

Abstract

Biometrics, a field of computer vision, is an emerging technology that offers new methods of control, identification and security. Biometric applications are often victims of threats. Current systems are based on multiple modalities such as face recognition, fingerprint, iris, hand, ears, gait, keystroke, foot, etc. But, the most popular of these biometric techniques is the face recognition that continues to progress. In the field of facial recognition, several existing approaches use images in the visible spectrum. These traditional visible spectrum systems suffer from several limitations due to changes in lighting, poses and facial expressions.

As a contribution, the present work uses multispectral imaging, more particularly infrared and visible imaging, to improve the performance of facial recognition and overcome the deficiencies of the visible spectrum. Also, the infrared spectra present some limitations including the wearing of glasses, the modifications of the ambient temperature and the metabolic processes of the subjects in conditions of anxiety, sport, disease, etc. The proposed solutions aim to contribute to these issues.

The methodology presented in this thesis is based on multispectral facial recognition. The multispectral images used in the study are obtained by fusing visible and thermal infrared images extracted from the IRIS multispectral database. The different recognition techniques used the features extraction that are texture and interest points. First, a first method is based on a hybrid extraction of these characteristics. Then a method uses the binary methods of features extraction. Finally, a method based on similarity measures for facial recognition is proposed by the use of a distance taking into account extracted features.

The results of the experiments demonstrate the performance of the proposed methods to build a more efficient face recognition system.

Keywords: face recognition, multispectral, visible, infrared, key-point, texture.

Liste des tableaux

Tableau (1- 1) : Avantages et inconvénients de la reconnaissance de visage. -----	11
Tableau (1- 2) : Présentation de base de données infrarouge [113, 35].-----	16
Tableau (1- 3) : Biométries multispectrales [22]. -----	19
Tableau (1- 4) : Bases de données multispectrales-----	21
Tableau (1- 5) : Atouts et insuffisances de la reconnaissance de visage dans le visible, l'infrarouge passif et l'infrarouge actif (+ indique un avantage et - une limitation).-----	24
Tableau (1- 6) : Systèmes biométriques multimodaux-----	27
Tableau (2- 1) : Résumé de quelques détecteurs de coins -----	36
Tableau (2- 2) : Résumé de quelques détecteurs de coins -----	37
Tableau (2- 3) : Résumé de quelques détecteurs de régions -----	39
Tableau (2- 4) : Présentation des caractéristiques de trois paires d'images et des images fusionnées. De gauche à droite, Contraste, Corrélation, Energie, Homogénéité, Point-clé. -----	55
Tableau(3- 1) : Présentations de principaux détecteurs et descripteurs binaires. -	60
Tableau(3- 2) : Récapitulatif des attributs des détecteurs et descripteurs binaires présentés plus haut. -----	65
Tableau(3- 3) : Paires de détecteurs et de descripteurs.-----	76
Tableau (4- 1) : Présentation des méthodes statistiques et structurelles-----	83
Tableau (4- 2) : Groupes de mesures de similarité.-----	86
Tableau (4- 3) : Fonction de similarité pour calculer le score de similarité de deux images X et Y. -----	98
Tableau (4- 4) : Résultats du score de similarité entre les images en utilisant les distances décrites précédemment et la fonction de similarité proposée. -----	101

Liste des figures

Figure (1- 1) : Différents modules d'un système biométrique [49].	7
Figure (1- 2) : Récapitulatif des modalités biométriques [52, 63, 84].	10
Figure (1-3) : Spectre électromagnétique.	14
Figure (1- 4) : Ensemble d'images. De gauche à droite : image originale couleur, couche correspondante rouge, couche correspondante verte, couche correspondante bleue.	17
Figure (1- 5) : Exemple d'image multispectrale.	17
Figure (1- 6) : Différences d'hyperspectral et de multispectral [22].	18
Figure (1- 7) : Ensemble d'images. De gauche à droite : image visible, image infrarouge, image infrarouge thermique.	19
Figure (1- 8) : Méthodologie de reconnaissance de visage multispectrale [15].	20
Figure (1- 9) : Ensemble d'images extraites de bases de données multispectrales [5].	23
Figure (1- 10) : Architecture de fusion. (a) en série, (b) en parallèle.	26
Figure (2- 1) : Quelques exemples de textures [13].	33
Figure (2- 2) : Présentation de SIFT et ses variantes [128].	40
Figure (2- 3) : Images visible/ infrarouge (a et b) puis (c et d) acquises sous différentes illuminations	48
Figure (2- 4) : De la gauche à la droite : image infrarouge thermique, image visible, carte de similarité de la somme des différences quadratiques SSD.	49
Figure (2- 5) : Comparaison de trois méthodes de fusion utilisant quatre métriques avec l'image infrarouge thermique choisie comme référence.	50
Figure (2- 6) : Extraits d'images de la base de données IRIS : Image de deux individus acquises dans des conditions d'éclairage différentes.	52
Figure (2- 7) : Principales étapes du détecteur MSER [104]	53
Figure (2- 8) : De la gauche vers la droite, Image visible, thermique infrarouge, image fusionnée correspondante	55
Figure (3- 1) : Principales étapes du détecteur et du descripteur BRISK [109].	61
Figure (3- 2) : Principales étapes du détecteur et du descripteur ORB [30].	64
Figure (3- 3) : Modèle d'échantillonnage de BRISK [109].	69

Figure (3- 4) : Image montrant le point d'intérêt testé et les 16 pixels du cercle [107].	70
Figure (3- 5) : Principales étapes du détecteur de coin FAST [107].	71
Figure (3- 6) : Modèle d'échantillonnage FREAK [2].	72
Figure (3- 7) : Paires sélectionnées pour calculer l'orientation du point d'intérêt [2].	73
Figure (3- 8) : Ensemble d'images. De gauche à droite : image visible, image infrarouge thermique, carte de similarité entre les images visibles. / infrarouges par SSD, image fusionnée par la méthode de fusion PCA avec l'illumination Lon	75
Figure (3- 9) : Graphique d'évaluation de détecteurs et descripteurs.	77
Figure (3- 10) : Extraction des points-clés sur des images d'éclairage unique	79
Figure (3- 11) : Extraction de points-clés sous différentes conditions	80
Figure (4- 1) : Chaîne classique de traitement pour la reconnaissance biométrique.	82
Figure (4- 2) : Différence entre les représentations basées sur les caractéristiques et la similarité.	84
Figure (4- 3) : Représentations de similarité entre images couleurs : de gauche à droite $d_1, d_2, d_3, d_4, d_5, d_6$ des mesures de similarité entre les images respectives et celle au- dessus.	87
Figure (4- 4) : Ensemble d'images extraites de la base de données. Chaque ligne de haut en bas présente des images visibles, des images infrarouges thermiques correspondantes, des images fusionnées correspondantes par la méthode de fusion PCA	95
Figure (4- 5) : Schéma proposé pour la reconnaissance de visage à l'aide de score de similarité	99
Figure (4- 6) : Images utilisées pour le calcul du score de similarité.	100

Liste des abréviations

2D	Deux dimensions et représentation selon deux axes
3D	Deux dimensions et représentation selon trois axes
ACP	Analyse en Composantes Principales
AGAST	Accelerated
ASIFT	Affine- Scale-Invariant Feature Transform
AVG	Average
BRIEF	Binary Robust Independent Elementary
BRISK	Binary Robust Invariant Scalable Keypoints
CSIFT	Color Scale-invariant feature transform
CSLBP	Center Symmetric-Local Binary Pattern
EBR	Edge-Based Regions
FAST	Features from Accelerated Segment Test
FREAK	Fast Retina Keypoint
FSIM	Feature SIMilarity
GLBP	Geometrical Local Binary Patterns
GLCM	Grey Level Co-occurrence Matrix
GLOH	Gradient Localization and Orientation Histogram
GMM	Gaussian Mixture Model
GSIFT	Global Scale-Invariant Feature Transform
HLP	Hétérogénéité des Longueur de Plage
HNG	Hétérogénéité des Niveaux de Gris
HOG	Histogram of Oriented Gradients
IBR	Intensity - Based Region
IR	InfraRouge
IRIS	Imaging Robotics Intelligent Systems
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis
LON	Light ON
LWIR	Long Wave Infrared
MAE	Mean Absolute Error
MS – FRHF	Multispectral Face Recognition using Hybrid Feature
MSCR	Maximally Stable Colour Regions

MSER	Maximally Stable Extremal Regions
MSVD	Multiresolution Singular Value Decomposition
MWIR	Medium Wavelength Infrared
NLP	Nombre de Longueur de Plage
ORB	Oriented FAST and Rotated BRIEF
PCA	Principal Component Analysis
PCA – SIFT	Principal Component Analysis - Scale-invariant feature transform
PFE	Percentage Fit Error
PGP	Pretty Good Privacy
PP	Pourcentage de Plage
PPP	Proportion de Petite Plage
PSNR	Peak Signal to Noise
RMSE	Root Mean Squared Error
RNA	Réseaux de Neurones Artificiels
RON	Right ON
RVB	Rouge – Vert – Bleu
SIFT	Scale-Invariant Feature Transform
SIG	Système d'Information Géographique
SSD	Sum Of Squared Differences
SURF	Speeded Up Robust Features
SUSAN	Smallest Univalve Segment Assimilating Nucleus
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TEE	Taux d'Egale Erreur
TFA	Taux de Fausse Acceptation
TFR	Taux de Faux Rejet
XCSLBP	an eXtended Center-Symmetric Local Binary Pattern

Introduction générale

La biométrie, un domaine de la vision par ordinateur, est un enjeu majeur pour les sociétés modernes confrontées à des problèmes d'identification, de contrôle et de sécurité. Les méthodes traditionnelles utilisent à cette fin des images dans le spectre du visible. Ces méthodes traditionnelles ont montré leurs limites. Et, depuis ces dernières années, face au besoin profond de sécurité, de nombreuses études ont été faites pour répondre à cette attente. Dans la littérature, ces études portent essentiellement sur des méthodes dites émergentes dont les plus utilisées sont les biométries multispectrale, multimodale et 3D.

1- Contexte et motivation

La méthode biométrique la plus utilisée est la reconnaissance de visage 2D utilisant des images couleurs [6, 48, 133]. Cependant, les images couleurs acquises dans le spectre visible sont sensibles à des paramètres tels que la variation d'éclairages, de poses, d'expressions faciales, d'occlusions, de fumée, de brouillard, etc. Ces différentes conditions difficiles détériorent considérablement les performances des systèmes de reconnaissance de visage 2D visible. Pour combler les insuffisances rencontrées par la biométrie faciale traditionnelle, les biométries 3D et multispectrale de visages se présentent comme une alternative [48]. Il est montré que les systèmes utilisant les images 3D sont robustes et insensibles à la variation d'illumination. Quant aux biométries utilisant des images infrarouges, ils sont peu sensibles aux variations d'illumination, de poses et d'expressions faciales.

Dans le but d'apporter une contribution à la biométrie faciale en générale et en particulier la biométrie multispectrale de visage, le sujet de thèse sur lequel porte nos travaux de recherche s'intitule : *extraction et fusion de point d'intérêts et texture spectraux pour l'identification, le contrôle et la sécurité.*

2- Objectif

L'objectif général de ce sujet consiste à améliorer les performances des systèmes d'identification, de contrôle et de sécurité par extraction de caractéristiques invariantes

du visage sur des images multispectrales acquises dans les spectres du visible et de l'infrarouge.

Les objectifs spécifiques qui se dégagent de cette problématique sont les suivantes :

- proposition d'une technique de fusion permettant de combiner des images acquise dans les spectres du visible et de l'infrarouge thermique afin d'obtenir une représentation qui tienne compte des avantages de chacun des spectres ;
- l'extraction de points d'intérêt et textures comme caractéristiques invariantes du visage pour la reconnaissance faciale ;
- l'étude de similarité d'images pour aider à la reconnaissance de visage ;
- le développement de technique de fusion d'images du visible et de l'infrarouge thermique ;
- l'utilisation de l'imagerie multispectrale dans la reconnaissance de visage par fusion d'images visibles et infrarouges thermiques.

3- Contribution

L'utilisation de l'imagerie multispectrale pourrait permettre d'améliorer les performances de la reconnaissance faciale classique 2D visible. L'imagerie multispectrale s'intéresse à des images de plusieurs bandes spectrales. Les expériences décrites dans cette thèse permettent d'utiliser les caractéristiques invariantes du visage dans le domaine multispectral plus précisément du visible et de l'infrarouge thermique pour la reconnaissance de visage. Les images fusionnées contiennent l'information utile, avantage des deux spectres et qui ont la qualité d'être robuste en reconnaissance de visage, même dans des conditions d'environnement difficiles

La première contribution propose une méthode d'extraction hybride de caractéristiques pour la reconnaissance faciale multispectrale. Après fusion d'images visible/infrarouge thermique, les images fusionnées obtenues sont plus riches en information que les images prises séparément. Au moyen de méthodes d'extraction de caractéristiques qui sont les points d'intérêt et la texture, une extraction de caractéristiques hybrides est utilisée pour la reconnaissance de visages.

Ensuite, la deuxième contribution est basée sur l'utilisation des méthodes récentes d'extraction de caractéristiques dites méthodes binaires et plus adaptées aux nouvelles

applications émergentes dans le multimédia, le mobile et le web. À travers cette contribution, il y a deux études menées qui sont présentées ci-dessous :

- une étude des détecteurs et des descripteurs pour l'extraction des meilleures caractéristiques;
- une application de bon compromis entre les détecteurs et les descripteurs pour l'extraction à points-clés dans la reconnaissance faciale multispectrale. Tout d'abord, les points d'intérêt sont extraits sur des images multispectrales visibles / infrarouges thermiques. Ensuite, une étape de correspondance d'image est faite pour la reconnaissance faciale.

Enfin, la troisième contribution consiste à utiliser les scores de similarité pour la reconnaissance de visage multispectrale. Dans la méthodologie utilisée à cette fin, les concepts des chapitres précédents sont repris notamment l'extraction de caractéristiques, la mise en correspondance, la carte de similarité.

Ces contributions ont donné lieu aux articles cités ci-dessous :

4- Liste des publications

- [1] MAMADOU, Diarra, GOUTON, Pierre, ADOU, Kablan J., A Comparative Study of descriptors and detectors in Multispectral Face Recognition. In: Signal-Image Technology and Internet-Based Systems (SITIS), 2016 International Conference on. IEEE, 2016. p. 209-214, 2016.
- [2] MAMADOU, Diarra, GOUTON, Pierre, ADOU, Kablan J., Multispectral face recognition using hybrid feature. Electronic imaging, Color Imaging XXII: Displaying, Processing, Hardcopy, and Applications, pp. 200-203, 2017.
- [3] MAMADOU, Diarra, GOUTON, Pierre, ADOU, Kablan J., Assessing the usefulness of similarity measures for multispectral face recognition. Electronic imaging, 2018. [accepté]

5- Organisation de la thèse

Cette thèse est organisée en quatre chapitres comme suit :

Chapitre 1 : Une revue de littérature présente sur la biométrie en générale et en particulier celle liée à la reconnaissance de visage. Les notions essentielles relatives à la reconnaissance de visage, la gamme de bandes spectrales, l'imagerie multispectrale, l'extraction de caractéristiques et la fusion d'images y sont également introduites.

Chapitre 2 : Ce chapitre présente l'imagerie multispectrale comme une alternative à l'imagerie couleur conventionnelle. L'imagerie multispectrale par la fusion d'informations des images visibles et des images infrarouges thermiques fournit de riches informations qui peuvent être utilisées dans la reconnaissance faciale. Enfin, une méthode de reconnaissance faciale multispectrale est proposée. Cette méthode se base sur une extraction hybride de caractéristiques qui sont les points d'intérêt et la texture.

Chapitre 3 : Dans ce chapitre, une reconnaissance de visage multispectrale basée sur les méthodes binaires d'extraction de caractéristiques est proposée. L'extraction de caractéristiques à l'aide des détecteurs de caractéristiques locales et des descripteurs nécessite un bon compromis entre les détecteurs et les descripteurs. Dans ce chapitre, nous avons décrit les outils BRISK, FAST, FREAK, HARRIS et SURF pour la reconnaissance faciale multispectrale.

Chapitre 4 : Ce chapitre présente une méthode de reconnaissance de visage basée sur les scores de similarité. L'analyse de similarité est un problème majeur dans la vision par ordinateur. Ce concept est désigné par un scalaire qui quantifie une mesure de ressemblance de deux objets. Cette étude présente d'abord une évaluation de certaines mesures de similarité dans la littérature. Ensuite, une fonction de similarité est proposée tenant compte des caractéristiques de l'image qui sont la texture et les points d'intérêt. Les données utilisées dans cette étude proviennent de l'imagerie multispectrale en utilisant des images infrarouges thermiques et visibles.

Chapitre 1. Etat de l'art de la reconnaissance de visage

Introduction

Le développement de systèmes biométriques connaît une évolution pour tenir les enjeux de gestion des identités, de la sécurité et du contrôle. Les systèmes traditionnels comme les cartes, les mots de passe présentent des insuffisances comme l'oubli de mot de passe, le vol de carte d'accès, etc. La biométrie se présente comme une alternative à ces systèmes traditionnels. Les systèmes biométriques sont robustes et ont l'avantage d'être basé sur des caractéristiques très représentatives telles que les propriétés biologiques, comportementales et physiologiques. Dans ces systèmes biométriques, l'identité d'une personne peut être vérifiée par les caractéristiques du visage, l'empreinte digitale, l'iris, la rétine, l'empreinte de la main, la géométrie de la main, la voix, la signature et la démarche etc., qui sont également appelés modalités ou traits biométriques [27, 52, 63].

Le principal avantage des systèmes biométriques est l'utilisation des modèles biométriques qui sont jugés être uniques et par conséquent, difficiles à falsifier.

Afin de faire face aux menaces croissantes contre la sécurité dans l'ère du numérique, les gouvernements et de nombreuses organisations s'intéressent à cet enjeu majeur de développement d'applications biométriques.

1.1. Fondamentaux de la biométrie

La biométrie a une origine lointaine qui remonte à l'ère de la reconnaissance anthropométrique. Les sociétés modernes dans un contexte d'émergence rencontrent des problèmes d'identification, de contrôle et de sécurité. Face à ces enjeux majeurs, la vision par ordinateur s'intéresse au développement d'applications biométriques. La biométrie permet d'identifier un individu par ses caractéristiques physiologiques ou comportementales. Les caractéristiques physiologiques concernées peuvent être le visage, la main, le pied, les oreilles, l'iris, les lèvres, la rétine, le doigt, etc. Quant aux caractéristiques comportementales, elles sont essentiellement la démarche, la frappe au clavier, etc. [52, 63, 65].

Les applications biométriques sont nombreuses avec pour objectif principal d'améliorer la sécurité d'accès aux systèmes (banques, aéroports, les systèmes numériques, etc.).

1.1.1. Architecture générale

Un système biométrique est généralement composé de trois modules qui sont l'enrôlement, l'identification et la vérification [27, 49, 52, 65].

Dans la phase d'enrôlement, les données biométriques de l'individu sont recueillies, prétraitées et les caractéristiques sont extraites selon un modèle. En vue d'être réutilisée lors d'une identification, les différentes données sont stockées.

Ensuite, les systèmes biométriques offrent un module d'identification à un individu qui permet de faire référence des données de cet individu à un ou plusieurs de la base de données afin de déterminer une réponse d'identification.

Enfin, dans la vérification, l'identité est réclamée, puis une correspondance du modèle de vérification avec les modèles maîtres est effectuée pour vérifier la demande d'identité de la personne. Cette vérification fait référence à une correspondance 1 à 1 ou à une authentification.

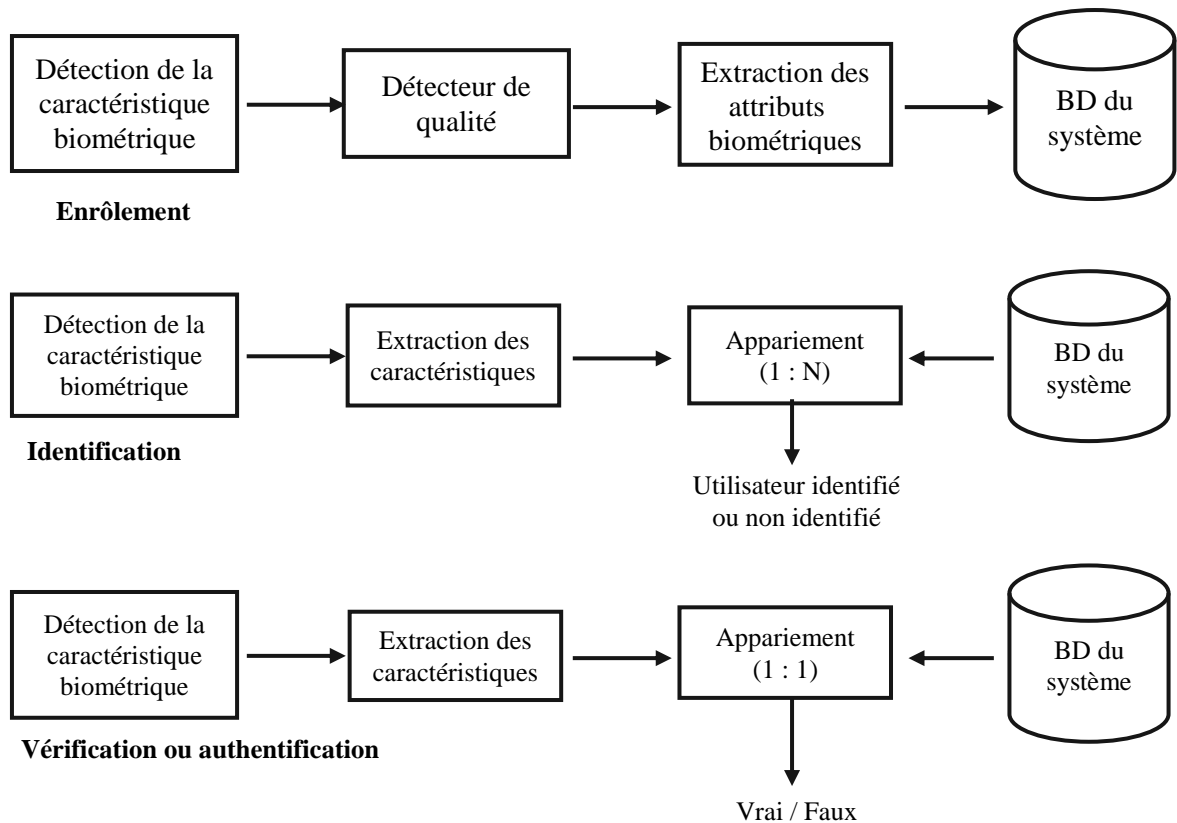


Figure (1- 1) : Différents modules d'un système biométrique [49].

1.1.2. Biométries monomodales

Les sociétés modernes intègrent le numérique dans de nombreuses activités et on parle de terminologie *smart* pour désigner l'intelligence numérique. Avec l'avènement d'internet, ce monde numérique bénéficie de nombreuses applications : email, e-learning, e-commerce, e-banking, domotique, smartphone ou tout simplement des e-services.

Cependant le développement des e-services rencontre des difficultés : attaque informatique, virus, vol informatique, pirates informatiques, accès non autorisé, vol des paramètres d'accès, etc.

Pour pallier ces insuffisances, la biométrie se présente comme un outil performant bien que tout système de sécurité peut souffrir de faille [49, 52]. Les systèmes biométriques renforcent les systèmes traditionnels utilisant des documents tels que, des clés, des codes, des mots de passe, des codes PIN, etc., [49]. L'intérêt de la biométrie est d'utiliser des propriétés des individus qui sont jugées être sensiblement uniques. Les différentes caractéristiques biométriques dans la biométrie monomodale

sont généralement les représentations numériques du visage, du doigt, de la main, du pied, des oreilles, des lèvres, la voix, la signature, la démarche, etc.

Vu que les données biométriques sont difficiles à contre-façonner et ne peuvent être empruntées, la biométrie est alors une alternative aux méthodes traditionnelles de vérification et d'authentification [52].

1.2.3. Enjeu de la biométrie

1.2.3.1. Avantages des systèmes mono-biométriques

Certaines parties du corps humain sont jugées disposer de caractéristiques uniques par individu. La biométrie utilise cette unicité de l'identité pour assurer la robustesse des systèmes de reconnaissance. Les avantages de la biométrie sont nombreux, dont [49] :

- l'augmentation de la sécurité par des techniques pratiques et peu coûteuses ;
- la réduction de la fraude en autorisant les individus ayant droit aux services ;
- l'élimination des pertes de codes d'accès et leur remplacement par des caractéristiques physiologiques ;
- la réduction des coûts de gestion de mot de passe ;
- la prévention des pertes de temps ;
- la rapidité d'identification.

1.2.3.2. Limites des systèmes mono-biométriques

Bien que les systèmes biométriques monomodaux présentent de nombreux avantages, ces systèmes souffrent de certaines insuffisances, dont [11, 22, 49] :

- le manque d'universalité de certaines caractéristiques : cas des visages, empreints, etc.
- la présence de conditions environnementales difficiles : humidité, poussière, fumée, neige, occultation, tatouage, maquillage, etc. ;
- la technologie de certains capteurs ;
- le taux d'erreur inacceptable pour certains systèmes biométriques uni-modaux ;
- la variabilité dans le temps des caractéristiques biométriques ;
- la possibilité de fraude volontairement ou involontairement, le clonage de caractéristiques biométriques.

En raison de ces problèmes mentionnés ci-dessus, les systèmes uni-modaux présentent des taux d'erreur quelques fois élevés. En résumé, chaque caractéristique biométrique a ses limites et aucun système biométrique n'est parfait [65, 133].

1.2.3.3. Différentes modalités biométriques

Il existe plusieurs modalités biométriques dans la littérature avec plusieurs présentations [27, 51]. Nous classons, ici, les modalités biométriques en trois grandes familles qui sont les modalités biométriques morphologiques, comportementales et biologiques. La figure (1-2) présente un récapitulatif des différentes modalités biométriques [52, 58, 63, 84].

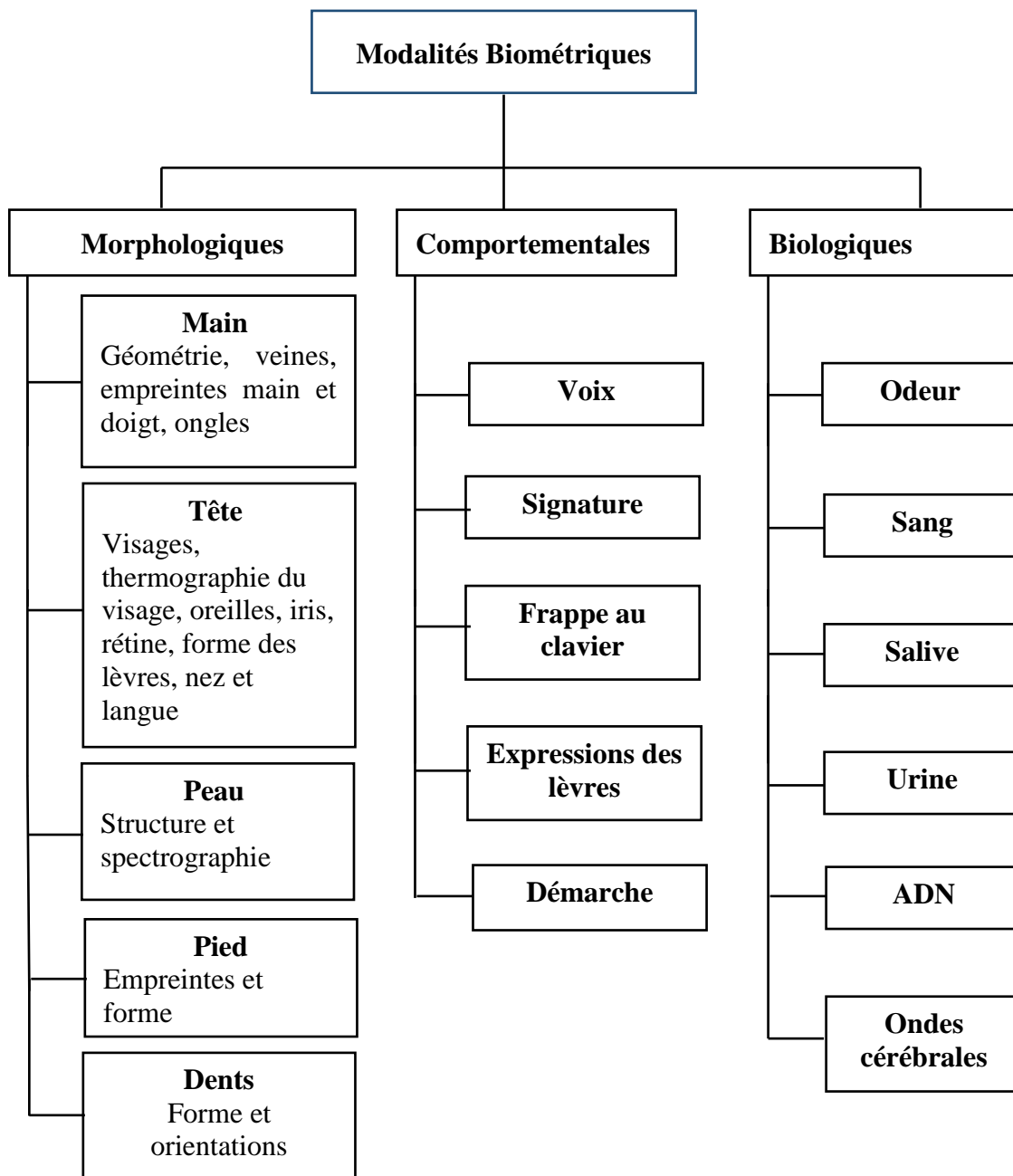


Figure (1- 2) : Récapitulatif des modalités biométriques [52, 63, 84].

1.2. Reconnaissance de visage

1.2.1. Motivation de la biométrie de visage

La biométrie faciale est l'une des techniques biométriques les plus utilisées [14, 49]. Elle est répandue, populaire et accepte l'adhésion des utilisateurs. C'est une méthode non intrusive (acquisition d'images à distance sans le consentement de

l'individu). Egalement le prix des caméras baisse au fil des ans. Une description des avantages et inconvénients de la biométrie de visages est présentée dans le tableau (1- 1) [49, 50, 63].

Tableau (1- 1) : Avantages et inconvénients de la reconnaissance de visage.

Avantages	Inconvénients
<ul style="list-style-type: none"> - Bien accepté par le public - Aucune action de l'utilisateur : peu intrusive - Pas de contact physique - Technique peu coûteuse - Plusieurs traits caractéristiques 	<ul style="list-style-type: none"> - Technologie sensible à l'environnement : éclairage, position, expression du visage, fumée, neige, etc. - Difficultés de différencier de vrais jumeaux. - Sensible aux changements : barbe, moustache, lunettes, piercing, âge, tatouage, chirurgie...)

