

Liste des travaux scientifiques

- El Mahdi Barraah, S. S., Malaoui, A. (2015). Extended Set of *DCT-TPLBP* and *DCT-FPLBP* for Face Recognition. World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9(8).
- Barraah, E. M., Ahdid, R., Safi, S., Malaoui, A. (2015). New Technique to Determination of Electronic Circuits for Semiconductor Components by Recognizing Nyquist Curve. Journal of Computer Sciences and Applications, 3(4), 100-104.
- Barraah, E. M., Ahdid, R., Safi, S., Malaoui, A. (2017). 3DLBP and SVD Fusion for 3D Face Recognition Using Range Image. Journal of Computer Science and Information Technology, 5(2), 61-65.
- Barraah, E. M., Safi, S., Malaoui, A. (2016, May). New technique for face recognition based on Singular Value Decomposition (SVD). In Cloud Computing Technologies and Applications (CloudTech), 2016 2nd International Conference on (pp. 96-99). IEEE.
- EL Mahdi BARRAH; Said SAFI; Abdessamad MALAOUI: Comparing of two 2D face Recognition Methods: Facial surface Analysis by Contours and new Descriptor Zernik Moment. First International Conference on Business Intelligence (CBI'14)Maroc, Beni Mellal, 29 - 30 April, 2014.
- EL Mahdi BARRAH; Rachid AHDID Said SAFI; Abdessamad MALAOUI: Comparative study of 2D face recognition algorithms using texture analysis using descriptive statistical. The second International Conference on Business Intelligence (CBI'14)Maroc, Beni Mellal, 23 - 25 April, 2015.
- EL Mahdi BARRAH; Said SAFI; Abdessamad MALAOUI: « Fusion of Singular Value Decomposition (SVD) and DCT-PCA for Face Recognition », International Arab Conference on Information Technology (ACIT'2016) Beni Mellal, Morocco, December 18-20 2016.
- R. AHDID; S.SAFI; S.ELKASSIMI; E.BARRAH: « Face Surface Reconstruction by Contours », The 4th International Workshop on codes, Cryptography and Communication Systems (WCCCS'13) EST-Meknes, Morocco, November 07-08 2013.

Résumé

Ce travail s'inscrit dans la thématique de la reconnaissance des visages. Il s'agit de décider de manière automatique de l'identité d'une personne en fonction de texture de son visage. Notre but étant de développer un système d'authentification de visage simple et efficace, l'extraction des caractéristiques à partir de l'image est une étape primordiale dans le processus de la reconnaissance faciale et de surmonter les faiblesses des systèmes classiques de reconnaissance 2D et 3D et d'améliorer le taux de reconnaissance obtenu par les différentes techniques utilisées dans ce domaine. Dans la première partie de notre thèse, nous proposons quatre systèmes de reconnaissance de visage 2D en utilisons deux méthodes d'extraction des caractéristiques basées sur le Motif binaire local (Local Binary Pattern LBP) à savoir : histogrammes des motifs binaires locaux (LBPH) + l'Analyse en Composante Principale(ACP), histogramme de motif binaires locaux multiéchelles(MLBPH) + l'Analyse en Composante Principale (ACP).

Les quatre algorithmes de reconnaissance de visage proposé est basées sur la représentation des régions faciales locales en utilisant la transformée en cosinus discrète (Discrete Cosine Transform DCT), l'analyse en composante principale (ACP). La représentation locale offre une robustesse face aux variations d'apparence dans les régions locales provoquées par l'occlusion partielle du visage ou l'expression du visage, alors que l'utilisation de l'information de fréquence fournit une robustesse à l'égard des changements d'illumination.

Nous présentons aussi un système de reconnaissance automatique de composants électroniques. Ce système est fondé sur l'image numérique des courbes de Nyquist tirées par des fonctions de transfert de leurs impédances équivalentes.

Dans l'étape de classification, nous avons utilisés les Réseaux de Neurones (RN), les K-Plus Proche Voisins (KPPV) et les Vecteurs à Support Machines (SVM). Pour tester nos méthodes et évaluer leurs performances, nous avons réalisé une série de