Dans la littérature, on distingue en général trois catégories de méthodes qui sont les méthodes globales, les méthodes locales et les méthodes hybrides [118].

1.2.2. Méthodes locales

Dans les méthodes locales toute la surface du visage est considérée comme source d'information sans tenir compte des caractéristiques locales comme les yeux, la bouche, etc...La méthode la plus utilisée reste la reconnaissance de visage par ACP. Ces méthodes jugées rapides, sont basées sur des principes mathématiques tels que les statistiques et l'algèbre linéaire. Les algorithmes globaux s'appuient sur des propriétés statistiques bien connues et utilisent l'algèbre linéaire. Cependant, les méthodes locales sont sensibles aux variations d'illumination, de pose, d'occlusion et d'expression faciale. Parmi ces méthodes locales on peut citer :

- l'Analyse en Composantes Principales (PCA ou Eigen Faces) [68] : l'analyse en composante principale est une méthode statistique qui permet de simplifier un ensemble de données en réduisant sa dimension. Cette méthode dans le contexte de la reconnaissance du visage, utilise les notions de visage standard (eigenpicture) et un espace de visage engendré par des visages propres

- (eigenfaces) qui représente les composantes principales de la distribution ou les vecteurs propres (eigenvectors) de la matrice de covariance de l'ensemble des images de visages ;
- l'Analyse Discriminante Linéaire (LDA) [53] : cette méthode est utilisée pour trouver la combinaison linéaire des caractéristiques qui séparent le mieux les classes d'objet ou d'événement. Les combinaisons linéaires obtenues en appliquant cette méthode s'appellent les Fisherfaces. L'analyse discriminante linéaire est employée comme classificateur linéaire et permet la réduction de caractéristiques tout en cherchant les directions qui sont efficaces pour la discrimination entre les données ;
 - la Machine à Vecteurs de Support (SVM) [91] : la méthode SVM permet de trouver le meilleur hyperplan qui sépare au mieux les ensembles de points dans un espace de grande dimension et tout en minimisant le taux d'erreur total de classification;
 - les Réseaux de Neurones Artificiels (RNA) [96] : C'est une méthode inspirée des réseaux de neurones biologiques pour exécuter des tâches calculatoires. . En vision par ordinateur, ce système est composé d'un grand nombre d'unités de traitement hautement reliées appelées neurones, travaillant ensemble pour exécuter une tâche de classification donnée. Il a la spécificité de s'adapter, d'apprendre, de généraliser pour classer les données en entrée ;
 - le Mélange de Gaussiennes (GMM) [115] : les modèles de mélanges de gaussiennes ou Gaussian Mixture Model (GMM) permettent de modéliser la distribution de caractéristiques extraites de manière dense sur tout le visage. Cette deuxième modélisation permet de mieux gérer les variations d'apparence. Ce modèle est optimisé par l'utilisation d'un critère de maximum de vraisemblance afin d'améliorer la reconnaissance;
 - le Modèle Surfaccique du Visage (3D) [56] : Ce modèle repose sur l'utilisation d'un modèle tridimensionnel du visage. Il utilise certaines mesures sur le modèle du visage telle que la distance entre des composantes du visage (distance entre les yeux, le nez et les lèvres, etc.) ou leurs dimensions. Les informations obtenues sont utilisées pour la reconnaissance de visages.

1.2.3. Méthodes globales

Les méthodes globales également appelées méthodes à traits, géométriques, à caractéristiques locales, ou analytiques dans leur principe de fonctionnement appliquent des transformations en des endroits spécifiques de l'image qui sont des points caractéristiques localisant des régions des lèvres, du nez, des yeux, etc. Ces méthodes ont des limites dans le cas des multiples vues. Il existe deux approches différentes dont l'une procède à une extraction des régions du visage et l'autre consiste à extraire des points particuliers du visage.

Les plus utilisées de ces méthodes sont :

- les Modèles de Markov Cachés (Hidden Markov Models : HMM) [92] : cette méthode est appliquée dans la reconnaissance du visage en considérant l'information du visage comme étant une séquence variable dans le temps. Il a l'avantage de prendre en compte les caractéristiques du visage comme une forme naturelle à reconnaître. ;
- le template matching [111] : c'est une technique globale de reconnaissance faciale. La corrélation est utilisée pour mesurer la similitude entre un template (masque ou gabarit) stocké et l'image à reconnaître.

1.2.4. Méthodes hybrides

Pour un gain de robustesse, les méthodes hybrides vont fusionnées plusieurs des méthodes des approches globales et locales. Cela permet de pallier les insuffisances rencontrées dans la reconnaissance de visage comme citées précédemment. Les méthodes hybrides sont des approches qui combinent les caractéristiques holistiques et locales afin d'améliorer les performances de la reconnaissance de visage [118].

1.3. Reconnaissance de visage en infrarouge

La reconnaissance de visage en infrarouge est basée sur l'utilisation des ondes infrarouges pour mesurer les radiations thermiques émises dans la gamme du spectre infrarouge. Ces systèmes sont peu sensibles à la couleur de la peau, le bronzage, l'usage de produits de beauté, les lentilles de l'œil coloré, la chirurgie plastique, et l'obscurité.

Néanmoins, le coût des équipements reste élevé avec une baisse depuis ces dernières années [6].

La luminosité est l'un des problèmes les plus affectant pour les algorithmes de reconnaissance de visage [66]. Les variations de condition d'éclairage détériorent les performances des systèmes de reconnaissance. L'amélioration de la qualité d'image acquise est une des solutions proposées mais reste limitée. L'imagerie infrarouge se présente comme une alternative à l'imagerie traditionnelle utilisant des images couleurs. La figure (1-3) présente le spectre électromagnétique [5, 61, 113].

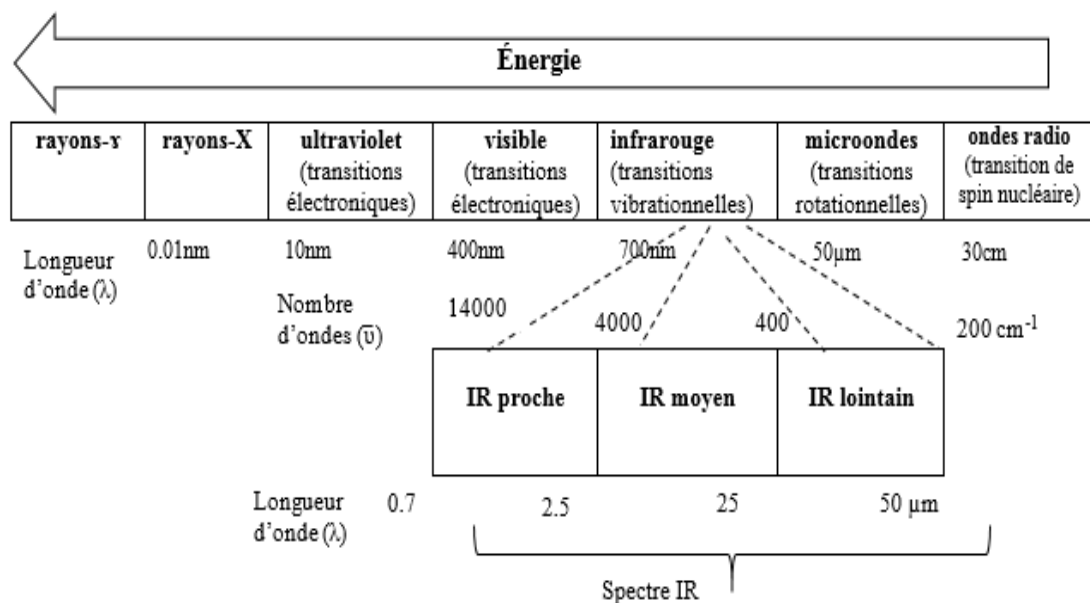


Figure (1-3) : Spectre électromagnétique.

La lumière visible utilise une petite partie du spectre électromagnétique (figure (1- 3)). Le spectre visible occupe la bande de longueur d'onde entre 0.4μm et 0.7 μm en-dessous de 0.4 μm commence le domaine de l'ultraviolet et au-delà de 0.7 μm c'est le domaine de l'infrarouge (IR). Celui-ci peut être divisé en 3 : proche infrarouge de 0.4 μm à 2.5 μm, IR moyen entre 2.5 μm et 25 μm et IR lointain entre 25 μm et 50 μm.

Plus on s'éloigne du domaine visible moins on capte les ondes infrarouges. Le proche infrarouge est réfléchi par la peau au niveau de l'épiderme alors que le visible est réfléchi à la surface de la peau. Une image en proche infrarouge contient alors moins de texture de la peau que les images en lumière visible. L'acquisition de l'IR moyen et lointain (infrarouge thermique) est faite par des capteurs thermiques [113, 117].

La capture thermique est une acquisition passive : on capte les ondes générées par la scène. Les images en infrarouge thermique ne sont pas influencées par l'éclairage, mais plutôt par la température des corps. Ce type de capture est utilisé pour la vision nocturne du fait que le corps humain émet des ondes en infrarouge moyen et lointain. Certains chercheurs exploitent cette caractéristique pour s'affranchir des conditions d'illumination et faire de la reconnaissance de visage avec des images infrarouges thermiques. Mais, il faut noter que les lunettes forment un obstacle pour les ondes thermiques [117]. Wilder et al. [127] ont montré que, dans le cas où il y a des variations d'illumination, les performances des algorithmes de reconnaissance sur images thermiques sont plus robustes que celles sur des images en visible. Mais, quand on a une variation temporelle, les images thermiques sont moins stables et les performances des algorithmes chutent [117]. Les tests de comparaison entre les images visibles et thermiques ont été multiples et effectués sur plusieurs scénarios d'acquisition. Mais les tests faits avec des images infrarouges thermiques donnent de meilleures performances que les tests faits avec les images dans le visible. Pour cela, des travaux de fusion des performances des images en visible et en thermique pour améliorer les taux de reconnaissance ont été effectués et cela a donné une amélioration par rapport aux performances de l'algorithme avec des images visibles [41, 48, 127].

Après avoir décrit les spécificités de l'infrarouge, une présentation de trois bases de données d'images infrarouges est faite dans le tableau (1-2).

Tableau (1- 2) : Présentation de base de données infrarouge [113, 35].

Base de données	Descriptions des bases de données	Références
UND Collection X1	<ul style="list-style-type: none"> - Université de Notre Dame - caméra LWIR non refroidie Merlin - haute résolution caméra couleur visible - images visibles et LWIR - acquis sur plusieurs sessions pour 82 sujets distincts - nombre total d'images : 2 293 	[44]
Base de données WSRI Dataset	<ul style="list-style-type: none"> - Institut de recherche de l'Université Wright - nombre de sujets 119 sujets - 3-visible, 3 images MWIR correspondantes - points fiduciaires annotés manuellement (œil gauche, œil droit, extrémité du nez et centre de la bouche) 	[45]
Base de données Equinox	<ul style="list-style-type: none"> - Equinox Corporation - 3 séquences de 40 images LWIR à partir de 91 personnes - Trois sources de lumières différentes: frontale, latérale gauche et latéral droit - images capturées pendant que les gens prononçaient les voyelles debout en position frontale - trois autres images de chaque personne avec le sourire, froncer les sourcils et des expressions de surprise - répété pour les personnes qui portaient des lunettes 	[43]

1.4. Reconnaissance de visage multispectrale

1.4.1. Notion d'images multispectrales

Les systèmes biométriques conventionnels utilisent soit des images en niveau de gris, soit des images couleurs.

L'imagerie multispectrale peut aider à améliorer un système biométrique traditionnel. L'imagerie multispectrale permet de mesurer la quantité de lumière absorbée par la peau et quantifier les propriétés de conduction électrique de cette peau à l'aide de capteurs de champ électrique.

Les capteurs en imagerie multispectrale capturent les données d'image à des longueurs d'ondes spécifiques du spectre électromagnétique. Ceci est contraire à l'imagerie conventionnelle qui capture trois valeurs de canal qui sont le rouge, le vert et le bleu comme présentée dans la figure (1-4).

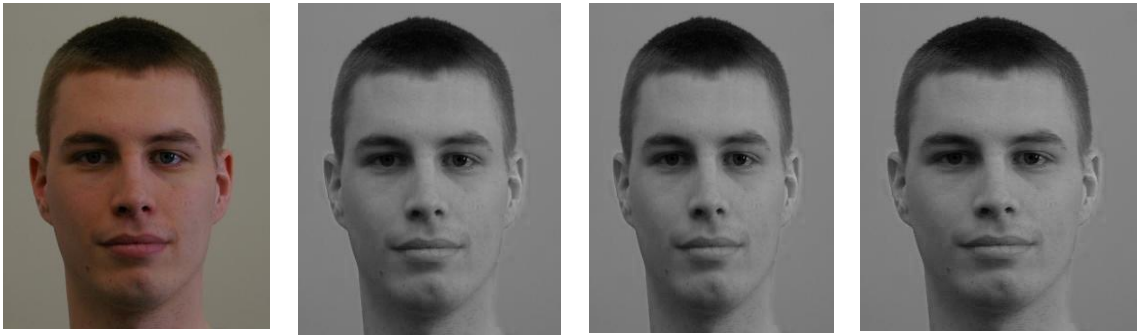


Figure (1- 4) : Ensemble d'images. De gauche à droite : image originale couleur, couche correspondante rouge, couche correspondante verte, couche correspondante bleue.

Une image en niveau de gris est une image à un canal et une image RVB est une image à trois canaux. Une image multispectrale est un multicanal c'est-à-dire plus de trois images qui peuvent être vues comme une pile d'images, chacune représentant l'image d'intensité à une longueur d'onde donnée. Le nombre de canaux en image multispectrale est en général compris entre 4 et 20 comme présenté dans la figure (1-5).

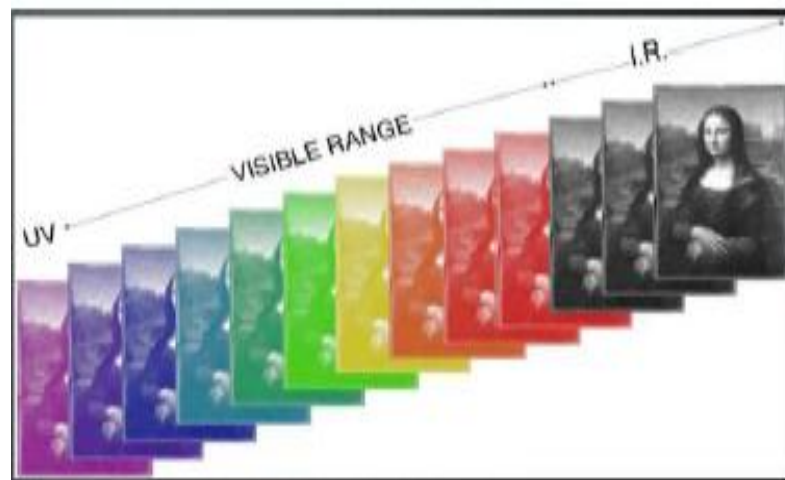


Figure (1- 5) : Exemple d'image multispectrale.

L'image multispectrale peut éviter le problème de métamérisme de l'image RVB. Deux couleurs métamères sont deux lumières visibles dont les spectres physiques sont différents, mais que la vision humaine ne différencie pas. Le multispectral fournit plus d'informations que la couleur. Il ne se limite pas à la portée visuelle, mais peut également être utilisé dans le proche infrarouge, le rayonnement infrarouge et le rayonnement ultraviolet. Ces propriétés permettent de construire des systèmes biométriques multispectraux plus efficaces [5, 11].

1.4.2. Notion de biométrie multispectrale

La biométrie multispectrale est basée sur des données comprenant plus de trois images distinctes du même trait biométrique. Ces images sont acquises à l'aide de capteurs dans différentes longueurs d'onde du spectre électromagnétique. Contrairement aux images couleurs, qui représentent une réponse intégrée d'un capteur unique sur une large gamme de bandes dans la même zone spectrale, les images multispectrales se réfèrent habituellement à des réponses multiples de différents capteurs dans des bandes spectrales relativement étroites. En télédétection, le multispectral a d'abord été introduit pour représenter des données acquises dans les spectres visibles et infrarouges. En biométrie, le multispectral est utilisé pour décrire les réponses dans plusieurs bandes étroites du visible, de l'infrarouge ou des deux. Dans la littérature, les notions d'hyperspectral et de multispectral sont utilisées de manière interchangeable, l'imagerie hyperspectrale se réfère généralement aux cas où le nombre de bandes est supérieur à 10 et lorsque ces bandes contiennent plus d'une région du spectre électromagnétique dont le visible et l'infrarouge. La figure (1.6) présente la différence entre l'hyperspectral et le multispectral.

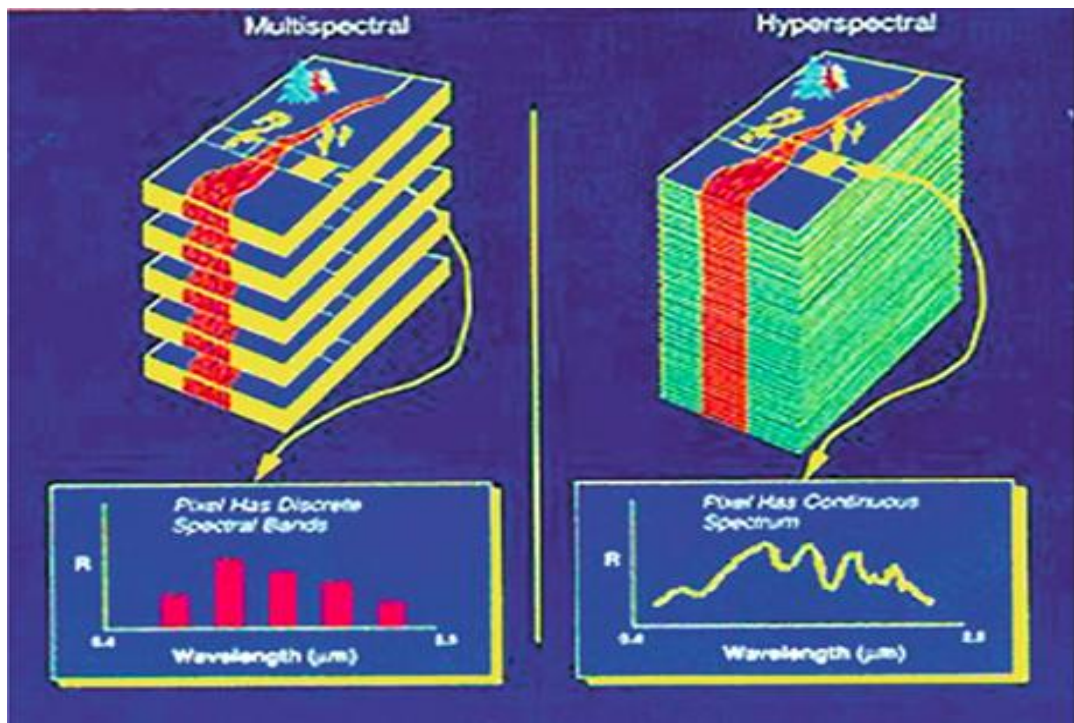


Figure (1- 6) : Différences d'hyperspectral et de multispectral [22].

1.4.3. Applications de la biométrie multispectrale

Il y a une floraison de systèmes biométriques. Ainsi, plusieurs travaux de recherches s'intéressent à la biométrie multispectrale proposant des méthodes biométriques multispectrales. Et chacune des méthodes biométriques a des forces et des faiblesses. Les plus utilisées de ses méthodes biométriques multispectrales sont données dans le tableau (1-3).

Tableau (1- 3) : Biométries multispectrales [22].

Biométrie	Références
Iris	[16]
Empreintes digitales	[22]
Paume	[108]
Main dorsal	[20]
Visage	[18]

1.4.4. Reconnaissance de visage multispectrale

1.4.4.1. Objectif

La reconnaissance de visage dans le spectre visible a montré ses limites par des problèmes tels que l'illumination, les poses, etc. Comme alternative, la reconnaissance de faciale multispectrale présente de meilleures performances par l'utilisation d'images acquises dans plusieurs bandes spectrales (visible, infrarouge, infrarouge thermique) dont un ensemble d'images est présenté dans la figure (1-7).



Image visible



Image infrarouge



Image thermique

Figure (1- 7) : Ensemble d'images. De gauche à droite : image visible, image infrarouge, image infrarouge thermique.

1.5.4.2. Principes

L'utilisation des spectres infrarouges et visibles permet de réduire les variations de couleur dans le visage occasionnées par les variations d'illuminations et de poses. Une reconnaissance faciale multispectrale a été présentée par Buddharaju et Pavlidis, en 2007 [15]. Cette étude est basée sur les informations du visible et celle de la physiologie obtenue par infrarouge thermique. La méthodologie utilisée est présentée dans la figure (1-8).

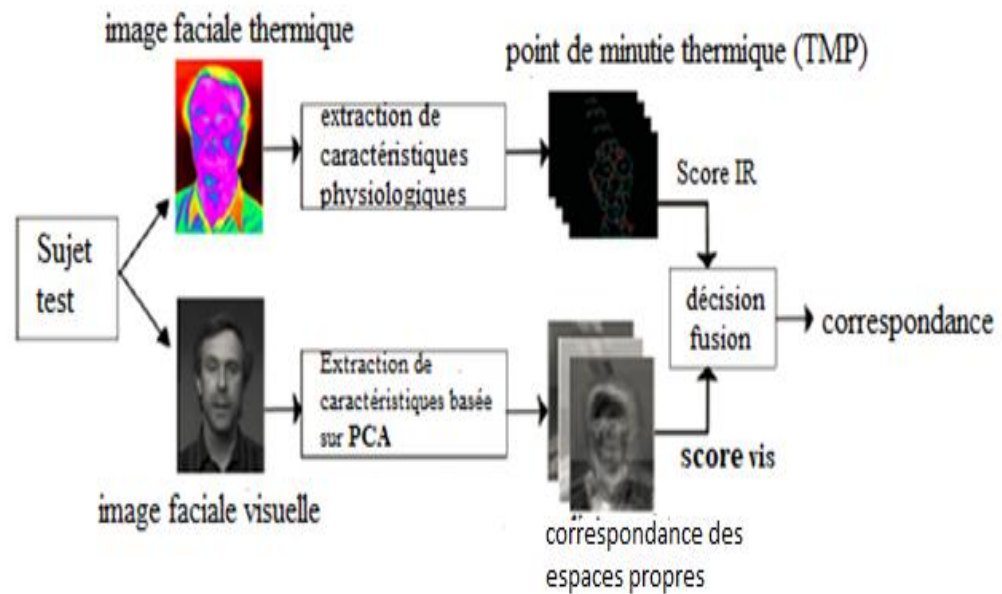


Figure (1- 8) : Méthodologie de reconnaissance de visage multispectrale [15].

L'imagerie multispectrale et l'imagerie infrarouge thermique sont vues comme un moyen pour améliorer les performances de la reconnaissance faciale dans des conditions d'illumination incontrôlée. Des travaux [6, 10, 18] ont montré l'efficacité de l'utilisation de l'infrarouge pour améliorer les performances de la reconnaissance faciale dans des conditions d'environnement difficiles. L'imagerie IR thermique peut être utilisée pour la reconnaissance de visages lorsqu'il y a peu ou pas de contrôle des conditions d'éclairage. La lumière dans la plage IR thermique est émise plutôt que réfléchi par l'objet. L'émission thermique de la peau est indépendante de l'illumination. Par conséquent, les images de visage capturées à l'aide de capteurs IR thermiques seraient presque invariables aux changements d'éclairage ambiant. La variabilité intra-classe est également nettement inférieure à celle qui a été observée dans l'imagerie visible.

Bendada et Akhloufi, ont introduit en 2010 une reconnaissance de visage multispectrale basée sur l'extraction de caractéristiques [10]. Les caractéristiques concernées sont la texture, elles ont été extraites à l'aide du descripteur local binaire (LBP).

Dans la littérature, les techniques d'évaluation de la performance reconnaissance du visage multispectrale basée sur l'extraction de la texture ont présenté des bonnes performances de reconnaissance [6, 10, 18].

1.5.4.3. Bases de données multispectrales

De nombreuses bases de données d'images multispectrales ont été construites et utilisées pour des études sur la reconnaissance de visage multispectrale. Les images issues de ces bases de données sont acquises dans le visible et l'infrarouge [5, 69].

Dans cette étude, les informations du visible et de l'infrarouge sont utilisées pour la reconnaissance de visage car chaque spectre a sa spécificité.

Il existe plusieurs bases de données multispectrales dont quelques-unes sont récapitulées dans le tableau ((1-4).

Tableau (1- 4) : Bases de données multispectrales

Base de données	Nombre de sujets	Paramètres d'études	Gamme spectrale (nm)	Références
CMU	54	Directions d'illumination et temps de latence	450-1100	[103]
IRIS-M	82	Conditions d'illumination et temps de latence	480-720	[47]
PolyU-HSFD	25	Poses et temps de latence	400-720	[132]
Standford	45	distance de visualisation	415-950	[119]
IRIS-HSFD-2014	130	Accessoires, poses, temps de latence	420-700	[100]

En plus de ces bases de données multispectrales, il en existe d'autres qui sont :

- base de données de visage "IRIS Thermal/Visible" [42] ;
- base de données infrarouge UND [44] ;
- base de données "Terravic Facial IR Database" [81] ;
- base de données PolyU-NIRFD [132] ;
- base de données Hyper-spectrale CMU [28] ;
- base de données USTC-NVIE [125] ;

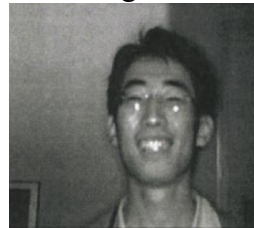
- base de données SCface [36].

Les bases de données présentées précédemment contiennent des images visibles et infrarouges acquises dans plusieurs spectres et avec des conditions de pose, d'éclairage, d'expression, de port de lunettes, etc. Des images extraites de ces bases de données sont présentes dans la figure (1-9).



Images de la base de données IRIS montrant des poses, des expressions faciales et des éclairages différents .

Images de la base de données UND montrant des expressions faciales différentes et un changement d'éclairage.



Images de la base de données Terravic montrant différentes poses et différents déguisements.

Images de la base de données PolyU-NIRFD montrant différentes poses et différentes expressions faciales.



Images de la base de données hyper-spectrale de CMU.

Image visible et infrarouge capturées simultanément.



Images d'une même personne dans la base de données SCface.

Figure (1- 9) : Ensemble d'images extraites de bases de données multispectrales [5].

1.5.4.4 Atouts et insuffisance de l'IR thermique dans la reconnaissance du visage

L'intérêt de l'infrarouge est de pallier les insuffisances de la reconnaissance faciale dans le visible. Cependant l'infrarouge présente également des insuffisances récapitulées dans le tableau (1-5).

Tableau (1- 5) : Atouts et insuffisances de la reconnaissance de visage dans le visible, l'infrarouge passif et l'infrarouge actif (+ indique un avantage et - une limitation).

Paramètres d'étude	Visible	Infrarouge passif (thermique)	Infrarouge actif
Changement d'éclairage	-	+	+
Expressions faciales	-	+	-
Température corporelle	+	-	+
Lunettes	+	-	+ -
Barbes, moustaches	+	-	
Déguisement	-	+	+ (pour le SWIR)
Usurpation, trucage	-	+	+

Des travaux récents [11,33, 77,115], ont initié la combinaison de différents modes afin d'améliorer la performance des systèmes de reconnaissance. Pour ce faire, les approches de fusion ont été utilisées pour combiner des bandes spectrales, notamment les spectres du visible et de l'infrarouge lointain. Aussi, plusieurs autres combinaisons font l'objet de recherche. Ces recherches portent sur la multimodalité telle que la combinaison 3D/infrarouge et d'autres types de combinaisons présentées dans la section suivante [11, 33].

1.5. Reconnaissance de visage et biométrie multimodale

Les systèmes biométriques monomodaux utilisent un seul trait biométrique présentant souvent des insuffisances dont les bruits, la non-universalité de certaines propriétés, la similitude, etc. font souvent face à des limitations qui influencent leurs performances [106]. Les systèmes d'identification, de contrôle et de sécurité sont victimes d'attaques dues souvent à l'exploitation de certaines de leurs lacunes. Comme alternative, les systèmes multi-biométriques sont développés pour surmonter ces problèmes rencontrés par les systèmes uni-modaux [33, 34].

Un système multi-biométrique est basé sur la multiplicité des sources d'informations biométriques. Selon la nature de ces sources, un système multi-biométrique peut être classé dans l'une des six catégories suivantes [33, 34, 77] :

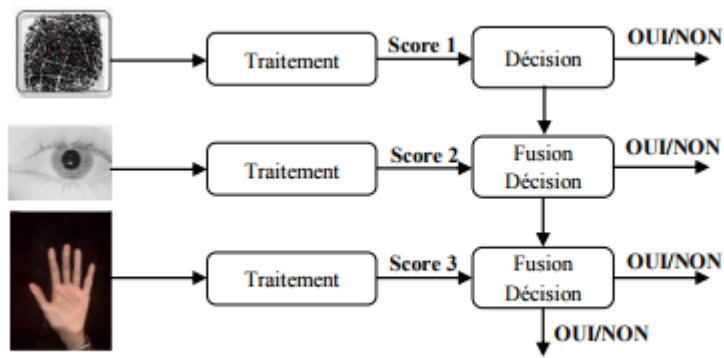
- systèmes multi-capteurs : ils utilisent plusieurs capteurs pour présenter un seul trait biométrique d'un individu ;
- systèmes multi-algorithmes : ils utilisent une combinaison de méthodes d'extraction de caractéristiques sur les mêmes données biométriques ;

- systèmes multi-instances : ces systèmes utilisent plusieurs instances du même trait biométrique;
- systèmes multi-échantillons : un capteur unique peut être utilisé pour acquérir plusieurs échantillons du même trait biométrique à des périodes données ;
- systèmes multimodaux : ces systèmes combinent plusieurs traits biométriques ;
- systèmes hybrides : le terme hybride est utilisé pour décrire les systèmes qui intègrent un module des cinq systèmes décrits précédemment.

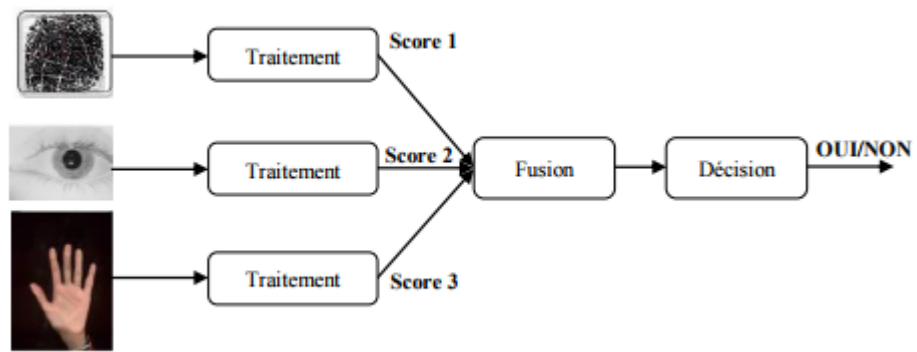
Les systèmes multi-biométriques sont classés en trois architectures selon les stratégies utilisées pour la fusion de l'information [77, 105]:

- fusion au niveau de l'extraction des caractéristiques : les informations extraites des différents capteurs sont encodées dans un vecteur de caractéristiques communes, qui est ensuite comparé à un modèle de base pour déterminer une valeur de similarité ;
- fusion au niveau de score correspondant : les vecteurs de caractéristiques sont créés indépendamment pour chaque capteur puis comparés aux modèles de base, qui sont stockés séparément pour chaque trait biométrique. Chaque module calcule son propre score correspondant. Ces scores obtenus sont combinés en un score total, qui est remis au module décisionnel ;
- fusion au niveau de la décision : une décision d'authentification distincte est faite pour chaque trait biométrique. Ces décisions sont ensuite combinées en un score final.

La figure (1-10) présente des architectures de fusion pour trois modalités biométriques qui sont l’empreinte digitale, l’iris et la main.



(a)



(b)

Figure (1- 10) : Architecture de fusion. (a) en série, (b) en parallèle.

La biométrie multimodale utilise plusieurs traits pour l'authentification. Dans la littérature plusieurs systèmes multimodaux ont été développés dont quelques-uns sont présentés dans le tableau (1–6) [37, 85].

Tableau (1- 6) : Systèmes biométriques multimodaux

Traits multimodaux	Bases de données	Références
Visage + veine du doigt	CARO	[90]
Visage + voix	UYVY. AVI	[86]
veine de la main + doigt	-	[114]
Mouvement de lèvres + gestuelle	-	[24]
Visage 2D + oreille 3D	West Virginia University Database	[46]
Visage + œil	FERET, AR database	[110]
Doigt + visage + main	-	[25]
Visage + main	-	[93]
Iris + visage	-	[32]
Visage + démarche	-	[79]

Pour surmonter les insuffisances des systèmes mono-biométriques, la biométrie multimodale se présente comme une alternative avec une meilleure sécurité et une plus grande fiabilité.

1.6. Reconnaissance de visage 3D

En général, les systèmes de reconnaissance de visages utilisent des images fixes. Pour améliorer cette technique, de nombreuses études se s'appuient sur la biométrie 3D du visage [12]. Dans le 3D, les images de visages sont capturées au moyen de plusieurs caméras, scanners ou caméras spécialisées. Le but de cette technologie est d'améliorer les performances de la reconnaissance de visage avec une qualité semblable à celle de l'empreinte digitale [12, 118]. Les avantages de la reconnaissance de visage 3D peut être synthétisés comme suit :

- insensibilité à la variation lumineuse : lorsque le sujet se présente de biais par l'utilisation de logiciel approprié un logiciel qui modifie les conditions lumineuses de l'image 3D pour les rendre plus proches de l'éclairage existant ;
- invariance de l'angle : la diversité des paramètres 3D fait que le dispositif d'enregistrement en temps réel est en mesure de reconnaître des sujets se trouvant dans un champ correspondant à une rotation de 30° de part et d'autre de son axe ;
- unicité des images 3D : la quantité de points de mesure et de données biométriques saisies est telle qu'elle permet d'identifier des jumeaux monozygotes.

L'avantage principal de la reconnaissance faciale 3D est obtenu par le fait que le modèle 3D conserve toutes les informations de la géométrie du visage, ce qui permet d'avoir une représentation réelle de ce dernier.

Cependant, les algorithmes de reconnaissance de visage en 3D sont généralement plus complexes. De plus, étant donné que les modèles de visage 3D décrivent la géométrie du visage, ils sont plus sensibles que les images 2D de texture aux expressions faciales.

Face au défi de sécurité, de nombreuses études sont faites sur les systèmes 3D. Les techniques de reconnaissance de visage 3D sont classées en deux grandes familles qui sont : l'approche de reconnaissance du visage 3D basée sur le profil et l'approche basée sur les caractéristiques géométriques faciales 3D.

Cependant, il existe des insuffisances de systèmes de reconnaissance de visage 3D qui sont les suivants:

- peu de systèmes 3D prennent en compte la variabilité des visages tout en tenant compte des positions et les expressions faciales comme le souhaite les applications biométriques ;
- peu de bases de données 3D du visage avec une variabilité importante dans la pose et les gestes.

1.7. Mesure de performance d'un système biométrique

L'utilisation de la biométrie est un enjeu majeur pour l'identification, le contrôle et la sécurité. Vu, son intérêt il est important de développer des systèmes biométriques performants afin d'évaluer cette efficacité, des métriques et normes ont été définies. Cette section présente, sept critères de robustesse [49, 52] :

- **Universalité**: la biométrie est un ensemble de caractéristiques extraites du corps humain ou du comportement. Cependant, des êtres humains peuvent perdre des traits biologiques. C'est le cas du travailleur qui peut perdre son empreinte digitale en raison du travail physique, du sourd-muet, etc.
- **unicité**: les traits biométriques d'un sujet sont uniques et peuvent être utilisés pour distinguer complètement deux personnes dans le monde. Le génotype est l'un des tests importants pour déterminer l'unicité des jumeaux identiques ayant

- la même génétique. Observant la similitude d'une biométrie dans une grande base de données est également un indicateur important de l'unicité ;
- **permanence**: des données biométriques peuvent changer au fil du temps. C'est le cas de la voix, du visage, etc. L'iris et l'empreinte digitale sont plus stables dans la durée. La permanence tient compte du fait que l'information biométrie collectée doit être présente pendant toute la vie d'un individu;
 - **collectabilité** : bien que certaines données biométriques aient une permanence élevée, un caractère unique et universel, elles ne peuvent pas être utilisées pour le public en raison de la possibilité de collecter ces données. Les données biométriques doivent être collectable et mesurable afin d'être employées pour les comparaisons. Le processus de collecte doit éviter d'être trop complexe ou employer des dispositifs coûteux. L'ADN et la rétine souffrent de ce problème car contraignants ;
 - **performance**: la Fausse Acceptation est un événement qui a lieu lorsqu'un système biométrique accepte une personne alors qu'elle n'est pas dans sa base d'utilisateurs. Cet événement doit être le plus rare possible dans un système biométrique sécurisé. Le Faux Rejet est un événement qui a lieu lorsqu'un système biométrique refuse une personne alors qu'elle est dans sa base d'utilisateurs. Cet événement est souvent dû à une mauvaise acquisition des données biométriques et est perçu comme une gêne par l'utilisateur. Le TFA (Taux de Fausse Acceptation) est la probabilité qu'un utilisateur inconnu soit identifié comme étant un utilisateur connu. Ce taux définit la sécurité du système biométrique. Le TFR (Taux de Faux Rejet), c'est la probabilité qu'un utilisateur connu soit rejeté par le système biométrique. Ce taux définit en partie le confort d'utilisation du système biométrique. Enfin, le TEE (Taux d'Egale Erreur) donne un point pour lequel le TFA et le TFR sont égaux ;
 - **acceptabilité** : système de biométrie convivial et non contraignant pour les sujets. En fait, presque tous les systèmes biométriques actuels ne sont pas physiquement intrusifs pour les utilisateurs, mais certains d'entre eux le sont tels que la biométrie rétinienne ;
 - **circonvention** : le terme contournement se réfère à la facilité pour contourner un système biométrique en utilisant un artefact ou un substitut. Par exemple, lors

d'une attaque par circonvension, le fraudeur peut accéder à des ressources protégées de l'application telles que les données personnelles d'un autre utilisateur.

Conclusion

Dans ce chapitre, nous avons présenté les technologies utilisées dans les systèmes biométriques pour l'identification de personnes. Cette étude a permis de constater que la reconnaissance de visage suscite de plus en plus un intérêt pour la communauté scientifique ; car elle présente encore plusieurs challenges. Les difficultés liées à chaque type de système biométrique ont été mis en exergue. Ces difficultés sont essentiellement l'invariance à l'illumination, aux poses, fumée, brouillard, l'occlusion, les expressions faciales, etc. Aussi, un aperçu sur les méthodes biométriques est fait avec les différentes métriques de performance. Après cette présentation introductive, les chapitres 2, 3 et 4 analyseront les méthodes de reconnaissance de visage multispectrale basée respectivement sur une extraction hybride de caractéristiques, une extraction de caractéristiques à l'aide de descripteurs binaires et l'utilisation de score de similarité pour l'identification.

Chapitre 2. Reconnaissance de visage multispectrale basée sur une extraction hybride de caractéristiques (MS-FRHF)

Introduction

La reconnaissance de visage traditionnelle a montré des insuffisances. Les images infrarouges thermiques ont des signatures utiles qui sont insensibles aux variations d'illuminations et de directions d'observation. L'imagerie multispectrale par la fusion d'informations des images visibles et des images infrarouges thermiques fournit de riches informations de données qui peuvent être utilisées dans la reconnaissance faciale. Comparativement à la reconnaissance faciale traditionnelle, l'imagerie multispectrale peut séparer les informations d'illumination et de réflectance des images de visages. Aussi, la méthode de reconnaissance de visage dans cette étude est basée sur une extraction hybride de caractéristiques qui sont la texture et les points d'intérêt dans des images fusionnées visibles/ infrarouges thermiques. Dans ce chapitre, après avoir fait une revue des techniques d'extractions de caractéristiques les plus courantes, nous présentons la méthode MS-FRHF. Cette méthode sera testée sur une base d'images multispectrales. L'utilisation de la fusion des images visibles et infrarouges thermiques dans la reconnaissance faciale montre de meilleures performances que l'imagerie traditionnelle [41].

2.1. Notion d'extraction de caractéristiques

Les caractéristiques d'une image sont des éléments permettant d'écrire une image de façon précise et dont l'intérêt dépend de l'étude caractéristique. Ces centres d'intérêt peuvent être portés sur des points d'intérêt, des régions d'intérêt, des coins, des lignes, la couleur, les bords, les blobs, la texture, etc. Ces différentes caractéristiques ont l'avantage d'avoir une signification sémantique riche en information, stable aux transformations géométriques.

Dans cette étude les caractéristiques utilisées sont la texture et les points d'intérêt avec sa spécificité d'analyse.

2.1.1. Points d'intérêt

Dans une image, un point d'intérêt peut correspondre à des doubles discontinuités de la fonction d'intensité. Ces discontinuités peuvent être représentées par des coins, des régions, des jonctions ou des points de fortes variations de texture. Ces points d'intérêt ayant des propriétés particulières sont souvent utilisés dans la reconnaissance de visage [118].

La détection de points d'intérêt est une étape préliminaire à de nombreux processus de reconnaissance dans le domaine de la vision par ordinateur. Les différents points extraits sont souvent utilisés dans les processus de reconstruction d'images, de suivi d'objets, de recherche d'images, de reconnaissance de mouvements, de reconnaissance de visage, etc.

Les points d'intérêt [19] dans la reconnaissance sont largement utilisés car ils interviennent dans plusieurs applications dont la biométrie, le contrôle qualité, les systèmes d'informations géographiques (SIG), la médecine, la robotique, etc.

Les points d'intérêts, dans une image, représentent des doubles discontinuités de la fonction d'intensités. Celles-ci peuvent être produites par des discontinuités de la fonction de réflectance ou des discontinuités de profondeur.

L'utilisation des points d'intérêt présente plusieurs avantages. D'abord ces points représentent une source d'information fiable due à la fonction d'intensité. Ensuite, ils sont jugés robustes car invariants aux transformations, localisés clairement. Enfin, de nombreuses images contiennent des points d'intérêt.

2.1.2. Texture

La reconnaissance visuelle d'une texture se montre aisée mais sa définition reste difficile. Dans la littérature, plusieurs définitions sont données suivant le contexte d'étude [72, 74].

Au travers de nombreux travaux, différentes définitions sont proposées en fonction du domaine d'application de leurs auteurs [74]. De façon succincte, cette définition présente la texture comme une région macroscopique dont la structure est

composée de motifs répétitifs dans lesquels les composants élémentaires (texels) ou les primitives sont disposés selon une règle de placement.

Bien que l'analyse de texture soit importante, il n'y a pas de définition précise de la texture et il en existe plusieurs dans la littérature [74]. Cette diversité de définitions est due au fait que chacune d'elle tient compte de critères spécifiques. Les textures se présentent comme des motifs visuels complexes composés d'entités ou de petits motifs ayant des caractéristiques telles que la luminosité, la couleur, la pente, la granularité, la taille, etc. Ainsi, la texture peut être considérée comme un groupement de petits éléments similaires dans une image. Les propriétés locales des petits motifs peuvent être la légèreté, l'uniformité, la densité, la rugosité, la régularité, la linéarité, la fréquence, la phase, l'orientation, la grossièreté, le caractère aléatoire, la finesse, la granularité, etc. Des exemples de textures sont présentés dans la figure (2-1).

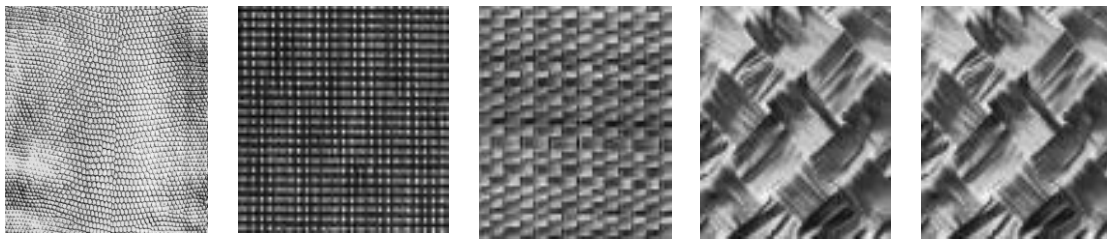


Figure (2- 1) : Quelques exemples de textures [13]

2.2. Méthodes d'extraction de caractéristiques

2.2.1. Extraction de points d'intérêt

Dans la reconnaissance d'objets, la détection de points d'intérêt d'une image met en évidence des zones de cette image jugée pertinentes en vue d'une analyse. Ces points présentent des propriétés locales remarquables. D'abord, les mêmes caractéristiques doivent être présentes dans l'image après une transformation géométrique. C'est une propriété d'invariance que les points d'intérêt doivent être robustes aux transformations. Aussi, ces zones d'intérêt doivent être riches en information et localisées de façon précise et ne peuvent varier d'une scène à une autre.

2.2.2. Méthodes d'extraction de points d'intérêt

Cette partie présente un aperçu des méthodes utilisées dans la vision par ordinateur pour la reconnaissance d'objets. Deux outils sont importants dans différentes pratiques de reconnaissance. Ils sont nommés détecteur et descripteur. La qualité du résultat d'un système de reconnaissance dépend du choix du détecteur et du descripteur car chaque méthode est spécifique en fonction d'un contexte donné. Ensuite, il est important de connaître les performances de ces outils, comme l'indique K. Mikolajczyk [54]. Il est important d'associer à un détecteur, un descripteur approprié.

Le développement et l'amélioration des systèmes de reconnaissance constituent un axe d'étude majeure dans les études récentes. De nombreuses méthodes ont été développées dans la reconnaissance d'objets. Les méthodes d'extraction de caractéristiques utilisées sont basées sur des algorithmes invariants aux transformations. Dans la dernière décennie l'algorithme le plus utilisé est Scale Invariant Feature Transform (SIFT) qui a eu plusieurs variantes dont Affine-SIFT (A-SIFT). Ces algorithmes s'utilisent dans plusieurs contextes comme l'imagerie multispectrale, la reconnaissance de visage sous les critères de variation de l'angle de pose, la variation d'illumination, d'expression du visage, de sorte que les performances d'un tel noyau d'extraction de caractéristiques soit capable d'extraire les paramètres.

2.3.2.1. Détecteurs

Un point d'intérêt dans une image est une zone de pixels aux propriétés remarquables souvent exprimées par des changements brusques d'intensité. Dans certains cas, il peut s'agir aussi d'une vaste zone sans changement d'intensité ou d'une zone à forme variée comme des points, des courbes, des régions, etc.

Les techniques de descriptions locales de contenu visuel utilisant des points d'intérêt sont supportées par deux outils que sont les détecteurs et les descripteurs. L'utilisation de ces deux outils complémentaires permet d'obtenir une description robuste des images dans le domaine de la reconnaissance d'objets. Cependant, l'un des problèmes de la détection de points-clés reste l'existence de diverses transformations dans l'image. L'efficacité d'une reconnaissance d'objets nécessite le bon choix de détecteurs et de descripteurs invariants aux transformations.

Les détecteurs sont utilisés pour isoler les zones d'intérêt dans une image. Ces zones d'intérêt peuvent être les coins ou les régions. Le coin est un changement brusque d'une caractéristique au voisinage d'un pixel. Aussi, la région est définie comme une zone de pixels adjacents qui satisfait une certaine propriété souvent exprimée en terme d'homogénéité, de couleur, de saturation, etc.

Au nombre des détecteurs de coins, il y a plusieurs dont quelques-uns sont présentés dans le tableau (2.1)

Tableau (2- 1) : Résumé de quelques détecteurs de coins

Détecteurs	Références
Moravec	[38]
Harris	[17]
SUSAN(Smallest Univalu Segment Assimilating Nucleus)	[73]
Harris couleur	[87]
Detector Harris Laplace	[55]
Detector Hessian Laplace	[55]
Detector Fast Hessian	[8]
Blob	[67]
SIFT (Scale Invariant Feature Transform)	[82]
SURF (Speeded-Up Robust Features)	[7]

Le détecteur de Moravec [38], détermine le changement d'intensité moyenne dans une région donnée de l'image. Du point de vue amélioration, le détecteur de Harris [17] est une variante du détecteur de Moravec qui tient compte de nouvelles invariances comme la correction de l'anisotropie, la correction de la forte réponse des côtés. Des détecteurs simplifient l'analyse du voisinage, l'optimisation des temps de calculs et le traitement afin d'utiliser de façon optimale la fonction d'intensité de l'image. Dans ce type de détecteur, il y a SUSAN [73](Smallest Univalu Segment Assimilating Nucleus) qui utilise une analyse locale circulaire pour la détection de point.

Pour des applications sur des images couleurs, le détecteur de Haris couleur [87] est invariant aux changements colorimétriques. Il existe une autre variante du détecteur de Harris nommé le détecteur Harris Laplace [55] qui combine le détecteur de coin Harris et la fonction de Laplace invariant aux changements d'échelle. Aussi, le détecteur de Hessian Laplace [55] est une variante rapide du détecteur de Haris Laplace.

Bien connu, le détecteur Fast Hessian [8] est rapide. Il est presque identique au détecteur Harris et est utilisé dans SURF [7]. De même pour le détecteur Harris, le détecteur de Hesse choisit des points d'intérêt au moyen d'un algorithme itératif à plusieurs échelles.

Concernant le détecteur de Blob [67], il applique la convolution pour détecter les régions dans une image, telles que la luminosité ou la couleur, par rapport aux zones

entourant ces régions. Un Blob est défini comme une région d'une image numérique dans laquelle certaines propriétés sont constantes ou varient dans une plage de valeurs prescrite ; tous les points d'un blob peuvent être considérés dans un certain sens comme similaires.

Le détecteur, SIFT [82] est le plus connu des détecteurs. Cette méthode combine un détecteur avec un descripteur. Dans son processus, il cherche les maxima locaux sur un ensemble de différences d'images gaussiennes. La détection de point d'intérêt de SIFT est basée sur un DoG (Difference of Gaussian). Il a l'avantage d'être stable à la translation, à la rotation, à la mise à l'échelle, à l'invariance par rapport aux changements de l'éclairage et au point de vue de la caméra.

Enfin, SURF est inspiré de SIFT et utilise quelques approximations pour augmenter la vitesse des traitements. Cette méthode est basée sur la matrice de Hesse qui résulte du produit de la convolution d'un gaussien, avec la matrice des dérivées secondes. À l'aide de matrice 2x2 et de matrice des dérivées partielles d'ordre deux, des calculs de surface de chacun des termes sont faits. Le calcul de ces surfaces est accéléré par la construction précédente de l'image complète. Pour le changement d'échelle, ce ne sont plus les nouvelles dimensions comme pour SIFT, mais le masque sur lequel est calculée la matrice de Hesse.

Au nombre des détecteurs de régions, il y a plusieurs dont quelques-uns sont présentés dans le tableau (2.2)

Tableau (2- 2) : Résumé de quelques détecteurs de coins

Détecteurs	Références
Harris Affine and Hesse Affine	[55, 83]
EBR (Edge-Based Region)	[122]
IBR (Intensity-Based Region)	[122]
MSER (Maximally Stable Extremal Regions)	[104]
MSCR (Maximally Stable Colour Regions)	[95]

Le détecteur Harris Affine et Hesse Affine [55] commencent par déterminer les points d'intérêt avec le détecteur de Harris-Laplace ou Hesse-Laplace. Cette détection a besoin de quatre étapes d'itérations pour chaque point détecté. Concernant le détecteur EBR (Edge-Based Region) [122], il combine le détecteur des angles de Harris avec le

détecteur des côtés de Canny pour construire des régions d'invariance en parallélogrammes. Quant au détecteur IBR (Intensity-Based Region) [122], il construit des régions elliptiques tout en étant basé sur l'intensité de l'image.

Concernant le détecteur de régions extrêmement stables nommé MSER (Maximally Stable Extremal Regions) [104], la détection des zones d'intérêt repose sur le principe de la ligne de partage des eaux. Il détecte des régions qui restent stables. Le détecteur MSCR (Maximally Stable Colour Regions) [95] est une extension du détecteur MSER aux images couleurs.

2.3.2.2. Descripteurs

Le descripteur a pour but de décrire les points d'intérêt obtenus par un détecteur. Is analyse le voisinage proche de chaque point, le support de l'extraction, pour produire un vecteur caractéristique, de cette zone. Ce vecteur est appelé vecteur descripteur. Le principe de descripteur est de rendre certains vecteurs de description uniques pour chaque point d'intérêt.

Deux aspects sont importants et indispensables pour un descripteur : l'invariance et la discrimination. L'invariance consiste à décrire un seul point d'intérêt de la même manière dans deux différentes images. Un descripteur peut être invariant aux transformations géométriques, photométriques, au bruit, etc. Quant à la discrimination, elle consiste à distinguer deux différents points d'intérêt dans deux images distinctes. La qualité d'un descripteur dépend du choix des notions d'invariance et de discrimination.

Au nombre des descripteurs, il y a plusieurs dont quelques-uns sont présentés dans le tableau (2.3)

Tableau (2- 3) : Résumé de quelques détecteurs de régions

Descripteurs	Références
Descriptor DAISY	[31]
GLOH (Gradient Location and Orientation Histogram)	[54]
HOG (Histogram of Oriented Gradients)	[23]
SIFT (Scale Invariant Feature Transform)	[82]
SURF (Speeded-Up Robust Features)	[7]
ASIFT (Affine-SIFT)	[88]
CSIFT (Color-SIFT)	[3]
PCA-SIFT	[57]
GSIFT	[89]

Les filtres ajustables peuvent être utilisés comme descripteur, il s'agit d'une combinaison linéaire de filtres de base pour déterminer l'orientation d'un point d'intérêt. Le descripteur HOG (Histogram of Oriented Gradients) [23] utilise une grille dense pour calculer les histogrammes locaux d'orientation. Quant au descripteur GLOH (Gradient Location and Orientation Histogram)[54], il améliore les performances de SIFT tout en construisant un histogramme de gradients orientés dans un plan circulaire. Le descripteur DAISY [31] est une nouvelle approche de description locale utilisant les avantages des SIFT et GLOH afin d'accélérer les temps de calculs et d'améliorer la gestion des invariances.

SIFT (Scale Invariant Feature Transform) est une méthode qui combine un détecteur avec un descripteur. Lors de la phase de détection, il y a une détection des extrema afin de localiser les points caractéristiques. Ensuite, une élimination des points de faible contraste et des points sur des contours de faible courbure est effectuée. Enfin, une assignation d'orientation est faite pour la description des points d'intérêt.

Dans une étude comparative sur SIFT et de ses variantes, J. Wu [128] évalue leur performance dans différentes situations : changement d'échelle, changement de rotation, changement de flou, changement d'illumination et changement affine. Ces variantes décrites sont SURF (Speeded-Up Robust Features), ASIFT (Affine-SIFT) [88], CSIFT (Color-SIFT) [3], PCA-SIFT [57], GSIFT [89]. Une présentation de cette famille de descripteurs est présentée dans la figure (2-2).

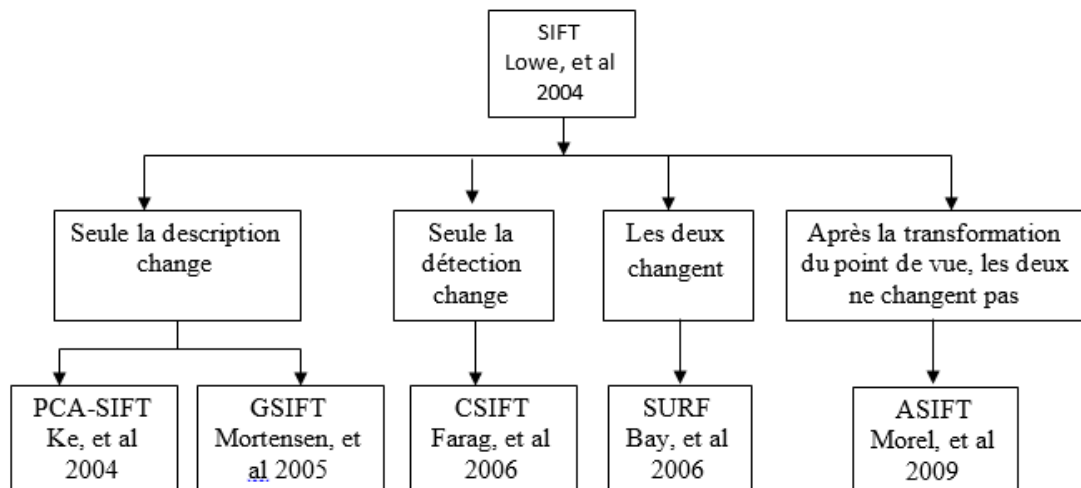


Figure (2- 2) : Présentation de SIFT et ses variantes [128]

2.2.3. Extraction de texture

L'analyse de texture peut permettre de quantifier la finesse, le contraste, les variations d'intensité, les discontinuités et les contours d'une image [102]. Compte tenu de la complexité de la reconnaissance faciale, il est judicieux d'associer l'analyse de texture à d'autres caractéristiques comme les points d'intérêt, la forme, etc. La combinaison de caractéristiques s'avère importante au vu de la complexité de la structure du visage tenant compte de la peau, les rides, les plis, les cicatrices, les poils, etc.

2.3.3.1. Extraction de texture basée sur les statistiques du premier ordre

Les statistiques du premier ordre se basent sur l'histogramme des niveaux de gris de l'image en le considérant comme une densité de probabilité. Ces statistiques sont utilisées pour l'analyse de texture. Cette méthode a l'avantage d'être simple et permet de caractériser les données. Cependant, cette approche présente des insuffisances du fait qu'elle ne tient pas compte des relations spatiales et des corrélations entre les pixels.

Par conséquent, elle n'est pas adaptée pour analyser tout type de texture. Deux images ayant les mêmes histogrammes peuvent être différentes du point de vue texture.

On utilise souvent, sept mesures statistiques pour décrire les propriétés d'un histogramme et par conséquent la texture d'une image. Ces mesures sont la moyenne, la variance, la rugosité, le skewness, le kurtosis, l'énergie et l'entropie dont les plus utilisés sont définies respectivement par les formules (1), (2), (3), (4), (5).

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

$$\text{var}(X) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 \quad (2)$$

$$\text{std}(X) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2} \quad (3)$$

$$\text{skewness}(X) = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^3}{s^3} \quad (4)$$

Où \bar{X} est la moyenne, s l'écart-type, n le nombre de données

$$\text{kurtosis}(x) = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^4}{s^4} \quad (5)$$

Où \bar{X} est la moyenne, s l'écart-type, n le nombre de données

Bien qu'insuffisantes pour décrire une texture à elles seules, ces mesures sont néanmoins très utilisées en reconnaissance de visage en combinant d'autres caractéristiques de plus haut niveau.

2.3.3.2. Extraction de texture basée sur les statistiques du second ordre

Les statistiques du second ordre mesurent la probabilité qu'une paire de pixels régie par une certaine relation qui existe dans l'image.

Les matrices de cooccurrence en niveaux de gris [80] nommées Grey Level Co-occurrence Matrix (GLCM) déterminent les propriétés d'une texture, mais elles ne sont pas directement utiles pour une analyse plus approfondie, telle que la comparaison de deux textures. Les fonctionnalités numériques sont calculées à partir de la matrice de cooccurrence qui représente la texture de manière plus compacte. Ces éléments caractéristiques standards dérivés d'une matrice de cooccurrence normalisée sont l'entropie, le contraste, la corrélation, l'énergie et l'homogénéité.

$$H_e = - \sum_{i=0}^L h_{I_f}(i) \log_2 h_{I_f}(i) \quad (6)$$

$$\text{Contraste} = \sum_{i,j} |i - j|^2 p(i, j) \quad (7)$$

$$\text{Corrélation} = \sum_{i,j} \frac{ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (8)$$

$$\text{Energie} = \sum_{i,j} p(i, j)^2 \quad (9)$$

$$\text{Homogénéité} = \sum_{i,j} \frac{1}{1 + (i-j)^2} p(i, j) \quad (10)$$

2.3.3.3. Statistiques d'ordre supérieur

Dans la famille des statistiques d'ordre supérieur, il y a la méthode dite longueur de plage. La plage se définit comme un ensemble de pixels consécutifs dans une orientation donnée et ayant le même niveau de gris. Ensuite la longueur de plage est le nombre de pixels composant une plage.

L'analyse de textures dans cette méthode est basée sur un ensemble de matrices $P_\theta = (p_\theta(i, j))$ où chaque élément (i, j) de la matrice P_θ contient le nombre de plages de longueur j et de niveau de gris i dans la direction θ . Pour l'extraction de textures, on considère L le nombre de niveaux de gris et n_θ la longueur maximale d'une région donnée et dans une orientation θ . Les caractéristiques de textures extraites sont décrites dans la suite.

Nombre de longueurs de plage (NLP), proportion de petites plages (PPP), proportion de grandes plages (PGP), hétérogénéité des niveaux de gris (HNG), hétérogénéité des longueurs de plage (HLP), pourcentage de plages (PP) sont décrites respectivement par les formules (11), (12), (13), (14), (15), (16).

$$NLP = \sum_{i=0}^{L-1} \sum_{j=1}^{n_\theta} p_\theta(i, j) \quad (11)$$

$$PPP = \frac{1}{NLP} \sum_{i=0}^{L-1} \sum_{j=1}^{n_\theta} \frac{p_\theta(i, j)}{j^2} \quad (12)$$

$$PGP = \frac{1}{NLP} \sum_{i=0}^{L-1} \sum_{j=1}^{n_\theta} j^2 p_\theta(i, j) \quad (13)$$

$$HNG = \frac{1}{NLP} \sum_{i=0}^{L-1} \left(\sum_{j=1}^{n_\theta} p_\theta(i, j) \right)^2 \quad (14)$$

$$HLP = \frac{1}{NLP} \sum_{j=1}^{n_\theta} \left(\sum_{i=0}^{L-1} p_\theta(i, j) \right)^2 \quad (15)$$

$$PP = \frac{NLP}{K} \quad (16)$$

Où K désigne le nombre total de pixels dans la région.

La méthode longueur de plage est utilisée pour la caractérisation des textures de régions avec de bons résultats. Cependant, elle est lente et appropriée pour les images en niveaux de gris [101].

2.3.3.4. Filtrés de Gabor

Les filtres de Gabor peuvent être utilisés pour l'analyse de texture d'une image tout en détectant des régions ayant une dimension et une orientation [76]. Ces filtres initialement utilisés pour le traitement du signal en dimension 1 ont été étendu par Daugman [26] en dimension 2. Ils interviennent dans de nombreuses applications de traitement d'image. Théoriquement, ils ont la forme d'un noyau gaussien modulé par une fonction sinusoïdale plane. La fonction G d'un filtre de Gabor est définie par la formule (17).

Pour tout couple (x, y) d'entier,

$$G(x,y) = \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \times \cos(2\pi\omega x + \phi) \quad (17)$$

où σ_x et σ_y sont les largeurs du noyau gaussien selon les directions respectives x et y , ω et ϕ sont respectivement la fréquence et la phase de la fonction sinusoïdale.

L'extraction de texture se fait en appliquant une série de filtres de Gabor sur une image en prenant en compte toutes les orientations, les échelles et les fréquences recherchées [26].

2.3.3.5. Local Binary Pattern (LBP)

Le descripteur LBP (Local Binary Pattern) peut être utilisé pour l'extraction de texture [94]. Il utilise une fenêtre 3×3 pixels pour le calcul des valeurs tenant compte de la formule (18).

$$LBP(x_0, y_0) = \sum_{i=0}^n 2^i s(i_n - i_0) \quad (18)$$

avec $s(u) = 1$ si $u > 0$ et 0 sinon, (x_0, y_0) les coordonnées du point où on calcule le descripteur, i_0 la valeur de ce point et les i_n parcourant le voisinage.

Il existe des variantes du descripteur LBP qui sont CLBP (Complete Local Binary Pattern), CSLBP (Center-symmetric local binary patterns) et XCSLBP (extended CSLBP).

Un descripteur LBP est invariant par une variation monotone de la valeur des pixels, ce qui lui permet de résister aux variations d'illumination.

2.3.2.6. Méthode fractale

Face à la complexité de modélisation des textures, la théorie des fractales a été proposée comme alternative à la géométrie euclidienne et les modèles usuels [78]. En théorie, les fractales représentent l'irrégularité d'une figure géométrique telle qu'une surface ou un volume. Cette méthode a l'avantage de décrire des textures tout en supposant que les motifs répétitifs ont une petite dimension fractale alors que les images stochastiques ont une grande dimension fractale.

2.3. Méthode MS-FRHF proposée pour l'extraction de caractéristiques en reconnaissance de visage multispectrale

Les images infrarouges thermiques ont des signatures utiles qui sont invariantes aux différentes illuminations et directions d'observation. L'imagerie multispectrale par la fusion d'informations des images visibles et des images infrarouges thermiques fournit des images riches en informations et qui peuvent être utilisées dans la reconnaissance faciale. Comparativement à la reconnaissance faciale traditionnelle, l'imagerie multispectrale peut séparer les informations d'illumination et de réflectance des images faciales [6, 98].

2.3.1. Objectifs

Dans l'imagerie multispectrale, plusieurs longueurs d'onde du visible et de l'infrarouge fournissent des informations utiles pour la reconnaissance faciale [41, 98]. D'une part, les images visibles fournissent des informations sur la réflectance de la peau du visage, et d'autre part, les images infrarouges thermiques produisent des caractéristiques uniques en raison de l'émissivité du rayonnement thermique des visages. L'intérêt principal des images infrarouges thermiques est de surmonter les faiblesses dues à plusieurs difficultés telles que : la fumée, la neige, la poussière, l'éclairage extrême, l'obscurité, les différentes poses de visage, le port de lunettes et l'expression faciale. Certaines expériences [6, 18] ont été réalisées en utilisant la base de données IRIS [42] qui contient des images de visage spectrales. En effet, dans ces expériences, l'imagerie multispectrale par la fusion d'images a permis d'obtenir des images composites en combinant des informations du visible et de l'infrarouge thermique. Ainsi, les images fusionnées fournissent des informations supplémentaires

avec la réflectance du visage et le rayonnement pour améliorer les méthodes traditionnelles de reconnaissance faciale.

Dans la méthode que nous proposons est fondée sur une reconnaissance faciale multispectrale utilisant la fonction hybride appelée MS-FRHF (Multispectral Face Recognition using Hybrid Feature). Après fusion des images multispectrales visibles / infrarouges thermiques, nous utilisons la fonction hybride de reconnaissance faciale basée sur l'extraction des points d'intérêt et de textures pour la reconnaissance faciale.

2.3.2. Reconnaissance de visage dans le visible et l'infrarouge thermique

L'imagerie multispectrale améliore la capacité et la fiabilité de l'interprétation de l'image. Cependant, plusieurs difficultés existent. Les problèmes peuvent être liés à la multiplicité des images, des capteurs, des séquences d'images, des images multi-dates, de la complexité des scènes. Pour atténuer ces difficultés, la fusion d'images est un élément important de la gestion de grands volumes d'images à partir de capteurs uniques, de multi-capteurs et de multimodalité. La fusion d'images permet de combiner des informations provenant de plusieurs images acquises par différentes sources. Compte tenu de son succès, la fusion d'images est utilisée dans de nombreux domaines de la vision par ordinateur, y compris la reconnaissance faciale [18, 41, 98]. Spécifiquement, la fusion d'images permet de combiner les informations pertinentes de plusieurs images afin d'obtenir une image plus riche d'informations ne contenant que des informations utiles. Cette information a un contenu riche en caractéristiques selon les spécificités de l'étude.

Notre étude se concentre sur la reconnaissance multispectrale du visage à l'aide d'images visibles et d'images infrarouges thermiques.

Premièrement, les images visibles donnent des informations sur la réflectance de la peau. Il a d'autres avantages tels que l'acceptation de l'utilisateur, la faible intrusion, l'utilisation d'une technologie peu coûteuse avec des capteurs à prix abordables. Cependant, des inconvénients existent :

- les problèmes de sensibilité lors de l'acquisition telles que les variations d'éclairage, la segmentation de l'arrière-plan ;

- la sensibilité aux expressions faciales, les mutations tel que l'âge, la pilosité, la barbe, la moustache, le piercing, le maquillage, le port de chapeaux, de port de lunettes, la chirurgie esthétique, etc.

- les variations d'éclairage ;
- les changements de poses.

L'imagerie infrarouge vient compenser les insuffisances des images visibles telles que les variations d'éclairage, le déguisement et la robustesse aux changements de poses. La caméra thermique capture l'émissivité thermique du corps humain qui est une caractéristique unique. Ceci est expliqué par la structure des vaisseaux sanguins qui transporte le sang chaud dans tout le corps, d'où le caractère unique des images thermiques infrarouges. À cet effet, les jumeaux identiques ont des propriétés thermiques différentes. En outre, la technologie thermique s'améliore au fil des ans avec une baisse de coût des caméras. Cependant, l'imagerie infrarouge présente des inconvénients, notamment :

- la modification de la température ambiante ;
- la modification des processus métaboliques des sujets dans des conditions de peur, de sport, de maladie, etc. ;
- la susceptibilité de la caméra en ce qui concerne les facteurs naturels tel que le vent, la poussière, la fumée, etc. ;
- la gestion du temps de réponse de la caméra pendant le fonctionnement sur une longue durée.

La reconnaissance du visage dans le visible et l'infrarouge thermique présente des avantages et des insuffisances. Parmi les méthodes de biométrie émergentes, l'imagerie multispectrale est une alternative à la biométrie classique. Plusieurs approches existent et notre étude se concentre sur la fusion d'images infrarouges thermiques et visibles pour la reconnaissance faciale multispectrale.

2.3.3. Fusion d'images

Il existe plusieurs définitions de la fusion d'images dans la littérature [99, 121, 123]. Dans [99], la fusion d'images se définit comme la combinaison de deux ou de plusieurs images différentes pour former à l'aide d'un algorithme une nouvelle image.

. La fusion d'images est une technique qui fusionne les informations en combinant plusieurs images d'une même zone, d'un même objet acquis dans différentes conditions. Ces conditions peuvent être de différentes dates d'acquisition, différentes conditions météorologiques, différents types de capteurs, etc.

Le but de la fusion d'images est d'intégrer une information complémentaire à partir de plusieurs images pour obtenir une image plus instructive pour les diverses utilisations dont la perception visuelle humaine, la reconnaissance, la médecine, la télédétection, la navigation, etc.

Le processus de fusion d'images consiste à intégrer des informations de différentes sources multi-capteurs, multi-vues, multi-focales, multi-dates dans une nouvelle image qui contient des caractéristiques de meilleure qualité. L'image fusionnée doit être plus informative. La fusion de différentes images permet d'améliorer la qualité visuelle et analytique de l'image.

Une image fusionnée est une nouvelle image qui présente des caractéristiques plus intéressantes pour les traitements dont la segmentation, l'extraction de caractéristiques et la reconnaissance d'objets.

2.3.4. Fusion d'images visible/infrarouge thermique

2.3.4.1. Méthodologie

La méthode MS-FRHF que nous proposons dans cette thèse est une approche de reconnaissance faciale utilisant des images fusionnées du visible et de l'infrarouge thermique. En ce qui concerne l'extraction de caractéristiques, deux caractéristiques sont utilisées en l'occurrence les points d'intérêt et la texture. L'ensemble du processus est articulé en trois étapes qui sont la fusion d'images, l'extraction de caractéristiques et la reconnaissance faciale.

2.3.4.2. Processus de fusion d'images visible/infrarouge

La fusion d'images est un processus permettant de combiner des informations pertinentes de plusieurs images en une image contenant une grande qualité d'information. Avant de fusionner des images, le recalage des images doit être pris en compte. Dans la base de données d'images utilisée, il y a une paire d'images faciales

acquises dans différentes poses et illuminations. Chaque paire est composée d'une image visible et d'une image infrarouge thermique.

Pour évaluer la quantité d'informations de chaque image, l'équation (19) définissant l'entropie est utilisée. L'entropie mesure la quantité d'information contenue dans une image. Il est sensible au bruit et à d'autres fluctuations rapides indésirables. Une image avec un contenu d'information élevé donne une entropie élevée.

$$H_e = - \sum_{i=0}^L h_{I_f}(i) \log_2 h_{I_f}(i) \quad (19)$$

où une image I_f est par exemple constituée d'une série d'événements, des pixels, ayant tous une probabilité h d'avoir une intensité i .

Deux paires d'images infrarouges visibles et thermiques sont présentées sur la figure (2-3). Le visage a la même position avec deux illuminations différentes. Ces images de visage provenant de la base de données IRIS sont acquises avec les illuminations Left Light /On (Lon) en ce qui concerne a et b et Left Light et Right Off en ce qui concerne c et d.



Figure (2- 3) : Images visible/ infrarouge (a et b) puis (c et d) acquises sous différentes illuminations

Dans un premier temps, le calcul d'entropie des images a montré que les images infrarouges thermiques contiennent plus d'informations que les images visibles correspondantes. Ensuite, en utilisant la formule (20), l'information mutuelle (MI) est calculée pour chaque paire d'images.

$$MI = \sum_{i=1}^M \sum_{j=1}^N h_{I_r I_f}(i, j) \log_2 \left(\frac{h_{I_r I_f}(i, j)}{h_{I_r}(i, j) h_j(i, j)} \right) \quad (20)$$

où $h_{I_x}(i, j)$ est la probabilité de l'image I_x

$$SSD = \sum_{(i,j) \in W} (I_1(i,j) - I_2(x+i, y+j))^2 \quad (21)$$

La corrélation des différentes paires est donnée par la mesure de similarité de la somme des différences quadratiques (SSD) dans la formule (21) et représentée à la figure (2-4).



Figure (2- 4) : De la gauche à la droite : image infrarouge thermique, image visible, carte de similarité de la somme des différences quadratiques SSD.

Plusieurs méthodes de fusion existent et certaines métriques de ces méthodes de fusion ont guidé le choix de l'une d'entre elles. Les méthodes étudiées sont MSVD, PCA et AVG [62, 121, 124].

La méthode de fusion basée sur l'analyse en composantes principales (ACP) est une méthode de sous-espace qui réduit les ensembles de données multidimensionnelles en dimensions inférieures pour l'analyse. Cette méthode détermine les poids pour chaque source d'image et utilise le vecteur propre correspondant à la plus grande valeur propre de la matrice de covariance de chaque image source.

La méthode de fusion basée sur la moyenne simple (AVG) consiste à produire une image fusionnée en prenant l'intensité moyenne des pixels correspondants des images considérées.

Enfin, la fusion d'images par une décomposition multiple des valeurs singulières (MSVD) est presque identique à celle des ondelettes. Le but de la fusion MSVD est de remplacer les filtres utilisés dans la transformée en ondelettes par des décompositions de la valeur singulière (SVD). Il a l'avantage d'être simple d'utilisation et s'applique convenablement dans des applications en temps réel comme la reconnaissance d'objet, la télédétection, etc.

Parmi toutes ces méthodes, seule la fusion PCA donne les meilleurs résultats.

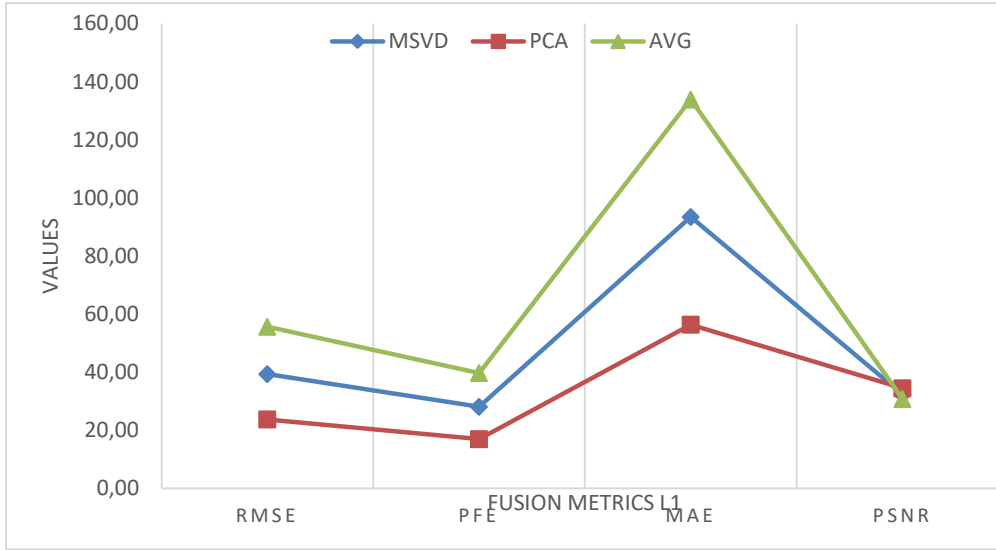


Figure (2- 5) : Comparaison de trois méthodes de fusion utilisant quatre métriques avec l'image infrarouge thermique choisie comme référence.

Les expérimentations faites dans le cadre de la comparaison des méthodes de fusion ont également montré qu'il est préférable d'utiliser l'image thermique comme image de référence et cela donne de meilleurs résultats.

Les différentes métriques utilisées sont : l'erreur quadratique moyenne (RMSE), l'erreur d'ajustement en pourcentage (PFE), l'erreur absolue moyenne (MAE), rapport signal sur le bruit (PSNR). Ils sont respectivement donnés par les formules (22), (23), (24) et (25) et décrits dans [121, 123].

$$RMSE = \sqrt{\left(\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_r(i,j) - I_f(i,j))^2\right)} \quad (22)$$

Où I_r et I_f sont deux images de taille MN

$$PFE = \frac{norm(I_r - I_f)}{norm(I_r)} * 100 \quad (23)$$

où la norme est l'opérateur pour calculer la plus grande valeur singulière.

$$MAE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |I_r(i,j) - I_f(i,j)| \quad (24)$$

$$PSNR = 20 \log_{10} \left(\frac{L^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_r(i,j) - I_f(i,j))^2} \right) \quad (25)$$

où L est la valeur maximale pour un pixel. Concernant une image, où les composantes d'un pixel sont codées sur 8 bits, la valeur de L est 255.

2.3.5. Description de la base de données multispectrale

La base de données de référence utilisée dans notre étude provient de la base de données Imaging, Robotics, and Intelligent Systems (IRIS) [42]. Il s'agit de la base de données 02 contenant des images visibles et infrarouges thermiques de visages. Les images de visages ont été acquises dans le visible et l'infrarouge thermique avec des variations d'illuminations, d'expressions et de poses.

Les capteurs utilisés sont Raytheon Palm-IR-Pro pour l'infrarouge et Panasonic WV-CP234 pour le visible. La base de données contient 4228 paires d'images de 30 individus avec les critères suivants :

- expression : ex1, ex2, ex3 - surprised, laughing, angry (varying poses);
- illumination : Lon (left light on), Ron (right light on), 2on (both lights on), dark (dark room), off (left and right lights off), varying poses.

La taille de la base de données est d'un volume de 1.83 GB. La figure (2-6) présente un extrait d'images de visages de deux individus acquise dans le visible et l'infrarouge, avec 4 variations de poses et deux conditions d'éclairage Lon et Ron.



Figure (2- 6) : Extraits d'images de la base de données IRIS : Image de deux individus acquises dans des conditions d'éclairage différentes.

2.3.6. Extraction de caractéristique par la méthode MS-FRHF

Pour l'extraction des points d'intérêt, MSER (Maximally Stable Extremal Region) a été comparé à SURF et BRISK pour extraire les points clés. C'est MSER qui donne les meilleurs résultats est robuste [104]. Les principales étapes de la méthode MSER sont présentes dans la figure (2-7).

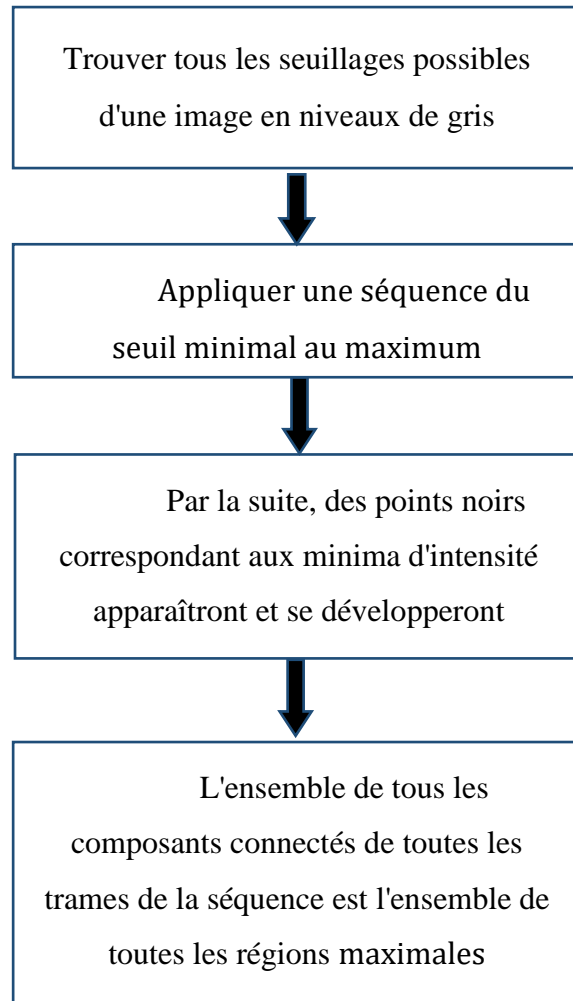


Figure (2- 7) : Principales étapes du détecteur MSER [104]

Cette méthode couplée à l'extraction de texture basée sur la matrice de cooccurrence de niveau de gris (GLCM) offre plus de possibilités grâce aux quatorze caractéristiques de texture qui peuvent être extraites. Dans cette expérience, quatre caractéristiques importantes [80]: contraste, corrélation, énergie et homogénéité ont été prises en compte. Les formules suivantes sont utilisées pour présenter différentes caractéristiques de texture :

$$\text{Contraste} = \sum_{i,j} |i - j|^2 p(i, j) \quad (26)$$

$$\text{Corrélation} = \sum_{i,j} \frac{ijp(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \quad (27)$$

$$\text{Energie} = \sum_{i,j} p(i, j)^2 \quad (28)$$

$$\text{Homogénéité} = \sum_{j,j} \frac{1}{1-(i-j)^2} p(i, j) \quad (29)$$

Les différentes formules (26), (27), (28) et (29) permettent de caractériser la texture en donnant respectivement des informations sur la quantité de variations locales, la mesure de la dépendance des niveaux de gris, l'uniformité de la texture et les distributions d'éléments.

Enfin, ces deux outils d'extraction de caractéristiques donnent une analyse de caractéristique hybride en utilisant des points d'intérêt et des textures.

La reconnaissance faciale est basée sur le réseau de neurone [39] pour la détection des visages. Et après détection, un test est effectué avec la métrique de similitude de fusion.

$$FSIM = \sum_{w \in W} sim(I_1, I_2, I_f | w) \left(\begin{array}{l} QI(I_1, I_f | w) \\ -QI(I_2, I_f | w) \end{array} \right) + QI(I_2, I_f | w) \quad (30)$$

$$\text{où : } sim(I_1, I_2, I_f | w) = \begin{cases} 0 & \text{if } \frac{\sigma_{I_1 I_f}}{\sigma_{I_1 I_f} + \sigma_{I_2 I_f}} < 0 \\ \frac{\sigma_{I_1 I_f}}{\sigma_{I_1 I_f} + \sigma_{I_2 I_f}} \text{ if } 0 \leq \frac{\sigma_{I_1 I_f}}{\sigma_{I_1 I_f} + \sigma_{I_2 I_f}} \leq 1 \\ 1 & \text{if } \frac{\sigma_{I_1 I_f}}{\sigma_{I_1 I_f} + \sigma_{I_2 I_f}} > 1 \end{cases}$$

I_1, I_2, I_f sont respectivement deux images et leurs images fusionnées

QI désigne la qualité d'image (Quality Image)

σ_{I_x} écart-type d'une image I_x

2.3.7. Résultats expérimentaux

Dans nos travaux nous avons utilisé des images faciales dans le visible et l'infrarouge de 14 visages regardant dans quatre directions avec des illuminations différentes: Lon (Gauche Allumé), Off (Gauche et Droite Off) et Ron (Droite Allumé). En tout, il y a 336 images faciales. Ensuite, la fusion donne 168 images fusionnées.

Pour l'extraction de caractéristiques, L1, L2 et L3 sont des images thermiques infrarouges, alors V1, V2 et V3 sont les images visibles correspondantes. Les images fusionnées obtenues sont PCA-L1V1, PCA-L2V2 et PCA-L3V3. Le tableau (2- 4) et la figure (2-8) récapitulent certains résultats.

Tableau (2- 4) : Présentation des caractéristiques de trois paires d'images et des images fusionnées. De gauche à droite, Contraste, Corrélation, Energie, Homogénéité, Point-clé.

Types d'images	Contraste	Corrélation	Energie	Homogénéité	Points-Clés
L1	0.62	0.94	0.12	0.84	117
V1	0.27	0.94	0.29	0.93	44
L2	0.40	0.92	0.16	0.86	147
V2	0.34	0.95	0.30	0.91	57
L3	0.50	0.92	0.10	0.92	107
V3	0.24	0.97	0.31	0.94	52
PCA-L1V1	0.43	0.94	0.15	0.84	84
PCA-L2V2	0.30	0.93	0.16	0.88	55
PCA-L3V3	0.23	0.96	0.19	0.91	51

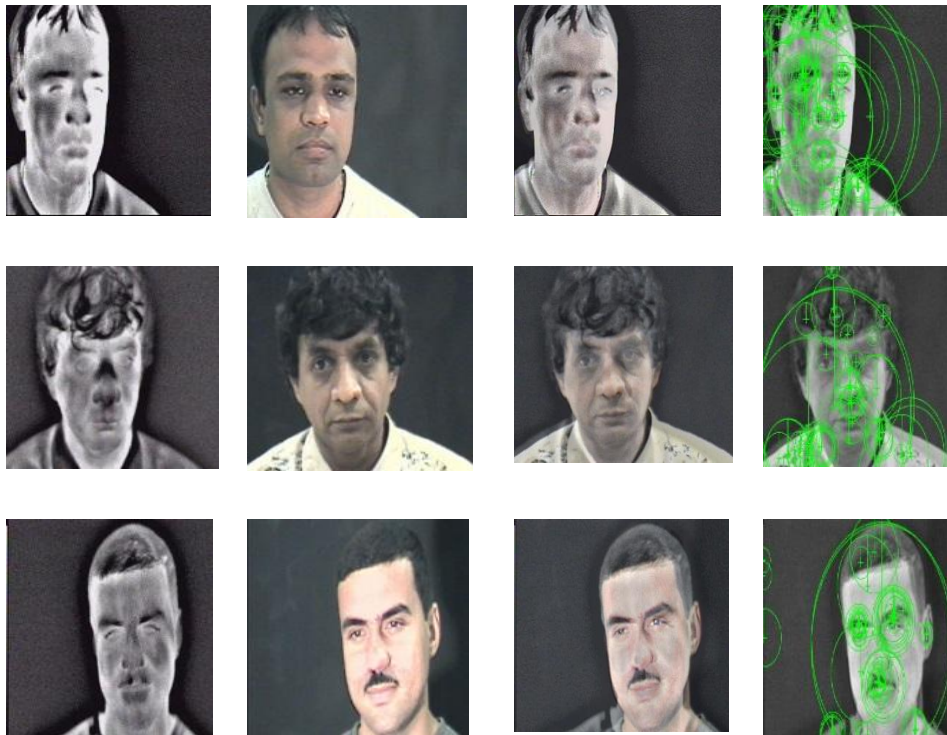


Figure (2- 8) : De la gauche vers la droite, Image visible, thermique infrarouge, image fusionnée correspondante

En observant le tableau (2- 4), les images fusionnées sont les meilleurs candidats pour la reconnaissance faciale. Par conséquent, MS-FRHF est une nouvelle façon de reconnaître le visage pour améliorer les méthodes traditionnelles. Les différents tests et résultats confirment cette hypothèse.

Conclusion

La performance de la reconnaissance faciale par l'imagerie multispectrale par rapport à la reconnaissance faciale classique est un problème majeur. Cette étude utilise des images multispectrales en fusionnant des images visibles et des images infrarouges thermiques. Les images thermiques infrarouges ont l'avantage d'avoir une signature unique d'émission thermique de la peau et sont robustes pour la reconnaissance faciale dans différents éclairages et poses diverses. Ces images thermiques sont combinées avec des images visibles qui utilisent les caractéristiques de réflectance de la peau. Les images composites obtenues par fusion donnent de bonnes performances en reconnaissance faciale.

Chapitre 3. Reconnaissance de visage multispectrale à l'aide de caractéristiques binaires

Introduction

Dans la dernière décennie est apparue un nouveau type de descripteurs dits binaires. Ces descripteurs se présentent comme une alternative aux descripteurs existants. Les descripteurs de caractéristiques binaires sont jugés robustes et ont une capacité de traitement élevé.

En plus de cette rapidité de traitement les détecteurs et descripteurs binaires se montrent performants, robustes et invariants à certains paramètres tel que le changement d'éclairage, la rotation, le changement d'échelle, etc. Dans ce chapitre, une étude est faite sur l'utilisation des détecteurs binaires dans la reconnaissance de visage multispectrale.

3.1. Motivation

La vision par ordinateur utilise des domaines qui manipulent de gros volume d'images. Et, de plus en plus il se pose des problèmes de stockage qui engendrent des problèmes de rapidité de traitement notamment en cartographie, biométrie, médecine, etc. Dans cette optique plusieurs études s'intéressent à ces méthodes pour améliorer les nouveaux applicatifs en termes de gain de temps, de ressources matérielles et de robustesse face à la complexité des données à traiter.

Dans les bases de données images volumineuses, la recherche d'image se fait généralement par l'utilisation de descripteurs de grande dimension qui sont gourmands en ressource temps. Aussi, le développement des applicatifs sur des outils intelligents tel que le mobile, les tablettes, les caméras, etc. nécessite de temps optimal pour les différents traitements. Dans le souci d'optimisation de temps, il est nécessaire de développer des outils adaptés qui s'exécutent rapidement.

Cette nouvelle approche de traitement a conduit à plusieurs travaux récents qui présentent des détecteurs et des descripteurs de caractéristiques binaires plus performants et adaptées aux nouvelles technologies [40].

Dans ce chapitre, une présentation des détecteurs et descripteurs est faite. Il s'agit dans cette étude de décrire ces outils et d'étudier les différentes combinaisons de détecteurs et de descripteurs binaires afin de déterminer leur performance. Comme application, une reconnaissance faciale multispectrale basée sur les descripteurs et descripteurs binaires est présentée.

3.2. Présentation de détecteurs et descripteurs binaires

Dans un algorithme de reconnaissance de visage et d'objet, trois composantes importantes sont à souligner : la détection, la description et la correspondance. De nombreux outils d'évaluations de ces composantes existent. Dans ce travail, il s'agira d'étudier les détecteurs et descripteurs binaires c'est-à-dire trouver un bon compromis entre détecteurs et descripteurs donnant une bonne performance.

3.2.1. Méthodes basées sur les caractéristiques locales

Les méthodes d'extraction de caractéristiques locales utilisent des caractéristiques distinctes (points-clés, bords, coins, etc.) dans les images, qui sont supposés être riches en information. Pendant l'étape d'extraction, un ensemble de caractéristiques distinctes est d'abord détecté en analysant des images et en recherchant des motifs de pixels distinctifs. Ensuite, une étape de description est effectuée, où certaines mesures permettent de considérer les voisinages de chaque point pour une description locale de la région considérée. En général, les entités sont formées sous la forme de vecteurs flottants multidimensionnels ou de chaînes de bits. Une caractéristique locale bien choisie doit généralement être invariante à une ou plusieurs transformations affines telles que l'échelle de l'image et les rotations. Enfin, une phase de mise en correspondance peut être utilisée pour observer la ressemblance de deux images.

3.2.2. Fonctionnement des détecteurs et descripteurs binaires

En vision par ordinateur de nombreuses applications sont basées sur l'extraction de caractéristiques. Dans le contexte des Technologies de l'Information et de la Communication (TIC) avec le développement d'applications émergentes dans les domaines du mobile, de la biométrie, des Systèmes d'Information Géographiques (SIG),

de nouvelles approches d'extraction de caractéristiques sont apparues. Au titre de ces méthodes, l'on peut citer l'extraction de caractéristiques par des outils binaires. L'une des caractéristiques les plus utilisée est le point d'intérêt. Cette approche consiste à détecter des points significatifs, appelés points-clés, point d'intérêt ou points caractéristiques et à procéder à une description du voisinage de chaque point d'intérêt. Une méthode robuste doit être invariante à certaines transformations. L'existence de détecteurs de points clés précis et fiables ainsi que de descripteurs adéquats est donc une étape importante dans la reconnaissance d'objets. Cette étude se concentre sur l'évaluation des détecteurs et descripteurs binaires pour la reconnaissance de visage multispectrale.

L'un des descripteurs le plus répandu est SIFT (Scale Invariant Feature Transform) [82] qui détecte les points-clés à l'aide d'une différence de Gaussienne (DoG). Bien que SIFT existe depuis plusieurs années, il a donné des résultats performants. Face aux succès de SIFT, plusieurs variantes de ce descripteur ont été élaborées. Ce sont ASIFT, CSIFT, GSIFT, PCA-SIFT et SURF qui est la variante la plus populaire. La méthode SURF a des performances similaires à SIFT, mais se présente plus rapide pour les calculs. Cependant, le temps de traitement de SIFT et ses variantes est encore trop élevé pour les applications en temps réel sur les nouvelles applications tels que les appareils mobiles ayant une puissance de calcul et une capacité de mémoire limitées.

Dans une littérature récente, plusieurs évaluations des descripteurs de caractéristiques binaires ont été présentées [40] : une évaluation sur les descripteurs de type SIFT, une évaluation sur SIFT et les descripteurs binaires, une évaluation sur les descripteurs binaires uniquement [40, 82].

Dans leur utilisation, le détecteur et le descripteur doivent être invariants par rapport aux mêmes transformations et les gains de vitesse obtenus par les descripteurs binaires s'obtiennent par le fait qu'ils sont binaires et courts.

Il est à remarquer que tous les descripteurs binaires récents possèdent les mêmes propriétés suivantes :

- le descripteur est construit à partir d'un ensemble de comparaisons d'intensité par paire ;

- chaque bit dans le descripteur est le résultat d'exactly une comparaison ;
- le modèle d'échantillonnage ;
- une compensation d'orientation.

En général, les descripteurs binaires respectent ces propriétés de base, et chacun ajoute ses propres propriétés uniques et spécifiques qui lui sont propres. Le tableau (3-1) présente un ensemble de détecteurs et descripteurs binaires.

Tableau(3- 1) : Présentations de principaux détecteurs et descripteurs binaires.

Détecteurs	Dimensions des détecteurs	Types	Références
BRISK (Binary Robust Invariant Scalable Keypoints)	512	Bit	[109]
BRIEF (Binary Robust Independent Elementary Features)	512	Bit	[70]
ORB(Oriented FAST and Rotated BRIEF)	256	Bit	[30]
FREAK (Fast Retina Keypoints)	512	Bit	[2]

3.2.3. BRISK

BRISK (Binary Robust Invariant Scalable Keypoints) a été introduit par Leutenegger et al. [109]. Il est invariant à l'échelle et à la rotation. Dans sa phase de détection, il utilise le détecteur de coin AGAST [75] qui est une variante améliorée de FAST [107]. Inspiré de FAST, AGAST est plus rapide et permet de calculer les emplacements des points-clés dans une pyramide d'échelle tout en faisant une suppression et une interpolation non-maxima à toutes les échelles. Il est invariant à l'échelle et à la rotation.

Ensuite, dans la phase de description, un modèle symétrique est utilisé pour décrire les caractéristiques. Les points d'échantillonnage sont localisés dans des cercles concentriques entourant l'entité, chaque point d'échantillon représentant un flou gaussien des pixels du voisinage. L'écart type de ce flou est augmenté avec la distance du centre de l'entité. Cela permet un appariement dense et capture plus d'informations utiles pour la description du point-clé.

Pour assigner une orientation, plusieurs comparaisons de points d'échantillons à longue distance sont utilisées. Lors des comparaisons des points d'échantillonnages, le déplacement du vecteur entre les points d'échantillonnage est stocké et pondéré par la

différence d'intensité relative. Ensuite, ces vecteurs obtenus sont utilisés pour calculer une moyenne déterminant la direction du gradient dominant du patch. Le motif d'échantillonnage est ensuite mis à l'échelle suivant une rotation. Le descripteur est constitué de 512 comparaisons de points d'échantillonnage à courte distance représentant les gradients locaux et la forme dans le patch.

Comparativement à BRIEF et ORB, BRISK est plus coûteux en temps de calculs et d'espaces. Les principales étapes de la méthode BRISK sont présentées par la figure (3-1).

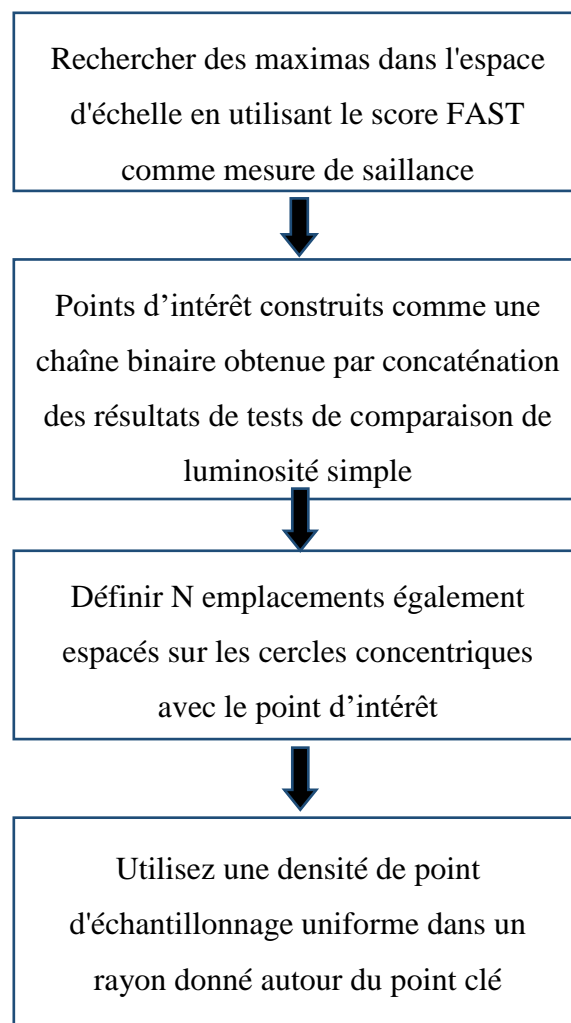


Figure (3- 1) : Principales étapes du détecteur et du descripteur BRISK [109].

3.2.4. BRIEF

Tout comme LBP et BRISK, BRIEF effectue un simple test de comparaison binaire et utilise la distance de Hamming en remplacement de d'autres distances telles que les distances euclidiennes ou de Mahalanobis. De façon rapide, la construction du descripteur binaire permet de comparer l'intensité entre deux positions de pixels situées au voisinage des points d'intérêt détectés. Cette méthode permet d'obtenir une description robuste et représentative à très faible coût de calcul.

L'algorithme BRIEF effectue un petit nombre de tests de différence d'intensité pour représenter un patch d'image en tant que chaîne binaire. Plus spécifiquement, un descripteur binaire pour un patch de pixels de taille $S \times S$ est construit en concaténant les résultats du test suivant

De façon théorique, un test T sur un patch P de taille $S \times S$ se définit comme suit :

$$T = \begin{cases} 1, & \text{si } I(P_j) > I(P_i) \\ 0, & \text{sinon} \end{cases} \quad (31)$$

où $I(P_i)$ indique la valeur d'intensité de pixel (lissée) en P et la sélection de l'emplacement de tous les P_i définit uniquement un ensemble de tests binaires. Pour augmenter la robustesse du descripteur, le patch de pixels est pré-lissé à l'aide d'un noyau gaussien avec une variance égale à 2 et une taille égale à 9×9 pixels. Le descripteur BRIEF possède deux paramètres de réglage qui sont le nombre de paires de pixels binaires et le seuil binaire.

BRIEF est très efficace à la fois pour calculer et stocker en mémoire. Cependant, le descripteur BRIEF n'est pas robuste contre une rotation supérieure à 35° environ, par conséquent, il n'est pas invariant pour la rotation.

En résumé, BRIEF utilise un modèle d'échantillonnage constitué de 128, 256 ou 512 comparaisons soient respectivement à 128, 256 ou 512 bits. BRIEF utilise généralement le détecteur SURF pour un gain de temps de calcul, mais il peut être couplé avec tout autre détecteur.

3.2.5. ORB

ORB (Oriented FAST et Rotated BRIEF) a été introduit par Rublee et al. [30]. C'est une variante de BRIEF pour combler le manque d'invariance de rotation de celui-ci. La méthode ORB calcule une orientation locale à l'aide d'un centroïde d'intensité qui est défini comme une moyenne pondérée des intensités de pixels dans le patch local supposé ne pas coïncider avec le centre de l'entité. L'orientation est caractérisée par le vecteur entre l'emplacement des caractéristiques et le centroïde.

Le modèle d'échantillonnage utilisé dans ORB est basé sur 256 comparaisons d'intensité par paires. Ces modèles d'échantillonnage sont construits par apprentissage automatique, maximisant la variance du descripteur tout en minimisant la corrélation sous divers changements d'orientation.

Généralement, le descripteur, ORB utilise le détecteur de coin FAST qui est limité quant à la détection de points et l'invariance à l'échelle. Pour pallier ces lacunes, le détecteur Harris est utilisé à chaque emplacement de point d'intérêt pour fournir une suppression non-maxima dans l'image et une pyramide d'échelle d'image limitée est utilisée pour détecter les points d'intérêt de tailles variables. Cependant, aucune suppression non-maxima n'est utilisée entre les échelles ; ce qui entraîne des détections potentielles en double dans les différents niveaux de la pyramide. Les principales étapes de la méthode ORB sont présentées par la figure (3- 2).

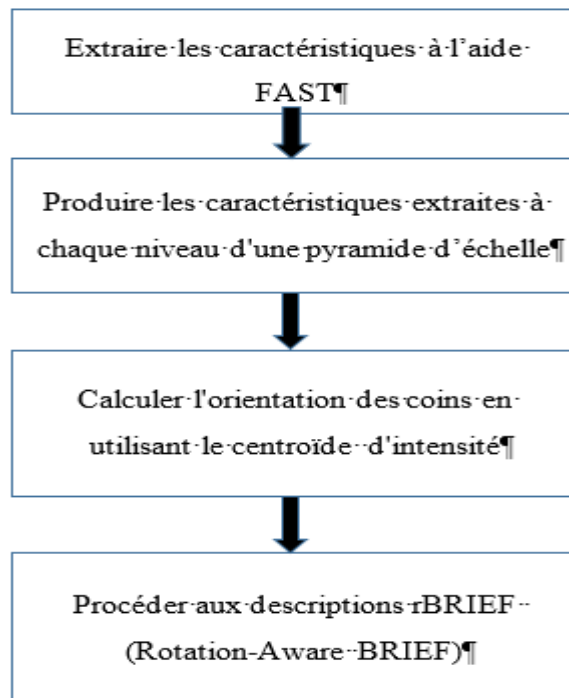


Figure (3- 2) : Principales étapes du détecteur et du descripteur ORB [30].

3.2.6. FREAK

FREAK (Fast Retina Keypoint) a été introduit par A. Alahi et al. [2]. Son principe de fonctionnement est basé sur le mécanisme biologique de l'œil humain. Les descripteurs FREAK utilisent une cascade de chaînes binaires, qui sont calculées en comparant les intensités d'image sur le modèle d'échantillonnage rétinien. Aussi, dans son fonctionnement, il utilise des cercles concentriques pour la méthode d'échantillonnage avec une différence de position des cercles. Les pixels en moyenne sont beaucoup plus concentrés au point d'intérêt, c'est-à-dire au centre. Cet échantillonnage s'inspire de la distribution des cellules rétiniennes avec leurs champs réceptifs correspondants. Dans la méthode, les champs réceptifs sont représentés par des cercles d'image gaussienne-convolutée permettant d'évaluer 43 gaussiennes pondérées autour de chaque point-clé. Enfin, la somme des gradients locaux est faite sur des paires sélectionnées pour l'assignation de l'orientation. La méthode FREAK est une méthode d'extraction de points d'intérêt rapide et robuste.

En résumé, les attributs des différents descripteurs et détecteurs binaires décrits précédemment sont récapitulés dans le tableau (3- 2).

Tableau(3- 2) : Récapitulatif des attributs des détecteurs et descripteurs binaires présentés plus haut.

Caractéristiques	BRISK	ORB	BRIEF	FREAK
Spectres	Binaire local + vecteur d'orientation	Binaire local + vecteur d'orientation	Binaire local	Binaire local avec passage de grossier à fin + vecteur d'orientation
Caractéristiques de forme	Carré	Carré	Carré centré au point d'intérêt	Carré
Caractéristiques de modèle	Comparaison des paires de points de pixels locaux formés	Comparaison des paires de points de pixels locaux formés	Comparaison de la paire de points pixel locale aléatoire	Comparaison de la paire de points pixel de la région 31x31
Densité des caractéristiques	Local 31x31 at points d'intérêt FAST	Local 31x31 at points d'intérêt	Local 31x31 at points d'intérêt	Local clairsemé AGAST points d'intérêt
Méthode de recherche	Fenêtre coulissante	Fenêtre coulissante	Fenêtre coulissante	Fenêtre coulissante sur l'espace d'échelle
Fonction de distance	distance de Hamming	distance de Hamming	Distance de Hamming	distance de Hamming
Robustesse	luminosité, contraste, rotation, échelle	luminosité, contraste, rotation, limité à l'échelle	Luminosité, contraste	luminosité, contraste, rotation, échelle, point de vue, flou

3.3. Méthode proposée pour la reconnaissance de visage multispectrale à l'aide de détecteurs et de descripteurs binaires

L'extraction de caractéristiques est largement utilisée dans de nombreuses applications de vision par ordinateur telles que la détection d'objets, la récupération d'images, la recherche d'images, la médecine, la reconnaissance d'objets et de scènes, etc. Le défi de ces recherches avancées sur ces domaines d'intérêt s'intéresse à des choix optimaux de nouvelles méthodes d'extraction de caractéristiques [1]. L'extraction de caractéristiques peut utiliser des détecteurs de caractéristiques locales et des descripteurs nécessitant un bon compromis entre les détecteurs et les descripteurs [9]. Cette étude présente les outils BRISK, FAST, FREAK, HARRIS et SURF pour la reconnaissance faciale multispectrale présentée comme alternative à la reconnaissance faciale

conventionnelle. Les images multispectrales sont obtenues par fusion d'images visibles et infrarouges thermiques.

L'extraction de caractéristiques présente un défi majeur en vision par ordinateur. Il est utilisé dans plusieurs applications tels que la reconnaissance de formes, le contrôle qualité, la reconnaissance d'objets dans plusieurs contextes, la récupération d'images, la biométrie, etc. Au cours de cette dernière décennie, face aux nouveaux défis de sécurité et aux nouvelles applications émergentes dans le multimédia, le mobile et le web, le recours à de nouvelles méthodes d'extraction de caractéristiques s'est avéré opportun. Spécifiquement, de nouveaux types de détecteurs et de descripteurs adaptés à cette fin, ont été largement utilisés. Il s'agit des descripteurs et détecteurs binaires. La présente étude montre un bon compromis entre les détecteurs et les descripteurs qui sont SURF, BRISK, FAST, FREAK et HARRIS vu comme une alternative à SIFT [9]. Ensuite, une étude de faisabilité est réalisée sur l'utilisation des descripteurs binaires dans la reconnaissance faciale multispectrale utilisant des images visibles / infrarouges thermiques. Cette reconnaissance faciale multispectrale avec de meilleures performances est une alternative à la reconnaissance faciale traditionnelle utilisant des images couleur [5].

Ces travaux nous ont permis d'apporter deux contributions principales qui sont présentées ci-dessous:

- l'identification des détecteurs et des descripteurs pour une meilleure extraction de caractéristiques;
- l'application de bon compromis entre les détecteurs et les descripteurs pour l'extraction de points d'intérêt en reconnaissance faciale multispectrale.

3.3.1. Présentation des méthodes d'extraction de caractéristiques utilisées dans la méthode proposée

Les points d'intérêt peuvent caractériser les zones les plus visuelles dans une image. Un point d'intérêt est un pixel, un coin ou une zone avec une position bien définie dans une image et peut être efficacement détecté. Les points d'intérêt présentent un contenu d'informations locales élevé et peuvent être répétés entre les images.

Pour leur robustesse, les méthodes d'extraction de point d'intérêt cherchent à être stables pour certains paramètres : la rotation, la translation, l'illumination et l'échelle. Le processus utilisé pour la reconnaissance peut se récapituler en trois étapes: la détection des points-clés, la description des points-clés et la correspondance des images. Les deux premières étapes consistent en l'extraction de points d'intérêt et la dernière étape est la correspondance d'image qui permet d'identifier des régions communes des images.

Pour la détection de caractéristiques, il existe plusieurs types de détecteurs. Quant à la détection des coins, les détecteurs d'angle les plus utilisés sont Harris, Moravec et SUSAN (Smallest Univalued Segment Assimilating Nucleus) [17, 38, 73]. Concernant la détection des blobs, les détecteurs de blobs les plus populaires sont généralement LoG (Laplacian of Gaussian) et DoG (Difference of Gaussian).

Depuis, ces dernières années, la recherche se concentre sur la robustesse des détecteurs. Ce sont les détecteurs invariants aux transformations affines. Ces méthodes prennent en compte certains paramètres qui sont : l'échelle, l'illumination, le bruit, etc. Les plus utilisées de ces outils d'extraction de caractéristiques sont FAST (Features from Accelerated Segment Test) [107] et SIFT (Scale Invariant Feature Transform) et de ses variantes [128] tels que ASIFT (Affine-SIFT) et SURF (Speeded Up Robust Features).

Les études récentes depuis 2010 proposent les détecteurs binaires [9] : BRISK (Binary Robust Invariant Scalable Keypoints), BRIEF (Binary Robust Independent Elementary Features), ORB (Oriented FAST and Rotated BRIEF) et d'autres. Ces nouveaux types de descripteurs ont l'avantage d'améliorer les propriétés des anciennes méthodes et de les adapter aux nouvelles applications de vision par ordinateur utilisant le multimédia, le web, le mobile, etc.

Pour l'extraction de caractéristiques conçue à l'aide des descripteurs binaires, de nombreuses études ont été faites [2, 9]. Notre étude se focalise sur les méthodes BRISK et FREAK. BRISK qui ont l'avantage de faire la détection et la description des points d'intérêt. Le détecteur utilisé est AGAST amélioré qui lui-même est inspiré de FAST. FREAK n'est qu'un descripteur binaire et fonctionne comme la rétine de l'œil humain.

Pour la mise en correspondance d'images, il est important de localiser les points d'intérêt qui ont des propriétés majeures. Ces propriétés sont :

- le caractère distinctif, le point d'intérêt doit être distinct localement ;
- l'invariance à différentes transformations comme l'illumination, l'échelle et la rotation ;
- la stabilité des points-clés détectés qui doivent apparaître dans les deux images ;
- la rareté qui consiste à avoir des points-clés globalement séparables.

Enfin, pour réussir une bonne reconnaissance, il est important de choisir un bon compromis entre les détecteurs et les descripteurs. Dans cette étude, une étude de FAST, SURF, SURF, BRISK et FREAK a permis de sélectionner un détecteur et un descripteur adaptés à la reconnaissance multispectrale du visage.

3.3.1.1. Présentation de BRISK

BRISK a été proposé par Leutenegger et al. [109] et contient un détecteur et un descripteur. Le détecteur utilisé est un AGAST inspiré de FAST.

Pour la détection de points-clés, les différentes opérations effectuées sont :

- calcul de score à l'aide de FAST sur les différentes échelles ;
- suppression des niveaux de pixel non-maxima ;
- calcul du maximum de sous-pixels sur le patch ;
- calcul du maximum continu sur plusieurs échelles.

Pour la description du point-clé, les étapes proposées sont:

- utiliser un modèle de pixels lissés autour de l'entité ;
- séparer les paires de pixels en deux sous-ensembles ;

- utiliser les paires de courte distance et longues distances pour calculer le gradient local entre les longues distances ;
- calculer les gradients pour déterminer l'orientation des fonctions ;
- faire pivoter des paires de courte distance en utilisant l'orientation des caractéristiques et construire un descripteur binaire à partir de paires de courte distance pivotées.

Le modèle d'échantillonnage utilisé dans l'étape de détection est présenté à la figure (3- 3).

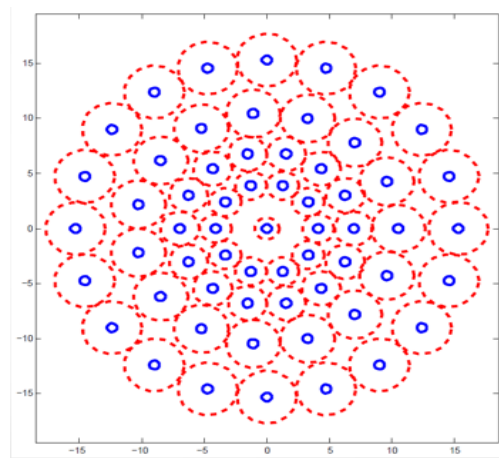


Figure (3- 3) : Modèle d'échantillonnage de BRISK [109].

3.3.1.2. Présentation de FAST

FAST est un détecteur proposé par Rosten et Drummond [107] pour détecter des points d'intérêt dans une image. Il est devenu un algorithme efficace pour la détection de coin. Un coin est une région avec un changement d'intensité dans deux directions différentes.

Pour tout point p donné dans l'image, un test est effectué pour vérifier s'il s'agit d'un coin en choisissant 16 pixels dans un rayon de 3 pixels autour de p et en trouvant la plus longue série de pixels dont l'intensité est comparée à l'intensité de p . Si, le résultat est d'au moins 12 pixels, p est considéré comme un coin. Pour la détection, le pixel p et ses voisins sont considérés comme le montre la figure (3- 4).

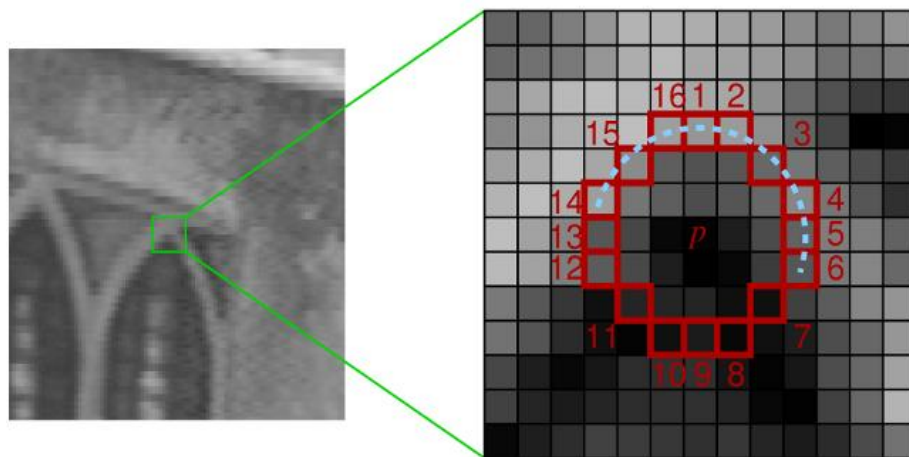


Figure (3- 4) : Image montrant le point d'intérêt testé et les 16 pixels du cercle [107].

L'algorithme FAST peut être rendu encore plus rapide en vérifiant d'abord les pixels 1, 5, 9 et 13. Si au moins trois d'entre eux remplissent la condition d'intensité, ils sont immédiatement exclus et le point donné est un coin. Pour éviter de détecter plusieurs coins proches du même pixel, il est exigé une distance minimale entre les coins. Ensuite, le score est calculé comme la somme des différences d'intensité entre p et les pixels de la série. FAST offre en général des performances nettement supérieures à celles de plusieurs détecteurs de caractéristiques. Ce détecteur est utilisé dans de nombreuses applications tels que le traitement vidéo, mobile, biométrique, etc.

Les principales étapes de la méthode FAST sont présentées dans la figure (3- 5).

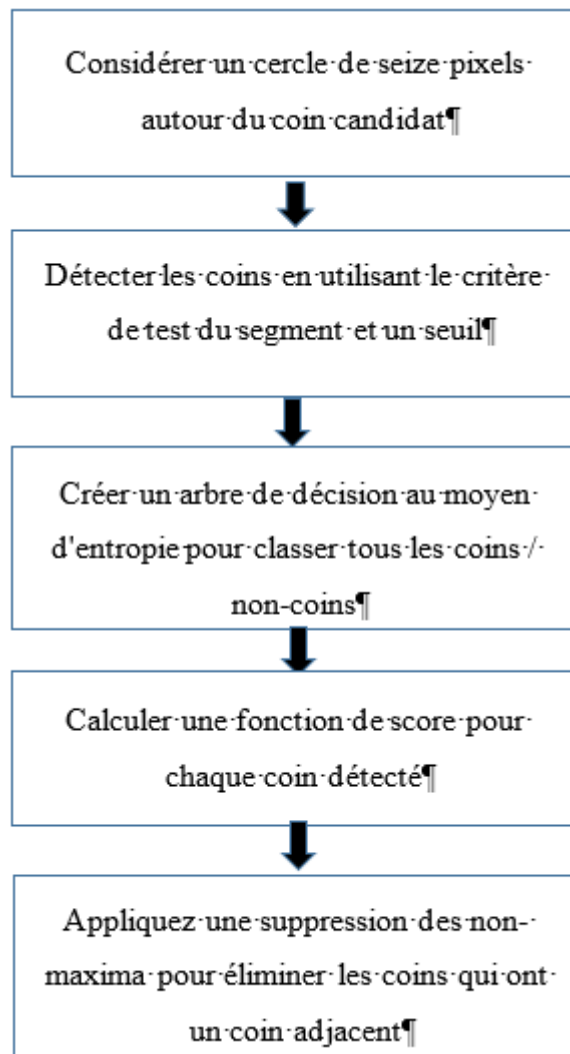


Figure (3- 5) : Principales étapes du détecteur de coin FAST [107].

3.3.1.3. Présentation de FREAK

FREAK est un descripteur binaire proposé par Alahi et al. [2]. Comme BRISK, ce descripteur utilise un modèle d'échantillonnage et une méthode de compensation d'orientation. Il s'agit d'une variante de BRISK améliorée en utilisant une sélection de paires de gabarits. FREAK organise des points d'échantillonnage analogues à la structure de la rétine biologique. Pour la description du point d'intérêt, les outils utilisés sont des gaussiennes pondérées, le motif fonctionnant comme la rétine et une assignation d'orientation est faite pour la description. Avec FREAK, les étapes de mise en correspondance sont basées sur la méthode grossière à fine.

Le modèle d'échantillonnage utilisé dans FREAK est présenté à la figure (3- 6).

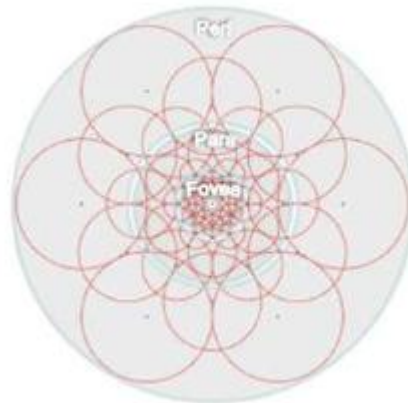


Figure (3- 6) : Modèle d'échantillonnage FREAK [2].

Les cercles rouges désignent les écarts-types des noyaux gaussiens utilisés pour les points d'échantillonnage correspondants. Un ensemble de 43 points d'échantillonnage est sélectionné pour le modèle d'échantillonnage du descripteur FREAK

Ensuite, la construction du descripteur est faite à l'aide de la comparaison binaire de l'intensité des paires des centres réceptifs préalablement lissés. Le récepteur a une grande capacité comparative de 512 paires (4 groupes de 128) et en fonction de l'orientation, un modèle symétrique est construit.

Enfin, l'orientation est normalisée par le choix du modèle basé sur 45 points d'échantillonnage des points d'échantillonnage de façon symétrique par rapport au centre comme le montre la figure (3- 7).

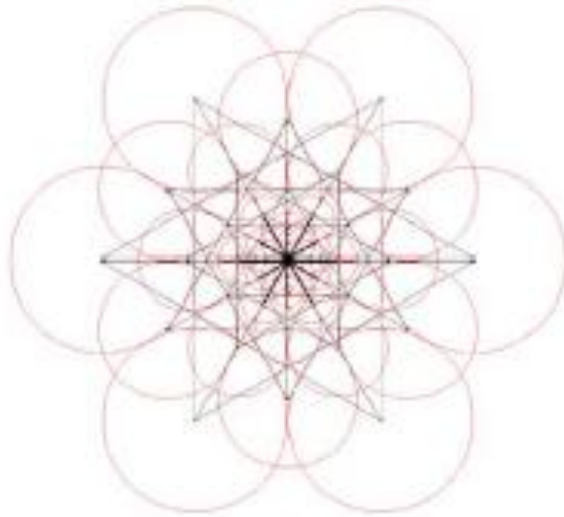


Figure (3- 7) : Paires sélectionnées pour calculer l'orientation du point d'intérêt [2].

En somme, FREAK est robuste à six paramètres qui sont la luminosité, le contraste, la rotation, l'échelle, le point de vue et le flou.

3.3.1.4. Présentation du détecteur Harris

Le détecteur de coin Harris est proposé par C. Harris et M. Stephens [55, 83] qui détecte facilement le point d'intérêt à travers une petite fenêtre et en déplaçant cette fenêtre dans n'importe quelle direction. L'algorithme utilisé par ce détecteur de coin est articulé en deux étapes principales qui consistent à :

- trouver des points avec la fonction de réponse R de grand coin ($R > \text{seuil}$);
- prendre les points des maxima locaux de R .

En tout, l'algorithme de détection de coin Harris est réalisé par le calcul du gradient de chaque pixel. Ensuite, si les valeurs absolues de gradient dans les deux directions sont toutes deux grandes, alors le pixel est supposé être un coin.

Le détecteur d'angle de Harris est devenu un détecteur de points d'intérêt populaire en raison de sa forte invariance : rotation, échelle, variation d'éclairage et bruit. Le détecteur de coin Harris utilise la fonction d'auto-corrélation locale d'un signal.

3.3.1.5 Présentation de SURF

SURF est une variante de SIFT [128]. Il est utilisé pour la détection et la description du point d'intérêt. Pour la détection de points-clés, tout d'abord, un espace

d'échelle est construit par un filtre de boîte à échelle différente pour convoluer avec l'image originale. L'utilisation d'une matrice de Hesse indique les points-clés candidats et effectue une suppression des non-maxima.

Pour la description de point-clé, d'abord une direction principale est obtenue en calculant une réponse d'ondelette de Haar dans deux directions x et y des secteurs dans une zone circulaire. La direction principale est la norme maximum. Ensuite, les réponses d'ondelettes de Haar et l'affectation d'orientation sont prises en compte pour construire le descripteur de point-clé, décrire le point d'intérêt et son voisinage. Le vecteur descripteur SURF mesure 64 points flottants.

3.3.2 Matériels et méthodes

3.3.2.1. Base de données d'images

La base de données est extraite de la base de données IRIS (Imaging Robotics Intelligent System) [42]. Cette base de données contient des images infrarouges thermiques et visibles. Cette étude utilise des images faciales dans le visible et l'infrarouge de 14 visages regardant dans quatre directions avec des illuminations différentes: Lon (Light On), Off (Left and Right Off) et Ron (Right Light On). Au total, il y a 336 images faciales.

3.3.2.2. Fusion d'images

Le processus de fusion d'images combine des informations pertinentes pour plusieurs images en une image avec une qualité d'information élevée. Une méthode de fusion d'image efficace prend en compte le recalage d'image. La base de données IRIS contient des paires d'images visibles et d'images infrarouges de visages acquises dans différentes directions et éclairages. Les images fusionnées sont composées d'une image visible et d'une image infrarouge thermique. Le processus de fusion a été présenté dans le chapitre précédent.

Les différentes étapes décrites dans le processus de fusion d'images permettent d'obtenir les images présentées dans la Figure (3- 8).



Figure (3- 8) : Ensemble d'images. De gauche à droite : image visible, image infrarouge thermique, carte de similarité entre les images visibles. / infrarouges par SSD, image fusionnée par la méthode de fusion PCA avec l'illumination Lon

Au cours des dernières années, plusieurs études ont présenté la reconnaissance multispectrale du visage comme une alternative à la reconnaissance faciale conventionnelle qui utilise généralement des images en couleurs. L'approche de l'imagerie couleur a montré des limites. Dans l'approche proposée, il est décrit une méthode de reconnaissance faciale multispectrale en utilisant des images fusionnées

visibles et infrarouges. Dans le visible et l'infrarouge, chaque longueur d'onde donne des informations utiles qui peuvent être interprétées. Les images visibles fournissent des informations sur la réflectance de la peau du visage et les images infrarouges thermiques utilisent l'émissivité du rayonnement thermique des faces donnant une caractéristique unique. La motivation en imagerie thermique est de surmonter les déficiences dues à des problèmes : bruit, fumée, neige, poussière, éclairage extrême, obscurité, pose dans différentes directions, déguisement, port de lunettes, port du chapeau et expression faciale. Les images composites sont obtenues par fusion d'images visibles et infrarouges afin de fournir des informations supplémentaires sur la réflectance et le rayonnement thermique du visage pour améliorer les méthodes traditionnelles de reconnaissance faciale [48].

3.3.2.3. Extraction de caractéristiques

Pour l'extraction de caractéristiques, les paires de détecteurs et de descripteurs sont présentées dans le tableau (3- 3).

Tableau(3- 3) : Paires de détecteurs et de descripteurs.

Détecteurs	Descripteurs
BRISK	BRISK
BRISK	FREAK
FAST	BRISK
FAST	FREAK
HARRIS	BRISK
HARRIS	FREAK
SURF	BRISK
SURF	FREAK

Pour la mise en œuvre de paires de détecteurs et de descripteurs, Matlab 2015 a été utilisé. Premièrement, ces techniques ont été appliquées sur des images visibles. Après, les mêmes méthodes ont été implémentées sur les images infrarouges. Enfin, chaque paire de détecteur et de descripteur a été testée sur des images multispectrales avec des résultats satisfaisants.

Pour les trois types d'images visibles, infrarouges, visibles / infrarouges, les mêmes résultats ont été observés. La figure (3-9) montre les résultats des points-clés

extraits sur V1 de l'image visible, L1 de l'image correspondante infrarouge, L1V1 de l'image fusionnée visible / infrarouge.

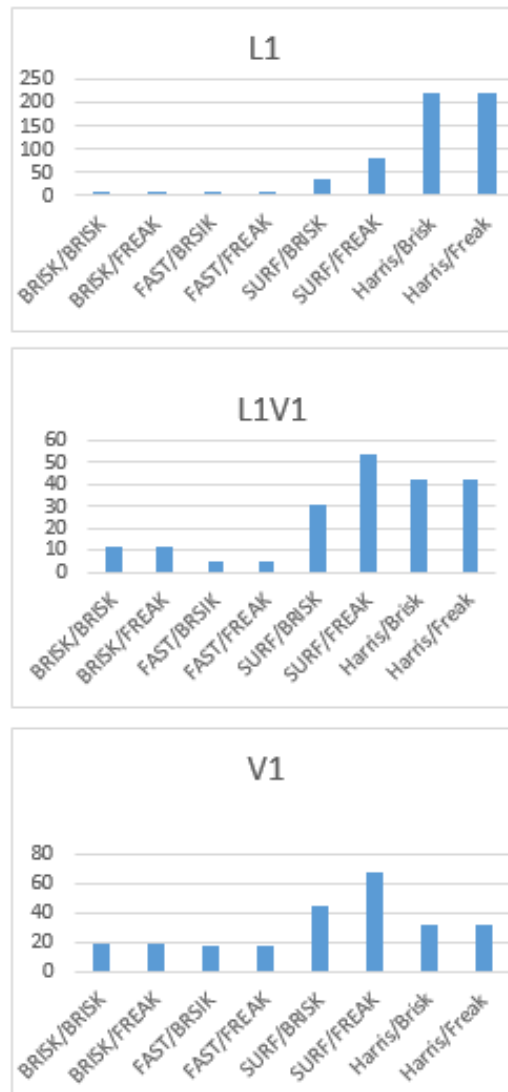


Figure (3- 9) : Graphique d'évaluation de détecteurs et descripteurs.

De haut en bas : application d'un couple de détecteur et de descripteur sur une image visible avec l'illumination Lon, implémentation d'un couple de détecteur et descripteur sur une image infrarouge avec illumination Lon, application d'un couple de détecteur et descripteur sur une image multispectrale avec illumination Lon.

L'analyse graphique montre que tous les couples de détecteur et de descripteurs peuvent être appliqués sur les trois types d'images qui sont des images visibles, infrarouges et visibles / infrarouges. La possibilité d'utiliser ces détecteurs et descripteurs sur des images multispectrales est prouvée. De bons résultats sont obtenus

avec les couples suivants SURF / BRISK, SURF / FREAK, Harris / BRISK et Harris / FREAK. Mais le meilleur d'entre eux, est le couple SURF / FREAK.

Dans la partie expérimentale, quelques résultats d'une application du détecteur SURF et du descripteur FREAK sur des images multispectrales seront présentés.

3.3.3. Résultats expérimentaux

Dans cette expérience, 168 paires d'images fusionnées ont été obtenues à partir de 336 images visibles et infrarouges thermiques. Ces images sont extraites de la base de données IRIS. Pour l'acquisition, il s'agit de 14 personnes regardant dans 4 directions avec 3 conditions d'éclairage qui sont Lon (Left Light On), Off (Left and Right Off) and Ron (Right Light On).

Dans cette expérience, après l'extraction des points d'intérêt, une étape de mise en correspondances des images est faite. Les résultats sont présentés dans la Figure (3-10).

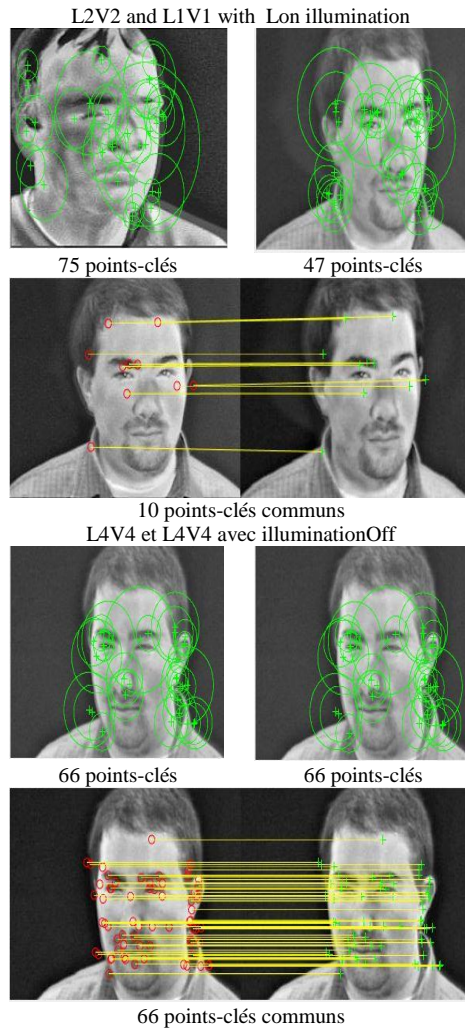


Figure (3- 10) : Extraction des points-clés sur des images d'éclairage unique

Une autre expérience a consisté à choisir différentes directions de pose, différentes personnes, différentes conditions d'illuminations. Les résultats de trois cas sont récapitulés à la figure (3- 11).

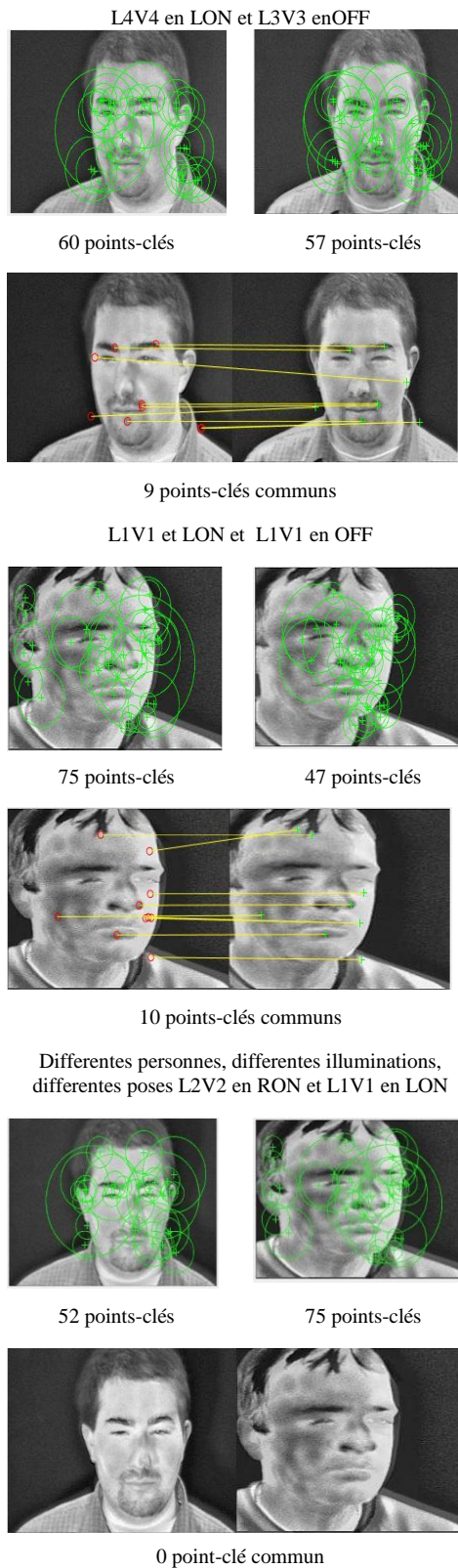


Figure (3- 11) : Extraction de points-clés sous différentes conditions

Enfin, les différentes expériences ont permis d'utiliser des descripteurs binaires. Ces expériences ont été mises en œuvre avec SURF qui est une variante améliorée de SIFT et FREAK un descripteur binaire récent et efficace.

Conclusion

. Les avantages de ces nouvelles méthodes d'extraction de caractéristiques sont nombreux devant un grand nombre de défis tels que les variations d'échelle, de rotation, d'éclairage, de point de vue, de qualité d'image et d'occlusion.

Notre étude présente une analyse d'un bon compromis entre les détecteurs et les descripteurs. Les descripteurs récents qui sont des descripteurs binaires surpassent plusieurs descripteurs cités dans la littérature. Une utilisation optimale de ces descripteurs nécessite un bon choix de détecteur. Une expérience différente dans cette étude montre la performance de SURF et FREAK respectivement descripteur de détecteur.

FREAK fonctionne comme un système oculaire humain. Ses performances et sa précision sont utilisées dans plusieurs applications tels que la reconnaissance de la rétine, l'application web, le mobile, etc. Cet outil peut également être implémenté sur des images multispectrales. Donc, la reconnaissance multispectrale du visage peut utiliser cet outil pour combler les faiblesses de la reconnaissance faciale traditionnelle.

Chapitre 4. Reconnaissance de visage multispectrale à l'aide de score de similarité

Introduction

Les progrès de la vision par ordinateur ont permis d'obtenir des résultats performants dans les domaines de l'identification, le contrôle et la sécurité. Vu, l'enjeu de la reconnaissance de visage, la recherche de précision suscite de nombreuses études dans les secteurs de la biométrie, l'authentification et l'identification des suspects. De façon générale, la reconnaissance biométrique selon les différentes modalités biométriques utilise un même processus de traitement. Lequel est composé essentiellement d'une extraction de caractéristiques, d'une modélisation de caractéristiques et d'une analyse comparative afin de déterminer un score de similarité.

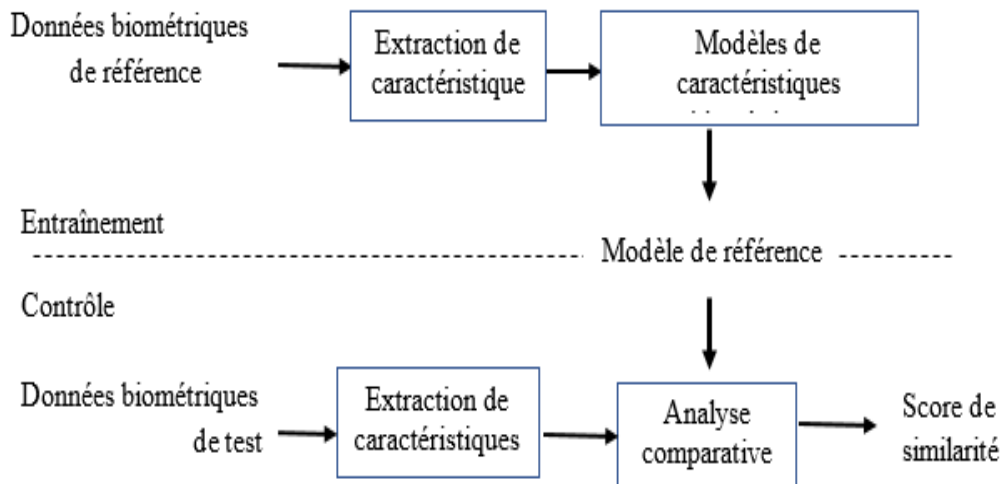


Figure (4- 1) : Chaîne classique de traitement pour la reconnaissance biométrique.

Dans le processus de traitement, l'analyse comparative consiste à déterminer un degré de similarité entre l'échantillon et le modèle de référence pour traduire la ressemblance des deux objets comparés.

Ce chapitre présente la notion de score de similarité et en proposons une détermination de score de similarité pour la reconnaissance multispectrale de visages.

4.1. Approches de comparaison d'objets

Il existe plusieurs approches pour comparer deux objets dont les méthodes statistiques et structurelles. Toutes ces deux approches utilisent dans le processus de traitement la notion de distance. En général, le choix d'une distance dépend de la spécificité du problème. Les propriétés des deux approches citées sont récapitulées dans le tableau suivant.

Tableau (4- 1) : Présentation des méthodes statistiques et structurelles

Caractéristiques	Méthodes statistiques	Méthodes structurelles
Fondement	Théorie mathématique des espaces vectoriels	Intuitivement attrayant: la perception humaine
Approche	Quantitative	Qualitative : structurelle/syntaxique
Descripteurs	Caractéristiques numériques : vecteurs de longueur fixe	Primitives morphologiques de taille variable
Syntaxe	Position de l'élément dans un vecteur	Processus d'encodage des primitives
Bruit	Facilement encode	Besoin de structures régulières
Apprentissage	Méthodes vectorielles	Graphes, arbres de décision, grammaire
Similarité	Métrique (exemple : distance Euclidienne)	Défini dans un processus de correspondance
Discrimination	Basée sur les distances ou les produits intérieurs dans un espace vectoriel	Grammaires pour reconnaître les objets valides (exemple : distance)
Chevauchement de classes	En raison de caractéristiques incorrectes et de modèles probabilistes	En raison de primitives incorrectes conduisant à l'ambiguïté

4.2. Motivation de l'utilisation des scores de similarité

La notion de ressemblance joue un rôle essentiel dans la classification. Et, les classes formées sont considérées comme des ensembles d'objets ayant les mêmes éléments d'observations et des propriétés communes. En fait, les objets similaires forment une même classe. Afin de déclarer objectivement la similarité de deux objet, il faut disposer d'une mesure. Il existe plusieurs mesures de similarité et chacune a sa spécificité. La similarité de deux objets se détermine toujours par rapport à des caractéristiques spécifiques choisis selon le contexte de l'étude. L'étude de la similarité

de deux objets utilise nécessairement une mesure basée sur deux types de représentations construites sur les caractéristiques ou sur la similarité. Ces représentations sont présentées dans la figure (4- 2).

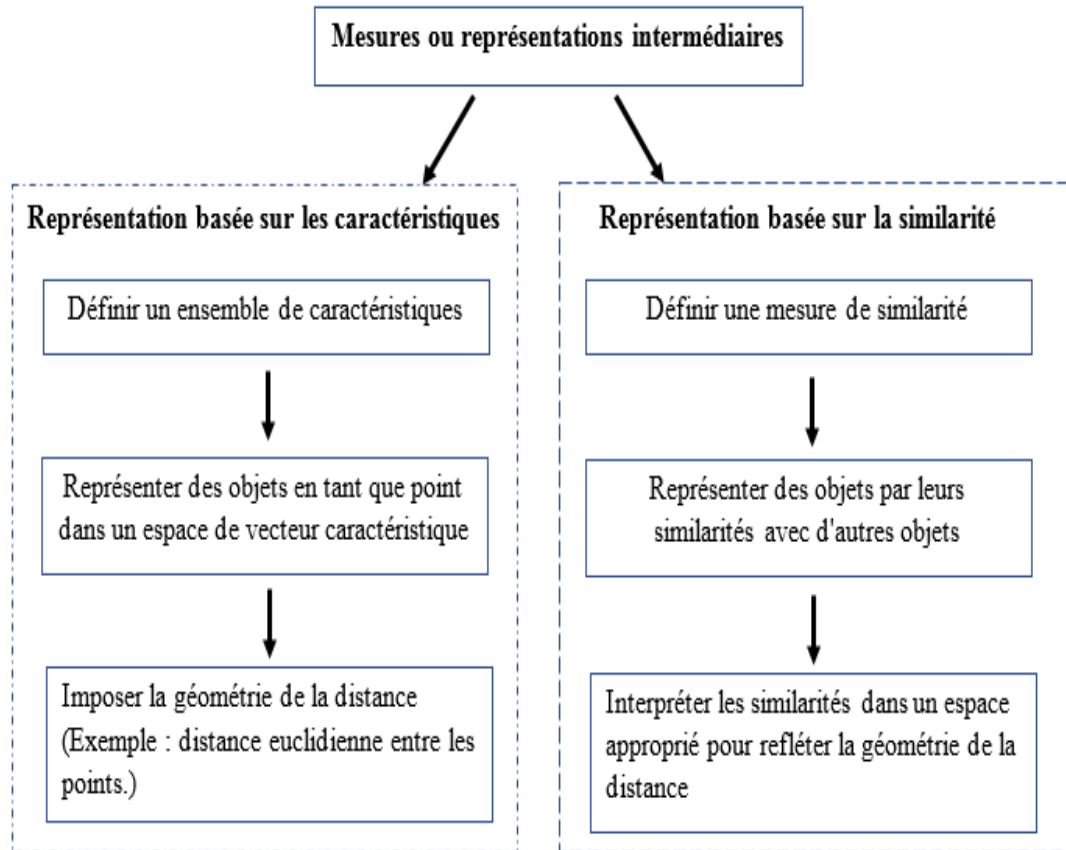


Figure (4- 2) : Différence entre les représentations basées sur les caractéristiques et la similarité.

La similarité peut être modélisée par une mesure de similarité ou de dissimilarité. Celles-ci sont intimement liées ; une petite dissemblance et une grande similitude désignent toutes deux une ressemblance étroite des objets. Il est possible de changer une valeur de similarité en une valeur de dissimilarité et vice versa, mais l'interprétation de la mesure pourrait être affectée.

En résumé, les représentations de similarité sont avantageuses pour l'identification et la reconnaissance, en particulier dans les cas suivants :

- les images numériques : fusion, reconnaissance, fouille, forme, contour, etc.
- les analyses de séquence d'ADN ;
- les phénomènes décrits par des fonctions de densité de probabilité ;
- les analyses de données ;

- les objets codés de manière structurelle par des arbres, des graphiques ou des chaînes ;
- les objets sont représentés comme des vecteurs ;
- les classificateurs.

Enfin, dans cette étude, les mesures de similarité seront utilisées pour la reconnaissance faciale multispectrale.

4.3. Mesures de similarité appliquées aux images

Dans le domaine de la vision par ordinateur, il existe plusieurs mesures de similarité dans les applications telles que la récupération d'images, la recherche d'images, la fusion d'images, la segmentation, la robotique, la médecine, la reconnaissance, etc.

4.3.1. Définitions

Le domaine de la reconnaissance rencontre de nombreuses difficultés comme la variation d'échelle, de contraste, d'énergie et de rotation. Pour un objet donné, ces problèmes peuvent être liés à des facteurs comme la multiplicité des capteurs, les conditions d'acquisition, la nature de la scène, etc. Par conséquent, la notion de l'invariance dans l'interprétation de l'image est étroitement liée à celle de la similarité. Une étude de ressemblance performante doit tenir compte de cette considération.

Les mesures de similarité classiques ne le permettent pas. Pour pallier cette insuffisance, une phase d'extraction de caractéristiques doit préalablement être faite à l'aide d'outils invariants aux transformations. Selon la représentation de l'image, des familles de mesures de similarité existent et sont :

- les mesures de similarité basées sur des matrices de pixels ;
- les mesures de similarité basées sur les caractéristiques ;
- les mesures de similarité structurelle.

Dans la littérature, une autre classification des mesures de similarité est faite et propose quatre groupes présentés dans le tableau ci-dessous.

Tableau (4- 2) : Groupes de mesures de similarité.

Mesures de Similarités	Descriptions	Références
Mesures de distance	mesures pouvant être utilisées avec tous les types de données (exemple : distance Euclidienne)	[21]
Coefficients d'association	mesures généralement utilisées avec des données binaires catégorielles	[116]
Coefficients de corrélation	Mesures de liaison entre deux objets	[116]
Coefficients de similarité probabiliste	mesures basées sur les probabilités et statistiques	[64]

4.3.2. Applications des mesures de similarité

La mesure de la similarité entre objets joue un rôle important dans de nombreux domaines de la vision par ordinateur : le traitement d'images, la recherche d'images, la compression d'images, la reconnaissance de formes, la classification, la biométrie, etc. Comme application, le score de similarité est utilisé pour étudier la ressemblance de deux objets. Avant de déterminer le score de similarité, il est indispensable de procéder à une extraction de caractéristique. La performance des systèmes de reconnaissance dépend d'une définition de mesure de similarité appropriée [112].

Dans la reconnaissance d'objets, la comparaison de deux images est une opération fondamentale et permet d'évaluer la ressemblance des images en question. Une image peut être représentée par un vecteur de caractéristiques où chaque élément est associé à un attribut donné de l'image. Ces attributs sont souvent des valeurs numériques uniques obtenues lors de l'extraction de caractéristiques. La ressemblance de deux images est représentée par le calcul de la mesure de similarité entre leurs vecteurs de caractéristiques. Du point de vue mathématique, la distance est définie comme une valeur quantitative de la distance entre deux points. Cette reconnaissance d'objets tient compte de l'extraction de caractéristiques et augmente par conséquent le taux de précision dans le processus de reconnaissance.

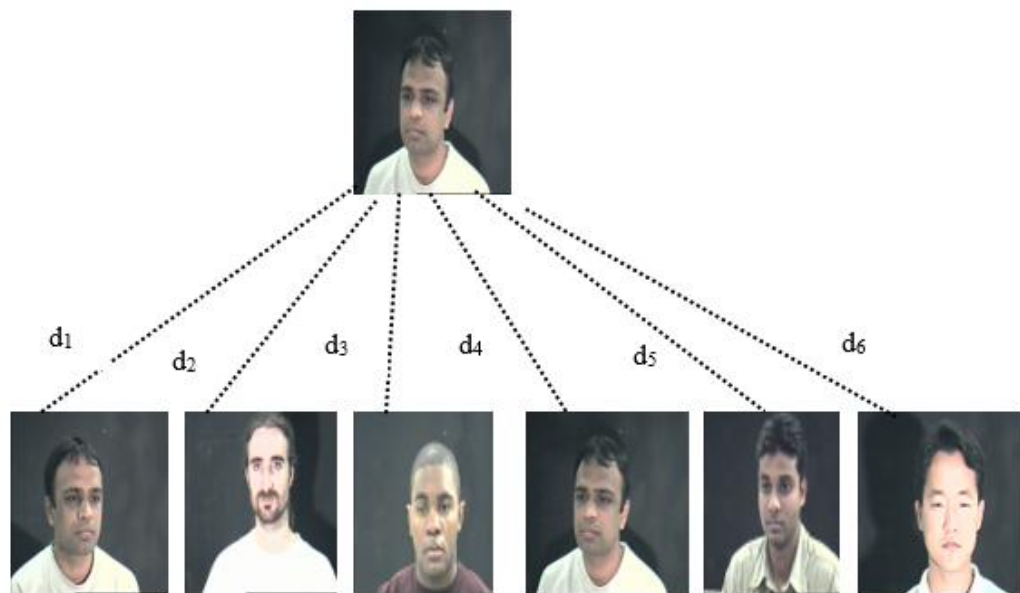


Figure (4- 3) : Représentations de similarité entre images couleurs : de gauche à droite d_1 , d_2 , d_3 , d_4 , d_5 , d_6 des mesures de similarité entre les images respectives et celle au- dessus.

Dans la section suivante, nous proposons une approche basée sur les scores de similarité pour la reconnaissance multispectrale de visage.

4.4. Approche proposée pour la reconnaissance de visage multispectrale à l'aide de score de similarité

La mesure de la similarité est représentée par un scalaire qui désigne une mesure de distance donnant la ressemblance de deux objets. Plus précisément, cette distance est utilisée dans de nombreux domaines : la compression d'image, la mise en correspondance d'images, la biométrie, la reconnaissance de formes, la reconnaissance d'objets, l'industrie manufacturière, l'analyse de données, la fusion d'images, etc. Dans notre étude, nous faisons une évaluation de certaines mesures de similarité rencontrées dans la littérature et proposons une fonction de similarité proposée prenant en compte des caractéristiques d'image. Les fonctionnalités concernées sont les textures et les points d'intérêt. Les données manipulées dans cette étude proviennent de l'imagerie multispectrale en utilisant des images visibles et infrarouges thermiques.

4.4.1. Introduction à la notion de similarité

Le concept de mesure de similarité est transversal. C'est un outil mathématique utilisé dans plusieurs sciences comme l'imagerie, l'analyse de données, la physique, etc. En vision par ordinateur, le concept de distance est utilisé dans plusieurs applications : la biométrie, la reconnaissance d'objets, l'industrie manufacturière, le contrôle qualité, la recherche d'images, l'évaluation de la qualité d'image, l'imagerie médicale, etc.

Vu l'importance de ce sujet, de nombreuses études s'intéressent aux concepts de distance qui sont la similarité et la dissimilarité [112]. Il a été prouvé dans plusieurs études que le choix et l'efficacité de ces métriques dépend du type de données à manipuler.

Une distance est définie comme une longueur entre deux points. La similarité est une valeur utilisée pour quantifier la ressemblance de deux objets en terme de points communs. De même, la mesure de similarité représente le nombre de correspondances tandis que la mesure de dissimilarité calcule le nombre de différences. Généralement, les mesures de similarité sont classées en trois groupes qui sont des mesures géométriques, des mesures de la théorie de l'information et des mesures statistiques. Une étude de Sun-Hyuk Cha [120] présente une classification décrivant de nombreuses familles incluant la famille L_p , la famille L_1 , la famille d'intersection, la famille de produits scalaires, la famille des cordes carrées (famille de fidélité), la famille χ^2 (famille L_2 au carré), la famille d'entropie de Shannon et d'autres.

En général, les mesures de similarité et de dissimilarité utilisées dans la vision par ordinateur sont s'appuient sur deux méthodes. La première méthode concerne l'extraction de caractéristiques tels que la forme, la couleur, la texture, le point clé, etc. Ensuite, la seconde méthode concerne la comparaison d'images par la mise en correspondance d'images, le score de similarité, la corrélation, etc. Notre étude se porte sur l'utilisation du score de similarité dans la reconnaissance faciale multispectrale qui est une méthode biométrique émergente. Pour ce faire, plusieurs distances sont utilisées et évaluées.

Notre contribution dans cette partie est double :

- une analyse de performance de mesures de similarité incluant une fonction de similarité proposée dans cette étude;
- une application du score de similarité dans la reconnaissance faciale multispectrale à l'aide d'images infrarouges visibles / thermiques.

4.4.2. Approche théorique

Cette partie présente d'abord les aspects théoriques de la métrique distance, des mesures de similarité et de dissimilarité. Ensuite, en se basant sur cinq familles de distance, une description des distances est faite.

4.4.2.1. Notion de distance, de similarité et de dissimilarité

La ressemblance de deux images I_1 and I_2 peut être déterminée par une valeur. Cette valeur peut être une mesure de distance, une mesure de similarité ou une mesure de dissimilarité. Une distance notée par *dist* sur un ensemble B est une application de $B \times B$ dans \mathbb{R}^+ satisfaisant les axiomes suivants :

$$(P_1) \text{ identité } \text{dist}(I_1, I_2) = 0 \Leftrightarrow I_1 = I_2 \quad (32)$$

$$(P_2) \text{ symétrie } \text{dist}(I_1, I_2) = \text{dist}(I_2, I_1) \quad (33)$$

$$(P_3) \text{ inégalité triangulaire } \text{dist}(I_1, I_3) \leq \text{dist}(I_1, I_2) + \text{dist}(I_2, I_3) \quad (34)$$

Une dissimilarité notée *dissimil* est une application de $B \times B$ dans \mathbb{R}^+ satisfaisant les axiomes suivants :

$$(P_4) \text{ dissimil}(I_1, I_2) = 0 \Rightarrow I_1 = I_2 \quad (35)$$

$$(P_5) \text{ symétrie } \text{dissimil}(I_1, I_2) = \text{dissimil}(I_2, I_1) \quad (36)$$

Une similarité notée *simil* est une application de $B \times B$ dans \mathbb{R}^+ vérifiant les axiomes suivants :

$$(P_6) \text{ simil}(I_1, I_2) \geq \text{simil}(I_1, I_1) \quad (37)$$

$$(P_7) \text{ symétrie } \text{simil}(I_1, I_2) = \text{simil}(I_2, I_1) \quad (38)$$

Les concepts définis précédemment obéissent aux remarques suivantes :

- une mesure de similarité élevée indique une forte ressemblance entre les images I_1 et I_2 ;
- une mesure de dissimilarité faible indique une forte ressemblance entre les images I_1 et I_2 ;
- une mesure de similarité notée *simil* peut être transformée en une mesure de dissimilarité notée *dissimil*. Ceci est formulé par $dissimil(I_1, I_2) = similmax - simil(I_1, I_2)$, où *similmax* désigne la similarité maximale.

On considère deux vecteurs $X(x_1, x_2, \dots, x_n)$ et $Y(y_1, y_2, \dots, y_n)$ pour les mesures de similarité et de dissimilarité.

4.4.2.2. Famille des distances L_p Minkowski

Les distances de Minkowski sont utilisées dans de nombreuses applications. En vision par ordinateur, les mesures de distance de cette famille sont largement appliquées dans le contexte de reconnaissance de visage, comme la reconnaissance de visage par score de similarité dans cette étude.

La distance de la famille Minkowski ou la distance usuelle est une métrique sur un espace vectoriel normé. La forme générale de cette distance est donnée par la formule (39):

$$d_{\text{mink}} = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p} \quad (39)$$

Plusieurs variantes de cette formule sont obtenues en faisant varier les valeurs de p . C'est la distance euclidienne pour $p = 2$, la distance Manhattan ou City-Block pour $p = 1$ et la distance Chebishev ou Taxicab pour $p = \infty$. Ces distances sont respectivement données par les formules (40), (41) et (42).

$$d_{\text{eucl}} = \sqrt{\sum_{i=1}^n |x_i - y_i|^2} \quad (40)$$

$$d_{\text{mana}} = \sum_{i=1}^n |x_i - y_i| \quad (41)$$

$$d_{\text{cheb}} = \max_i |x_i - y_i| \quad (42)$$

La distance de Minkowski est la mesure de distance la plus répandue, non seulement dans le domaine de la reconnaissance faciale, mais aussi dans de nombreux domaines.

4.4.2.3. Famille des distances L_1

La famille L_1 est très grande et contient des distances dont celles de Sorensen, Gower, Lerentzian, Canberra, Soergel et Kuczynski. Dans cette section, les distances de Kaczynski et Canberra décrites sont présentées respectivement par les formules (43) et (44).

$$d_{\text{kulc}} = \sum_{i=1}^n \frac{|x_i - y_i|}{\sum_{i=1}^d \min(x_i, y_i)} \quad (43)$$

$$d_{\text{canb}} = \sum_{i=1}^n \frac{|x_i - y_i|}{x_i + y_i} \quad (44)$$

Il est à noter que ces distances utilisent la valeur absolue de la différence dans la description de chaque variante.

4.4.2.4. Famille des distances d'intersections

Il existe une multitude de distances dans cette famille. En vision par ordinateur, deux notions essentielles de cette famille sont utilisées. Ces deux notions sont les fonctions de densité de probabilité (PDF) et l'histogramme. L'intersection entre deux fonctions de densité de probabilité est décrite par la formule (45).

$$d_{\text{inte}} = \sum_{i=1}^n \min(x_i, y_i) \quad (45)$$

Concernant la distance entre les histogrammes, elle est présentée par la formule (46).

$$d_{\text{histo}}(X, Y) = \frac{\sum_{i=0}^{l-1} \min(H_X(i), H_Y(i))}{\sum_{i=0}^{l-1} H_Y(i)} \quad (46)$$

où H_X et H_Y sont deux histogrammes.

Issue de cette famille, la distance de Tanimoto est donnée par la formule (47).

$$d_{\text{tani}}(X, Y) = \frac{\sum_{i=1}^n (\max(x_i, y_i) - \min(x_i, y_i))}{\sum_{i=1}^d \max(x_i, y_i)} \quad (47)$$

4.4.2.5. Famille des distances de produit scalaire

Le produit intérieur ou produit scalaire est déterminée par un scalaire. Cette distance est présentée par la formule (48).

$$d_{\text{inne}}(X, Y) = X \bullet Y = \sum_{i=1}^n x_i y_i \quad (48)$$

Plusieurs distances utilisent le produit intérieur pour calculer le score de similarité. Dans cette section, est présentée la similarité cosinus (mesure cosinus) décrite par la formule (49).

$$d_{\text{cosi}}(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (49)$$

La similarité cosinus est basée sur le calcul de l'angle entre deux vecteurs, c'est simplement une métrique angulaire. La mesure cosinus compare les vecteurs caractéristiques de deux images en retournant le cosinus de l'angle entre ces deux vecteurs.

4.4.2.6. Famille des distances de la corde au carré

La famille de la corde au carré est également appelée famille de fidélité. La formule de base de la corde au carré est donnée par la formule (50).

$$d_{\text{sqeu}} = \sum_{i=1}^d (x_i - y_i)^2 \quad (50)$$

La famille de fidélité contient les distances de Matusita, Fidélité et Battacharrya respectivement présentées par les formules (51), (52) et (53).

$$d_{\text{matu}} = \sqrt{\sum_{i=1}^n (\sqrt{x_i} - \sqrt{y_i})^2} \quad (51)$$

$$d_{\text{batta}} = -\ln \sum_{i=1}^d \sqrt{x_i y_i} \quad (52)$$

$$d_{\text{fide}} = \sum_{i=1}^n \sqrt{x_i y_i} \quad (53)$$

4.4.2.7. Famille des distances L_2 carré

La famille L_2 carré ou famille χ^2 est aussi appelée famille statistique. La mesure de base utilisée est la distance euclidienne carrée donnée par la formule (54).

$$d_{\text{sqeu}} = \sum_{i=1}^n (x_i - y_i)^2 \quad (54)$$

Plusieurs autres distances sont construites à partir de cette distance comme la distance Pearson χ^2 et la distance chi-carré définie par les formules (55) et (56).

$$d_{\text{pea2}}(\mathbf{P}, \mathbf{Q}) = \sum_{i=1}^n \frac{(x_i - y_i)^2}{y_i} \quad (55)$$

$$d_{\text{chi2}}(\mathbf{H}_X, \mathbf{H}_Y) = \sum_{i=0}^{l-1} \frac{(H_X(i) - m(i))^2}{m(i)} \quad (56)$$

où $m(i) = \frac{H_x(i) + H_y(i)}{2}$, H_x et H_y sont deux histogrammes.

4.4.2.8. Famille des distances d'entropie de Shannon

La famille d'entropie de Shannon est également appelée famille des mesures de divergence.

Dans les probabilités et les théories de l'information, la divergence, appelée aussi fonction de contraste, est une fonction qui donne la mesure de deux distributions de probabilités en utilisant des notions statistiques. La divergence peut ne pas être symétrique et obéir à l'inégalité triangulaire.

En théorie, la divergence de Kullback-Leibler, aussi appelée entropie relative, est une mesure de dissimilarité entre deux distributions de probabilités X et Y . Elle considère les histogrammes comme des distributions et donne la dissemblance en calculant l'entropie relative. La divergence de Kullback-Leibler est décrite par la formule (57).

$$d_{\text{kuLe}}(X, Y) = \sum_{i=1}^n x_i \ln \frac{x_i}{y_i} \quad (57)$$

Le principe de la divergence de Jeffrey consiste à mesurer la dissemblance de deux variables aléatoires en théorie de l'information. Il utilise la notion d'information discriminante moyenne entre les deux signaux aléatoires. La divergence de Jeffrey donnée par la formule (58) a l'avantage d'être la version symétrique de la divergence de Kullback-Leibler.

$$d_{\text{jeff}}(X, Y) = \sum_{i=1}^n (x_i - y_i) \ln \frac{x_i}{y_i} \quad (58)$$

La divergence de Jensen-Shannon appelée divergence totale à la moyenne permet de mesurer la similarité entre deux distributions de probabilité. La divergence de Kullback-Leibler a l'avantage d'être symétrique et donne une valeur finie. La formule (59) décrit la divergence de Jensen-Shannon.

$$d_{\text{jesh}}(X, Y) = \frac{1}{2} (\sum_{i=1}^n (x_i - y_i) \ln \frac{x_i}{y_i}) \quad (59)$$

En plus des distances décrites précédemment, cette étude présente deux autres distances qui sont la distance de Mahalanobis et la distance du cantonnier appelée. Earth Mover's Distance (EMD).

Utilisée en statistique, la distance de Mahalanobis a été introduite par Prasanta Chandra Mahalanobis en 1936. Son principe repose sur des corrélations entre les différentes caractéristiques des vecteurs caractéristiques. La distance de Mahalanobis est une version généralisée de la distance euclidienne qui est invariante à l'échelle et tient en compte les corrélations d'ensembles de données. En résumé, la distance de Mahalanobis donne la corrélation entre deux séries de mesures. Il a l'avantage de prendre en compte la variance et la corrélation des séries de données. Cette distance décrite par la formule (60) permet de mesurer la dissemblance entre deux vecteurs de même distribution.

$$d_{\text{Maha}}(X, Y) = \sqrt{(x - y)^T S^{-1} (x - y)} \quad (60)$$

ou S est la matrice de covariance et $(x-y)^T$ est la transposé de la différence.

La distance EMD peut être utilisée pour comparer deux histogrammes afin d'évaluer une mesure de similarité à l'aide d'un algorithme d'optimisation. En recherche opérationnelle cet algorithme est utilisé pour des problèmes de transport optimal. Dans la vision par ordinateur, cette distance est l'énergie minimale requise pour transformer une distribution en une autre. Cette méthode s'appuie sur la recherche du débit minimum F en considérant tous les d_{ij} de transport pour transformer l'histogramme X en un histogramme Y . Cette notion est présentée dans la formule (61).

$$d_{\text{EMD}}(X, Y) = \min_F \sum_i \sum_j f_{ij} d_{ij} \quad (61)$$

Les concepts de distance sont décrits dans la littérature avec plusieurs applications [29].

4.4.3. Méthode proposée pour l'étude de similarité

4.4.3.1. Corrélacion et fusion d'images

L'imagerie multispectrale améliore la capacité et la fiabilité de l'interprétation de l'image.. Compte tenu de son succès, la fusion d'images est utilisée dans de nombreux domaines de la vision par ordinateur, y compris la reconnaissance faciale [59, 60]. Cette étude s'intéresse à la reconnaissance multispectrale du visage à l'aide d'images visibles et infrarouges thermiques.

Pour la fusion d'images, des paires d'images infrarouges visibles et thermiques sont utilisées. Tout d'abord, le calcul de l'entropie montre que les images infrarouges sont plus riches en informations que les images visibles. La formule d'entropie nommée H_e est donnée par la formule (19) au chapitre 2.

Ensuite, la corrélation des paires d'images a été prouvée par l'information mutuelle (MI) [129, 130, 131] donnant l'information commune et la mesure de similarité par la somme des carrés de différences (SSD) pour vérifier la similarité. Ces deux notions sont décrites par les formules (20) et (21) au chapitre 2.

Pour l'évaluation de la qualité de l'image, quatre différentes métriques sont utilisées [4, 131]. Ces métriques sont l'erreur quadratique moyenne (RMSE), l'erreur d'ajustement en pourcentage (PFE), l'erreur absolue moyenne (MAE), le rapport d'un signal de crête sur bruit (PSNR).

Les différentes métriques utilisées, montrent que la méthode de fusion de PCA [62, 124] donne la meilleure fusion d'image avec quelques résultats présentés dans la figure 1.



Figure (4- 4) : Ensemble d'images extraites de la base de données. Chaque ligne de haut en bas présente des images visibles, des images infrarouges thermiques correspondantes, des images fusionnées correspondantes par la méthode de fusion PCA

4.4.3.2. Extraction de caractéristiques

Dans le domaine de la reconnaissance d'objets, l'extraction de caractéristiques consiste à quantifier les propriétés mesurables dans une image. Les expériences de cette étude ont porté sur les caractéristiques qui sont les points d'intérêt et la texture.

En ce qui concerne la texture, des notions statistiques sont utilisées pour estimer les textures par la mesure locale et la mesure de la dispersion.

D'abord, la moyenne permet de faire une mesure locale. La moyenne est obtenue par la somme de tous les composants divisés par le nombre total de composants comme indiqué par la formule (62).

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad (62)$$

S'agissant de la mesure de la dispersion, l'écart-type et la variance sont indiqués. Quant à la variance, elle mesure la dispersion variable. Il est défini comme la racine carrée de la variance. La variance et l'écart type sont donnés par les formules (63) et (64)

$$\text{var}(X) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 \quad (63)$$

$$\text{std}(X) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2} \quad (64)$$

Concernant les caractéristiques de texture, le coefficient d'aplatissement (skewness) et le coefficient de dissymétrie (kurtosis) sont utilisés. Ces deux mesures caractérisent l'emplacement et la variabilité de l'ensemble de données.

Le coefficient d'aplatissement est une mesure classique de symétrie. La valeur de symétrie est utilisée pour interpréter l'étalement de la distribution. La formule (65) présente le skewness.

$$\text{skewness}(X) = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^3}{s^3} \quad (65)$$

où \bar{X} est la moyenne, s la variance, n le nombre de données.

Dans la théorie des probabilités et statistiques, kurtosis est une mesure de la «pointe» de la distribution de probabilité d'une variable aléatoire à valeur réelle. C'est un descripteur de la forme d'une distribution de probabilité. La valeur du kurtosis est

une mesure de dissymétrie d'une distribution variable. La formule (66) donne kurtosis qui est un paramètre de forme.

$$\text{Kurtosis (x)} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^4}{s^4} \quad (66)$$

où \bar{X} est la moyenne, s la variance, n le nombre de données.

Pour l'extraction de caractéristiques, des points d'intérêt ont également été utilisés. L'extraction du points-clés est classiquement composée de trois étapes : la détection du point du point d'intérêt, la description du point d'intérêt et la mise en correspondance des images. L'un des détecteurs les plus connus est Scale Invariant Feature Transform (SIFT) [82].

Une étude comparative des descripteurs et des détecteurs dans la reconnaissance multispectrale de visages a montré l'utilisation de descripteurs binaires. Les résultats de notre étude ont montré qu'il existe un bon compromis en utilisant le détecteur SURF (Speed Up Robust Features) [128] et le descripteur de points d'intérêt Fast Retina Key-point (FREAK). Ceci justifie l'utilisation du détecteur SURF et du descripteur FREAK pour l'extraction des points d'intérêt.

La signature d'une image peut être représentée par un vecteur, un ensemble de vecteurs, des graphes, etc. Dans cette expérience, pour la construction des signatures, nous avons utilisé la notion de vecteurs dont les composantes sont std, skewness, kurtosis, entropie, points d'intérêt désignant respectivement l'écart-type, le coefficient d'aplatissement, la mesure de dissymétrie, l'entropie et les points-clés obtenus par la méthode SURF / FREAK.

De plus, les caractéristiques extraites sont prises en compte pour le calcul de la fonction de similarité notée d_{MSF} , en utilisant la distance euclidienne comme indiqué dans le tableau (4-3).

Tableau (4- 3) : Fonction de similarité pour calculer le score de similarité de deux images X et Y.

Fonction distance= $d_{MSF}(X, Y)$

$x_1 \leftarrow \text{std}(X)$
 $x_2 \leftarrow \text{skewness}(X)$
 $x_3 \leftarrow \text{kurtosis}(X)$
 $x_4 \leftarrow \text{entropy}(X)$
 $x_5 \leftarrow \text{surffreak}(X)$
 $A = [x_1, x_2, x_3, x_4, x_5]$
 $y_1 \leftarrow \text{std}(Y)$
 $y_2 \leftarrow \text{skewness}(Y)$
 $y_3 \leftarrow \text{kurtosis}(Y)$
 $y_4 \leftarrow \text{entropy}(Y)$
 $y_5 \leftarrow \text{surffreak}(Y)$
 $B = [y_1, y_2, y_3, y_4, y_5]$
distance $\leftarrow \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2 + (x_5 - y_5)^2}$
retourner (distance)

Dans la méthode que nous proposons, les caractéristiques extraites sont la texture et les points d'intérêt. La combinaison de ces deux caractéristiques caractérise mieux une image avec l'avantage d'être robuste et invariante. Pour ce faire, des vecteurs de caractéristiques et une fonction de similarité basée sur la distance euclidienne sont construits. Ensuite, une correspondance d'image utilise un calcul de similarité appliquant la fonction de similarité d_{MSF} , une autre distance, l'information mutuelle, une correspondance d'image à partir des points clés obtenus et l'édition d'une carte de similarité. Enfin, les résultats obtenus permettent de prendre des décisions. Ce processus peut être récapitulé dans la figure (4-5).

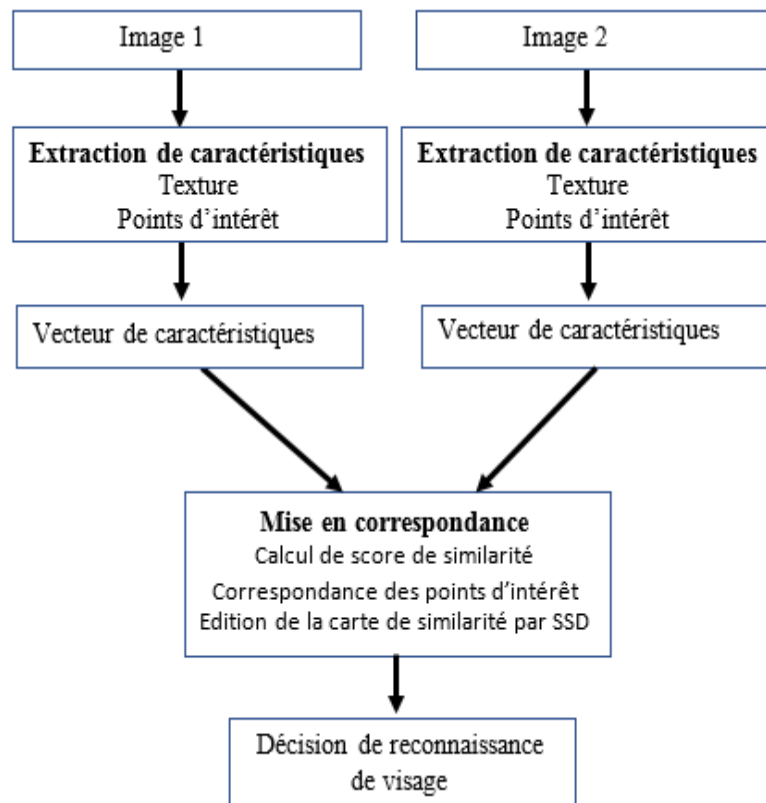


Figure (4- 5) : Schéma proposé pour la reconnaissance de visage à l'aide de score de similarité

4.4.4. Résultats expérimentaux

4.4.4.1. Base de données

Pour les expériences, la base de données est extraite de la base de données IRIS (Imaging Robotics Intelligent System) [47]. Cette base de données IRIS contient des images visibles et infrarouges thermiques de visages. Dans cette expérience, sont utilisées des images de visage acquises dans le visible et l'infrarouge. Sont concernées 14 personnes avec 4 poses de visages et 3 illuminations qui sont Lon (Left Light On), Off (Left and Right Off) and Ron (Right Light On). Au total, 336 paires d'images ont été fusionnées soient 168 images visibles/ infrarouges.

4.4.4.2. Description de l'expérience

Afin d'évaluer les méthodes de reconnaissance faciale multispectrales analysées, trois types d'expériences ont été réalisées : la variation d'illumination, la variation de poses et la combinaison de poses et d'illuminations. Les expériences sont présentées ci-dessous.

Variation d'illuminations. Dans cette expérience, pour un sujet donné, une position est fixée et les conditions d'illumination Lon, Ron et Off sont variées. La même expérience est répétée pour chaque direction.

Variations de poses. Concernant cette expérience, pour un sujet donné, une illumination est fixée et les quatre poses sont variées. La même expérience est répétée pour chaque illumination.

Variations d'illuminations et de poses. Dans cette expérience, sont prises en compte la variation du sujet, l'illumination et la pose.

4.4.4.3. Résultats

Pour l'expérience, soixante cas ont été testés. Un ensemble d'images est présenté à la figure (4- 5). Cet ensemble présente :

- un même individu, la même pose et une variation d'illumination;
- un même individu, la même illumination, une variation de poses;
- différents individus, la même pose, la même illumination;
- différents individus, différentes poses et illuminations.

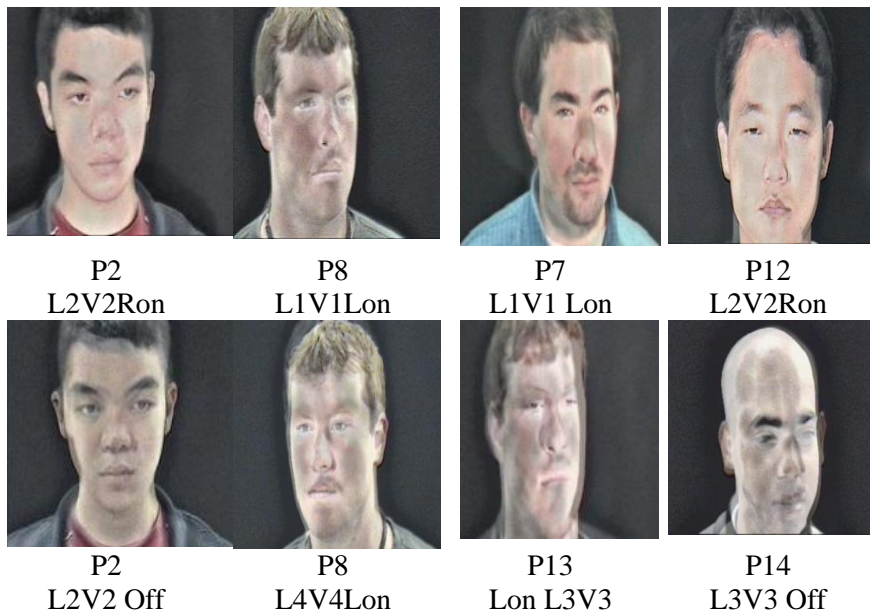


Figure (4- 6) : Images utilisées pour le calcul du score de similarité.

Les résultats obtenus sont récapitulés ans le tableau ci-dessous.

Tableau (4- 4) : Résultats du score de similarité entre les images en utilisant les distances décrites précédemment et la fonction de similarité proposée.

Images Distances	P2 L2V2 Ron et P2 L2V2 Off	P8 L1V1 Lon et P8 L4V4 Lon	P7L1V1 Lon et P8 L4V4 Ron	P12 L2V2Ron et P14 L3V3 Off
Euclidean	4.14	1.53	0.38	3.23
Manhattan	7.94	2.82	0.71	6.11
Chebishev	2.70	1.07	2.26	2.05
Kulzinski	54.21	17.72	3.45	70.27
Canberra	15.95	6.72	1.64	15.14
Intersection	6.03	8.59	9.64	6.94
histogram intersection	3.97	1.41	0.35	3.06
Tanimoto	22.72	11.04	3.13	20.84
Cosine	2.70	0.30	0.22	1.34
Matusita	9.02	0.97	0.41	3.70
Battacharrya	0.87	0.14	0.01	0.71
fidelity	9.16	9.87	9.99	9.32
Jeffrey divergence	1.65	0.27	0.02	1.32
Pearson chi2	6.91	0.92	0.07	11.36
Chi2	1.6	0.26	0.02	1.23
Kullback-leiber	3.86	0.50	0.03	3.26
Jensen-shanon	0.82	0.14	0.01	0.66
EMD	85.11	29.53	16.61	64.19
d_{MSF}	12.88	14.01	49.77	16.27

En observant le tableau (4- 4), les images fusionnées visibles/infrarouges sont les meilleures candidates pour la reconnaissance faciale. Par conséquent, le score de similarité est un moyen efficace de reconnaissance faciale multispectrale pour améliorer les méthodes traditionnelles. Les différents tests et résultats confirment cette hypothèse

par l'utilisation de dix-neuf distances incluant la distance proposée. Cependant, les difficiles conditions d'environnement ont une influence sur la reconnaissance faciale.

Conclusion

La performance de la reconnaissance de visage multispectrale par rapport à la reconnaissance faciale traditionnelle est un problème majeur. Cette étude se porte sur une évaluation de la mesure de similarité dans la reconnaissance faciale multispectrale en utilisant des images fusionnées à partir d'images visibles et des images infrarouges thermiques. Une distance peut désigner la ressemblance de deux images. Cette similarité peut être robuste en prenant en compte des propriétés spécifiques. Toutes les distances présentées sont utilisables dans l'imagerie multispectrale mais leur qualité varie selon les conditions d'environnement telles que l'éclairage et les poses. La distance proposée suit la même tendance mais a l'avantage d'utiliser des caractéristiques comme des textures et des points d'intérêt qui donnent plus de précision.

Enfin, comme travaux futurs de cette étude, la reconnaissance faciale émergente comme la reconnaissance multispectrale du visage s'intéresse à plusieurs aspects tels que l'obscurité totale, les déguisements (maquillages, chapeaux, lunettes, etc.), l'occlusion partielle, la distance, les expressions de visage. À l'avenir, nous travaillerons sur différentes images spectrales séparément et leur combinaison.

Conclusion générale et perspectives

Dans ce travail, il a été étudié la reconnaissance des visages dans le domaine multispectral. L'objectif a consisté à faire une analyse de la performance des techniques de reconnaissance de visage dans les spectres visibles, infrarouges et de proposer des solutions susceptibles d'améliorer les résultats de la reconnaissance faciale traditionnelle. Les approches proposées ont permis d'améliorer de manière significative les résultats de reconnaissance faciale.

L'extraction des caractéristiques en reconnaissance de visage permet de rendre plus robuste les méthodes de reconnaissance. Les caractéristiques utilisées sont la texture et les points d'intérêt. Il s'agit d'extraire des caractéristiques invariantes du visage dans le domaine multispectral par l'utilisation du visible et de l'infrarouge thermique. A cet effet, une technique de fusion a permis de combiner des spectres visibles et infrarouges thermiques afin d'obtenir une représentation qui tient compte des avantages de chacun des spectres. Toutefois, sa performance dans des conditions difficiles de variations d'éclairage et de poses la qualité de reconnaissance se détériore.

Dans nos travaux de recherche, les contributions suivantes ont été apportées :

- une étude de la reconnaissance des visages dans le domaine multispectral : visible, et infrarouge thermique;
- élaboration d'une nouvelle base de données multispectrale des visages par fusion d'images visibles et infrarouges thermiques à l'aide de paires d'images extraites de la base de données IRIS;
- une nouvelle approche nommée MS-FRHF pour l'extraction des caractéristiques des visages basée sur une extraction hybride de caractéristiques qui sont la texture et les points d'intérêt;
- une stratégie de reconnaissance basée sur l'utilisation de techniques d'extraction binaire de caractéristiques. D'abord, une analyse a permis de trouver un bon compromis de choix de détecteurs et descripteurs. Ensuite, la combinaison de SURF comme détecteur et FREAK comme descripteur sont appliqués pour extraire les points d'intérêt avec une bonne performance afin de permettre une mise en correspondance des images multispectrales ;

- une méthode de reconnaissance de visage basée sur les mesures de similarité. L'analyse de similarité est un problème majeur en vision par ordinateur. Après une évaluation de certaines mesures de similarité, une fonction de similarité nommée d_{MSF} est proposée tenant compte des caractéristiques de l'image qui sont la texture et les points d'intérêt. Dans le processus de reconnaissance, les dites caractéristiques sont extraites et ensuite des vecteurs de caractéristiques sont construits et utilisés pour calculer la distance proposée. En sus dans une phase de mise en correspondance, sont prises en compte la mesure de similarité, la correspondance des points d'intérêt et l'édition de la carte de similarité. Les indicateurs issus de cette phase de correspondance permettent la prise de décision de la reconnaissance de visage.

Perspectives

Les résultats des expériences permettent d'améliorer la performance de reconnaissance du visage et de réduire l'impact de diverses conditions qui la détériore habituellement. Des travaux futurs peuvent être initiés pour approfondir certaines approches tels que la multimodalité et le 3D tout en exploitant l'imagerie multispectrale.

Enfin, en prolongement de ce travail, l'imagerie multispectrale s'intéresse à d'autres défis tels que l'obscurité totale, les déguisements (maquillages, chapeaux, lunettes, etc.) et les expressions faciales. Dans nos travaux futurs, nous allons explorer la façon de combiner différentes images spectrales ou exploiter des bandes spectrales séparément. D'autres techniques dans la littérature sont en développement pour améliorer la qualité de reconnaissance de visage notamment le 3D, l'extraction des caractéristiques physiologiques, la tomographie ou la multi-modalité du visage avec un autre trait biométrique.

Bibliographie

- [1] “A battle of three descriptors: SURF, FREAK, and BRISK”, <http://computer-vision-talks.com/2012/08/a-battle-of-three-descriptorsurf-freak-and-brisk/> [dernière visite le 10/01/2018]
- [2] Alahi, R. Ortiz, and P. Vandergheynst. FREAK: Fast Retina Keypoint. In IEEE Conference on Computer Vision and Pattern Recognition, 2012.
- [3] Abdel-Hakim, A.E., Farag, A.A. CSIFT: A-SIFT descriptor with color invariant characteristics. In Computer Vision and Pattern Recognition (CVPR 2006), IEEE, Vol. 2, pp 1978-1983. 17-22, June 2006.
- [4] Ahmet M. Eskicioglu, Paul S. Fisher. Image Quality Measures and Their Performance. IEEE Transactions on Communication, Vol. 43, No. 12, pp. 2959-2965, December 1995.
- [5] Akhloufi Moulay abdellatif. Reconnaissance des visages par imagerie Multispectrale, Thèse présentée à l’université de Laval, 2013.
- [6] Akhloufi, Moulay, Bendada, Abdelhakim, et Batsale, Jean-Christophe. State of the art in infrared face recognition. Quantitative InfraRed Thermography Journal, vol. 5, no 1, pp. 3-26, 2008.
- [7] Bay, H., Tuytelaars, T., Gool, L.V. SURF: Speeded up robust features. In *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision*, pp.7-13 May 2006. Springer, Part II, 404-417.
- [8] Bay, Herbert, Fasel, Beat, et Van Gool, Luc. Interactive museum guide: Fast and robust recognition of museum objects. 2006.
- [9] Bekele, Dagmawi, Teutsch, Michael, et Schuchert, Tobias. Evaluation of binary keypoint descriptors. In: Image Processing (ICIP), 2013 20th IEEE International Conference on. IEEE, pp. 3652-3656, 2013.
- [10] Bendada, Abdelhakim, Akhloufi, Moulay A., Multispectral face recognition in texture space, CRV 2010 - 7th Canadian Conference on Computer and Robot Vision, pp 101-106, 2010.
- [11] Bhanu, Bir et Govindaraju, Venu. *Multibiometrics for Human Identification*. Cambridge University Press, 2011.

- [12] Bowyer, Kevin W., Chang, Kyong, et FLYNN, Patrick. A survey of approaches to three-dimensional face recognition. In : Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. IEEE, pp. 358-361, 2004.
- [13] Brodatz, Textures: A Photographic Album for Artists and Designers, 1966.
- [14] Bruce, Vicki et Young, Andy. Understanding face recognition. British journal of psychology, vol. 77, no 3, p. 305-327, 1986.
- [15] Buddharaju, Pradeep et Pavlidis, Ioannis. Multispectral face recognition: fusion of visual imagery with physiological information. *Face biometrics for personal identification*, pp. 91-108, 2007.
- [16] BURGE, Mark J. et MONACO, Matthew K. Multispectral iris fusion for enhancement, interoperability, and cross wavelength matching. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. pp. 73341D-73341D, 2009.
- [17] C. Harris and M. Stephens. A combined corner and edge detector. Alvey Vision Conference, pp. 147-151, 1988.
- [18] Chang H, Koschan A, Abidi B, Abidi M. Fusing continuous spectral images for face recognition under indoor and outdoor illuminants. *Mach Vis Appl* 21(2): pp 201–215, 2010.
- [19] Chang, Kyong I., Bowyer, Kevin W., et Flynn, Patrick J. An evaluation of multimodal 2D+ 3D face biometrics. *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no 4, pp. 619-624, 2005.
- [20] Chen L, Zheng H, Li L, Xie P, Liu S. Near-infrared dorsal hand vein image segmentation by local thresholding using grayscale morphology. In: 1st international conference bioinformatics and biomedical engineering, 2007.
- [21] Cormack, Richard M. A review of classification. *Journal of the Royal Statistical Society. Series A (General)*, pp. 321-367, 1971.
- [22] D. Zhang et al., *Multispectral Biometrics*, Springer International Publishing Switzerland 2016.
- [23] Dalal, Navneet et Triggs, Bill. Histograms of oriented gradients for human detection. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE*, pp. 886-893, 2005.

- [24] Dalka, Piotr et Czyzewski, Andrzej. Lip movement and gesture recognition for a multimodal human-computer interface. In : Computer Science and Information Technology, 2009. IMCSIT'09. International Multiconference on. IEEE. pp. 451-455, 2009.
- [25] Dass, Sarat C., Nandakumar, Karthik, et JAIN, Anil K. A principled approach to score level fusion in multimodal biometric systems. In: *AVBPA*. pp. 1049-1058, 2005.
- [26] Daugman, John G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A*, vol. 2, no 7, pp. 1160-1169., 1985.
- [27] Delac, Kresimir et GRGIC, Mislav. A survey of biometric recognition methods. In: *Electronics in Marine*, 2004. Proceedings Elmar. 46th International Symposium. IEEE, pp. 184-193, 2004.
- [28] Denes, Louis J., Metes, Peter, et LIU, Yanxi. *Hyperspectral face database*. Carnegie Mellon University, The Robotics Institute, 2002.
- [29] Deza E. and Deza M.M., *Dictionary of Distances*, Elsevier, 2006
- [30] Rublee, Ethan, Rabaud, Vincent, Konolige, Kurt, et Bradski, G. ORB: An efficient alternative to SIFT or SURF. In : Computer Vision (ICCV), 2011 IEEE international conference on. IEEE, 2011. pp. 2564-2571, 2011.
- [31] E. Tola, V. Lepetit, and P. Fua. A fast local descriptor for dense matching. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [32] Eskandari, Maryam Toygar, Önsen Demirel, Hasan. A New Approach for Face-Iris Multimodal Biometric Recognition Using Score Fusion. *International Journal of Pattern Recognition and Artificial Intelligence*. Vol 27, issue 03, 2013.
- [33] G. Angeline Prasanna, K. Anandakumar, A. Bharathi . Multi Modal Biometric Systems: A State of the Art Survey. *International Research Journal of Engineering and Technology (IRJET)*. Volume: 03, Issue: 04 ,Apr-2016.
- [34] Gandhimathi Amirthalingam, Radhamani, Research Scholar. A Multimodal Approach for Face and Ear Biometric System. *IJCSI International Journal of Computer Science Issues*, Vol. 10, Issue 5, No 2, September 2013.

- [35] GHIASS, Reza Shoja, ARANDJELOVIĆ, Ognjen, BENDADA, Abdelhakim, *et al.* Infrared face recognition: A comprehensive review of methodologies and databases. *Pattern Recognition*, 2014, vol. 47, no 9, p. 2807-2824.
- [36] Grgic, Mislav, Delac, Kresimir, et GRGIC, Sonja. SCface—surveillance cameras face database. *Multimedia tools and applications*, vol. 51, no 3, pp. 863-879, 2011.
- [37] Gudavalli, Madhavi, Raju, S. Viswanadha, Babu, A. Vinaya, et al. Multimodal Biometrics--Sources, Architecture and Fusion Techniques: An Overview. In : *Biometrics and Security Technologies (ISBAST)*, International Symposium on. IEEE, pp. 27-34, 2012.
- [38] H. Moravec. Towards automatic visual obstacle avoidance. *International Joint Conferences on Artificial Intelligence*, pp. 584, 1977.
- [39] Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt, Gang Hua. A Convolutional Neural Network Cascade for Face Detection, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 5325-5334, 2015.
- [40] Heinly, Jared, Dunn, Enrique, et Frahm, Jan-Michael. Comparative evaluation of binary features. In : *Computer Vision—ECCV 2012*. Springer, Berlin, Heidelberg, pp. 759-773, 2012.
- [41] Hong Chang, Andreas Koschan, Mongi Abidi, Seong G. Kong, Chang-Hee Won. Multispectral Visible and Infrared Imaging for Face recognition, in *Computer Vision and Pattern Recognition Workshops (CVPRW 08)*. IEEE Computer Society conference, pp 1-6, 2008
- [42] <http://vcipl-okstate.org/pbvs/bench/Data/04/download.html> [dernière visite le 10/01/2018]
- [43] <http://www.equinoxsensors.com/products/HID.html> [dernière visite le 10/01/2018]
- [44] <https://sites.google.com/a/nd.edu/public-cvrl/data-sets> [dernière visite le 10/01/2018]
- [45] <https://wsri.wright.edu/> [dernière visite le 10/01/2018]
- [46] Husken, Michael, Brauckmann, Michael, Gehlen, Stefan, et al. Strategies and benefits of fusion of 2D and 3D face recognition. In : *Computer Vision and*

- Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on. IEEE, pp. 174-174, 2005.
- [47] IRIS-M database. <http://imaging.utk.edu>[dernière visite le 10/01/2018]
- [48] J. Heo, B. Abidi, S. G. Kong, and M. Abidi, Performance Comparison of Visual and Thermal Signatures for Face Recognition. Biometric Consortium Conference, pp 164, 2003
- [49] Jain, Anil K. et LI, Stan Z. Handbook of face recognition. New York: springer, 2011.
- [50] Jain, Anil K. et Park, Unsang. Facial marks: Soft biometric for face recognition. In : Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE, pp. 37-40, 2009.
- [51] Jain, Anil K., Dass, Sarat C., et Nandakumar, Karthik. Soft biometric traits for personal recognition systems. In : Biometric Authentication. Springer, Berlin, Heidelberg, pp. 731-738, 2004.
- [52] Jain, Anil K., Ross, Arun, et Prabhakar, Salil. An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology, vol. 14, no 1, p. 4-20, 2004.
- [53] Juwei Lu, K.N. Plataniotis, and A.N. Venetsanopoulos. Face Recognition Using LDA Based Algorithms. in IEEE transactions on neural networks, accepted in may 2002.
- [54] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. IEEE Pattern Analysis and Machine Intelligence, 27(10):1615- 630, 2005.
- [55] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. International Journal of Computer Vision, 1(60) :63 - 6, 2004.
- [56] Kakadiaris, Ioannis A., Passalis, Georgios, Toderici, George, *et al.* 3D Face Recognition. In : *BMVC*. pp. 869-878, 2006.
- [57] Ke, Y., Sukthankar, R. PCA-SIFT: A more distinctive representation for local image descriptors. In Computer Vision and Pattern Recognition (CVPR 2004), 27 June – 2 July 2004. IEEE, Vol. 2, pp. 506-513, 2004.
- [58] Khalil MS, Muhammad D, AL-Nuzaili Q. Fingerprint verification using the texture of fingerprint image. In: Second international conference, machine vision, 2009.

- [59] Kong, Seong G., HEO, Jingu, Boughorbel, Faysal, et al. Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. *International Journal of Computer Vision*, vol. 71, no 2, pp. 215-233, 2007.
- [60] Koschan A, Yao Y, Chang H, Abidi M. Multispectral face imaging and analysis. *Handbook of face recognition*. Springer, London, pp 401–428, 2011.
- [61] Krishna Rao Kakkirala, Srinivasa Rao Chalamala, SantoshKumar Jami, Thermal Infrared Face Recognition: A review: UKSim-AMSS 19th International Conference on Modelling & Simulation, 2017.
- [62] Kusum Rani¹, Reecha Sharma. Study of Different Image fusion Algorithm *International Journal of Emerging Technology and Advanced Engineering* Volume 3, Issue 5, May 2013
- [63] Le, Chien et JAIN, R. A survey of biometrics security systems. *EEUU. Washington University in St. Louis*, 2009.
- [64] LERMAN, Israël-César et PETER, Philippe. Indice probabiliste de vraisemblance du lien entre objets quelconques: analyse comparative entre deux approches. *Revue de statistique appliquée*, 2003, vol. 51, no 1, p. 5-36.
- [65] Li, Stan Z. et JAIN, Anil. *Encyclopedia of biometrics*. Springer Publishing Company, Incorporated, 2015.
- [66] Li, Stan Z., Chu, Ru Feng, LIAO, Sheng Cai, et al. Illumination invariant face recognition using near-infrared images. *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no 4, pp. 627-639, 2007.
- [67] Lindeberg, Tony. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, vol. 11, no 3, p. 283-318, 1993.
- [68] Liton Chandra Paul , Abdulla Al Sumam. Face Recognition Using Principal Component Analysis Method. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* Volume 1, Issue 9, November 2012.
- [69] Louis J. Denes, Peter Metes and Yanxi Liu. Hyperspectral Face Database. Tech. Report, CMU-RI-TR-02-25, Robotics Institute, Carnegie Mellon University, October, 2002

- [70] M. Calonder, V. Lepetit, C. Strecha, and P. Fua. Brief: Binary robust independent elementary features. *Computer Vision– ECCV 2010*, pages 778–792, 2010.
- [71] M. kowalski, A. grudzien, N. Palka, M. Szustakowski, Face recognition in thermal infrared domain. *Proc. SPIE 10441, Counterterrorism, Crime Fighting, Forensics, and Surveillance technologies 1044109*. 5 October 2017.
- [72] M. Petrou and P.G. Sevilla. *Image Processing: Dealing with Texture*. Wiley, Chichester, UK, 2006.
- [73] M. Smith and J. Brady. Susan - a new approach to low level image processing. *International Journal of Computer Vision*, 23 :45-78, 1997.
- [74] M. Tuceryan and K. Jain. Texture analysis. In C. H. Chen and P.S P. Wang L.F. Pau, editors, *The handbook of pattern recognition and Computer vision*. World Scientific Publishing Co. 1998.
- [75] Mair, Elmar, Hager, Gregory D., Burschka, Darius, et al. Adaptive and generic corner detection based on the accelerated segment test. In: *European conference on Computer vision*. Springer, Berlin, Heidelberg, pp. 183-196, 2010.
- [76] Mallat S., . A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989.
- [77] Mamta Ahlawat, Chander Kant. An Introduction to Multimodal Biometric System: An Overview, *IJSRD - International Journal for Scientific Research & Development* | Vol. 3, Issue 02, pp 1150, 1154, 2015
- [78] Mandelbrot, Benoit B. *Fractals*. John Wiley & Sons, Inc., 1977.
- [79] Manhal S. Almohammad, Gouda I. Salama and Tarek A. Mahmoud. Face and Gait Fusion Methods: A Survey. *International Journal of Computer Science and Telecommunications*, Volume 4, Issue 4, April 2013.
- [80] Metty Mustikasari, Sarifuddin Madenda. Texture Based Image Retrieval Using GLCM and Image Sub-Block, *IJARCSSE*, Volume 5, Issue 3, 2015.
- [81] Mieziako, Roland. Terravic research infrared database. *IEEE OTCBVS WS Series Bench*, 2005.

- [82] Mikolajczyk, Krystian et Schmid, Cordelia. A performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no 10, pp. 1615-1630, 2005.
- [83] Mikolajczyk, Krystian et Schmid, Cordelia. Scale & affine invariant interest point detectors. *International journal of computer vision*, vol. 60, no 1, pp. 63-86, 2004.
- [84] Mir, A. H., Rubab, S., et Jhat, Z. A. Biometrics verification: a literature survey. *International Journal of Computing and ICT Research*, vol. 5, no 2, pp. 67-80, 2011.
- [85] Mishra, Ashish. Multimodal biometrics it is: need for future systems. *International Journal of Computer Applications*, vol. 3, no 4, pp. 28-33, 2010.
- [86] Mohamed Soltane, Nouredine Doghmane, Nouredine Guersi. Face and Speech Based Multi-Modal Biometric Authentication. *International Journal of Advanced Science and Technology*, Vol. 21(8), pp. 41-46, 2010.
- [87] Montesinos, Philippe, Gouet, Valérie, et Deriche, Rachid. Differential invariants for color images. In: *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*. IEEE, pp. 838-840, 1998.
- [88] Morel, J.M., Yu, G. ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2 (2), 438-469, 2009.
- [89] Mortensen, E.N., Deng, H., Shapiro, L. (2005). A SIFT descriptor with global context. In *Computer Vision and Pattern Recognition (CVPR 2005)*, 20-25 June 2005. IEEE, Vol. 1, 184-190, 2005.
- [90] Muhammad Imran Razzak, Rubiyah Yusof and Marzuki Khalid. Multimodal face and finger veins biometric authentication. *Scientific Research and Essays*, Vol.5(17), pp. 2529-2534, 2010.
- [91] Navin Prakash, Dr.Yashpal Singh. Support Vector Machines for Face Recognition. *International Research Journal of Engineering and Technology (IRJET)*. Volume: 02 Issue: 08, pp. 1517-1529. Nov-2015. FR_SVM
- [92] NEFIAN, Ara V. et HAYES, Monson H. Hidden Markov models for face recognition. In : *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*. IEEE, 1998. p. 2721-2724.

- [93] Noushath, S. Imran, Mohammad Jetly, Karan Rao, Ashok Hemantha Kumar, G. Multimodal biometric fusion of face and palmprint at various levels. Proceedings of the 2013 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2013, pp. 1793-1798. 2013.
- [94] Ojala, Timo, Pietikäinen, Matti, et Harwood, David. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, vol. 29, no 1, p. 51-59, 1996.
- [95] Orssén, Per-Erik. Maximally stable colour regions for recognition and matching. In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, pp. 1-8, 2007.
- [96] P.Latha, Dr.L.Ganesan & Dr.S.Annadurai. Face Recognition using Neural Networks. *Signal Processing: An International Journal (SPIJ) Volume (3) : Issue (5)*. pp. 153-160. 2009.
- [97] Palanivel, S. Jain, Anil K., Hong, Lin, et Kulkarni, Yatin. A multimodal biometric system using fingerprint, face and speech. In : *Proceedings of 2nd Int'l Conference on Audio-and Video-based Biometric Person Authentication, Washington DC*. pp. 182-187, 1999.
- [98] Pierre Buysens and Marinette Revenu. IR and visible face identification via sparse representation, In *Biometrics Theory Applications and Systems (BTAS)*, pp. 1-6, 2010.
- [99] Pohl, Cle et Van Genderen, John L. Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International journal of remote sensing*, 1998, vol. 19, no 5, p. 823-854. [Pohl et Van Genderen, 1998]
- [100] PolyU-HSFD database. http://www4.comp.polyu.edu.hk/*biometrics
- [101] R. Galloway, Mary M. Texture analysis using grey level run lengths. NASA STI/Recon Technical Report N, 1974, vol. 75, 1974.
- [102] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification”, *IEEE Trans. System Man. Cybernetics*, vol. SMC-3, pp. 610–621, 1973.
- [103] Ralph Gross. Face Databases. The Robotics Institute, Carnegie Mellon University. In S.Li and A.Jain, (ed). *Handbook of Face Recognition*. Springer-Verlag, 2005.

- [104] Ron Kimmel, Cuiping Zhang, Alexander M. Bronstein, and Michael M. Bronstein. Are MSER features really interesting?, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.33, Issue 11, pp 2316-2320, 2011.
- [105] Ross, Arun A., Jain, Anil K., et Nandakumar, Karthik. Information fusion in biometrics. *Handbook of Multibiometrics*, p. 37-58, 2006.
- [106] ROSS, Arun et JAIN, Anil K. Multimodal biometrics: An overview. In : *Signal Processing Conference, 2004 12th European*. IEEE, p. 1221-1224, 2004.
- [107] Rosten, E., Porter, R. and Drummond, T., Faster and better: A machine learning approach to corner detection. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on* 32(1), pp. 105–119, 2010.
- [108] Rowe RK, Uludag U, Demirkus M, Parthasaradhi S, Jain AK (2007) A multispectral whole-hand biometric authentication system. In: *Biometrics Symposium*, pp 1–6
- [109] S. Leutenegger, M. Chli, and R. Siegwart. Brisk: Binary robust invariant scalable keypoints, 2011.
- [110] S. Prabhakar, S. Pankanti, A. K. Jain, “Biometric Recognition: Security and Privacy Concerns”, *IEEE Security & Privacy Magazine*, vol. 1, no. 2, pp. 33-42, April 2003.
- [111] Sadhna Sharma. Template Matching Approach for Face Recognition System. *International Journal of Signal Processing Systems* Vol. 1, No. 2 December 2013
- [112] Santini S. et Jain R. (1999), Similarity Measures, *IEEE Trans. on PAMI*, Vol.12, No. 9, pp. 871 – 883, September 1999.
- [113] Seong G. Kong, Jingu Heo, Besma R. Abidi, Joonki Paik, Mongi A. Abidi, Recent advances in visual and infrared face recognition :a review, *Computer Vision and Image Understanding*, elsevier, Volume 97, Issue 1, 2005, Pages 103-135, January 2005
- [114] Shahin, M. K. Badawi, A. M. Rasmy, M. E. .A multimodal hand vein, hand geometry, and fingerprint prototype design for high security biometrics. 2008 Cairo International Biomedical Engineering Conference, CIBEC 2008, 2008.

- [115] Sinjini Mitra. Gaussian Mixture Models for Human Face Recognition under Illumination Variations. *Applied Mathematics*, pp. 2071-2079, 2012. (<http://www.SciRP.org/journal/am>) [dernière visite le 10/01/2018]
- [116] SNEATH, Peter HA, SOKAL, Robert R., et al. Numerical taxonomy. The principles and practice of numerical classification. 1973.
- [117] Socolinsky, Diego A., et al. "Illumination invariant face recognition using thermal infrared imagery." *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. IEEE, 2001.
- [118] Souhila Guerfi Ababsa. Thèse : Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D. Université d'Evry Val d'Essonne. Pp 37 – 54. 2008.
- [119] Stanford database. <http://scien.stanford.edu/index.php/hyperspectral-image-data/>
- [120] Sung-Hyuk Cha, Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions, Issue 4, Volume 1, pp 300-307, 2007.
- [121] Swathy Nair, Bindu Elias and VPS Naidu. Pixel Level Image Fusion: A Neuro-Fuzzy Approach, *International Journal of Computer Science and Business Informatics*, Vol. 12, No. 1, pp 71-86, 2014.
- [122] T.Tuytelaars and L. Van Gool, IBR & EBR: Matching widely separated views based on affine invariant regions In *IJCV* 59(1):61-85, 2004.
- [123] V.P.S. Naidu. Image Fusion Technique using Multi-resolution Singular Value Decomposition. *Defense Science Journal*, Vol. 61, No. 5, pp. 479-484, September 2011.
- [124] V.P.S. Naidu and J.R. Raol. Pixel-level Image Fusion using Wavelets and Principal Component Analysis. *Defense Science Journal*, Vol. 58, No. 3, pp 338-352, 2008.
- [125] Wang, Shangfei, LIU, Zhilei, LV, Siliang, *et al.* A natural visible and infrared facial expression database for expression recognition and emotion inference. *IEEE Transactions on Multimedia*, vol. 12, no 7, pp. 682-691, 2010.
- [126] Wark, Tim et Sridharan, Sridha. Adaptive fusion of speech and lip information for robust speaker identification. *Digital Signal Processing*, vol. 11, no 3, pp. 169-186, 2001.

- [127] Wilder, Joseph, Phillips, P. Jonathon, Jiang, Cunhong, et al. Comparison of visible and infra-red imagery for face recognition. In: Automatic Face and Gesture Recognition, 1996. Proceedings of the Second International Conference on. IEEE, 1996. p. 182-187, 1996.
- [128] Wu, Jian, Cui, Zhiming, Sheng, Victor S., *et al.* A Comparative Study of SIFT and its Variants. *Measurement Science Review*, vol. 13, no 3, pp. 122-131, 2013.
- [129] Yang, C., Zhang, J., Wang, X., Liu, X.: A Novel Similarity Based Quality Metric for Image Fusion. *Inf. Fusion* 9(2), 156–160 (2008).
- [130] You, C., Liu, Y., Zhao, B., & Yang, S. An Objective Quality Metric for Image Fusion based on Mutual Information and Multi-scale Structural Similarity. *JSW*, 9(4), pp. 1050-1054, 2014.
- [131] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity”, *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [132] Zhang, Baochang, ZHANG, Lei, ZHANG, David, *et al.* Directional binary code with application to PolyU near-infrared face database. *Pattern Recognition Letters*, vol. 31, no 14, p. 2337-2344, 2010.
- [133] Zhao, Wenyi, Chellappa, Rama, Phillips, P. Jonathon, *et al.* Face recognition: A literature survey. *ACM computing surveys (CSUR)*, vol. 35, no 4, p. 399-458, 2003.

Table des matières

Dédicace -----	V
Remerciements -----	VI
Résumé -----	VII
Abstract -----	VIII
Liste des tableaux -----	IX
Liste des figures -----	X
Liste des abréviations -----	XII
Introduction générale -----	1
1- Contexte et motivation-----	1
2- Objectif-----	1
3- Contribution-----	2
4- Liste des publications-----	3
5- Organisation de la thèse-----	3
Chapitre 1. Etat de l'art de la reconnaissance de visage -----	5
Introduction-----	5
1.1. Fondamentaux de la biométrie-----	5
1.1.1. Architecture générale-----	6
1.1.2. Biométries monomodales-----	7
1.2.3. Enjeu de la biométrie-----	8
1.2.3.1. Avantages des systèmes mono-biométriques-----	8
1.2.3.2. Limites des systèmes mono-biométriques-----	8
1.2.3.3. Différentes modalités biométriques-----	9
1.2. Reconnaissance de visage-----	10
1.2.1. Motivation de la biométrie de visage-----	10
1.2.2. Méthodes locales-----	11
1.2.3. Méthodes globales-----	13
1.2.4. Méthodes hybrides-----	13
1.3. Reconnaissance de visage en infrarouge-----	13
1.4. Reconnaissance de visage multispectrale-----	16
1.4.1. Notion d'images multispectrales-----	16
1.4.2. Notion de biométrie multispectrale-----	18
1.4.3. Applications de la biométrie multispectrale-----	19
1.4.4. Reconnaissance de visage multispectrale-----	19
1.4.4.1. Objectif-----	19
1.5.4.2. Principes-----	20
1.5.4.3. Bases de données multispectrales-----	21

1.5. Reconnaissance de visage et biométrie multimodale	24
1.6. Reconnaissance de visage 3D	27
1.7. Mesure de performance d'un système biométrique	28
Conclusion	30
Chapitre 2. Reconnaissance de visage multispectrale basée sur une extraction hybride de caractéristiques (MS-FRHF)	31
Introduction	31
2.1. Notion d'extraction de caractéristiques	31
2.1.1. Points d'intérêt	32
2.1.2. Texture	32
2.2. Méthodes d'extraction de caractéristiques	33
2.2.1. Extraction de points d'intérêt	33
2.2.2. Méthodes d'extraction de points d'intérêt	34
2.3.2.1. Détecteurs	34
2.3.2.2. Descripteurs	38
2.2.3. Extraction de texture	40
2.3.3.1. Extraction de texture basée sur les statistiques du premier ordre	40
2.3.3.2. Extraction de texture basée sur les statistiques du second ordre	41
2.3.3.3. Statistiques d'ordre supérieur	42
2.3.3.4. Filtres de Gabor	43
2.3.3.5. Local Binary Patern (LBP)	43
2.3.2.6. Méthode fractale	44
2.3. Méthode MS-FRHF proposée pour l'extraction de caractéristiques en reconnaissance de visage multispectrale	44
2.3.1. Objectifs	44
2.3.2. Reconnaissance de visage dans le visible et l'infrarouge thermique	45
2.3.3. Fusion d'images	46
2.3.4. Fusion d'images visible/infrarouge thermique	47
2.3.4.1. Méthodologie	47
2.3.4.2. Processus de fusion d'images visible/infrarouge	47
2.3.5. Description de la base de données multispectrale	51
2.3.6. Extraction de caractéristique par la méthode MS-FRHF	52
2.3.7. Résultats expérimentaux	54
Conclusion	56
Chapitre 3. Reconnaissance de visage multispectrale à l'aide de caractéristiques binaires	57
Introduction	57
3.1. Motivation	57
3.2. Présentation de détecteurs et descripteurs binaires	58

3.2.1. Méthodes basées sur les caractéristiques locales	58
3.2.2. Fonctionnement des détecteurs et descripteurs binaires	58
3.2.3. BRISK	60
3.2.4. BRIEF	62
3.2.5. ORB	63
3.2.6. FREAK	64
3.3. Méthode proposée pour la reconnaissance de visage multispectrale à l'aide de détecteurs et de descripteurs binaires	65
3.3.1. Présentation des méthodes d'extraction de caractéristiques utilisées dans la méthode proposée	67
3.3.1.1. Présentation de BRISK	68
3.3.1.2. Présentation de FAST	69
3.3.1.3. Présentation de FREAK	71
3.3.1.4. Présentation du détecteur Harris	73
3.3.1.5. Présentation de SURF	73
3.3.2. Matériels et méthodes	75
3.3.2.1. Base de données d'images	75
3.3.2.2. Fusion d'images	75
3.3.2.3. Extraction de caractéristiques	76
3.3.3. Résultats expérimentaux	78
Conclusion	81
Chapitre 4. Reconnaissance de visage multispectrale à l'aide de score de similarité	82
Introduction	82
4.1. Approches de comparaison d'objets	83
4.2. Motivation de l'utilisation des scores de similarité	83
4.3. Mesures de similarité appliquées aux images	85
4.3.1. Définitions	85
4.3.2. Applications des mesures de similarité	86
4.4. Approche proposée pour la reconnaissance de visage multispectrale à l'aide de score de similarité	87
4.4.1. Introduction à la notion de similarité	88
4.4.2. Approche théorique	89
4.4.2.1. Notion de distance, de similarité et de dissimilarité	89
4.4.2.2. Famille des distances L_p Minkowski	90
4.4.2.3. Famille des distances L_1	91
4.4.2.4. Famille des distances d'intersections	91
4.4.2.5. Famille des distances de produit scalaire	91
4.4.2.6. Famille des distances de la corde au carré	92

4.4.2.7. Famille des distances L_2 carré	92
4.4.2.8. Famille des distances d'entropie de Shannon	93
4.4.3. Méthode proposée pour l'étude de similarité	94
4.4.3.1. Corrélation et fusion d'images	94
4.4.3.2. Extraction de caractéristiques	96
4.4.4. Résultats expérimentaux	99
4.4.4.1. Base de données	99
4.4.4.2. Description de l'expérience	99
4.4.4.3. Résultats	100
Conclusion	102
Conclusion générale et perspectives	103
Perspectives	104
Bibliographie	105
Table des matières	117

Titre : Extraction et fusion de points d'intérêt et texture spectraux pour l'identification, le contrôle et la sécurité

Mots clés : reconnaissance de visage, multispectral, visible, infrarouge, point d'intérêt, texture

Résumé : La biométrie est une technologie émergente qui propose de nouvelles méthodes de contrôle, d'identification et de sécurité. Les systèmes biométriques sont souvent victimes de menaces. La reconnaissance faciale est populaire et plusieurs approches existantes utilisent des images dans le spectre visible. Ces systèmes traditionnels opérant dans le spectre visible souffrent de plusieurs limitations dues aux changements d'éclairage, de poses et d'expressions faciales. La méthodologie présentée dans cette thèse est basée sur de la reconnaissance faciale multispectrale utilisant l'imagerie infrarouge et visible, pour améliorer la performance de la reconnaissance faciale et pallier les insuffisances du spectre visible.

Les images multispectrales utilisées cette étude sont obtenues par fusion d'images visibles et infrarouges. Les différentes techniques de reconnaissance sont basées sur l'extraction de caractéristiques telles que la texture et les points d'intérêt par les techniques suivantes : une extraction hybride de caractéristiques, une extraction binaire de caractéristiques, une mesure de similarité tenant compte des caractéristiques extraites.

Title : Extraction and fusion of spectral interest points and texture for identification, control and security

Keywords : face recognition, multispectral, visible, infrared, key-point, texture

Abstract : Biometrics is an emerging technology that proposes new methods of control, identification and security. Biometric systems are often subject to risks. Face recognition is popular and several existing approaches use images in the visible spectrum. These traditional systems operating in the visible spectrum suffer from several limitations due to changes in lighting, poses and facial expressions. The methodology presented in this thesis is based on multispectral facial recognition using infrared and visible imaging, to improve the performance of facial recognition and to overcome the deficiencies of the visible spectrum.

The multispectral images used in this study are obtained by fusion of visible and infrared images. The different recognition techniques are based on features extraction such as texture and points of interest by the following techniques: a hybrid feature extraction, a binary feature extraction, a similarity measure taking into account the extracted features.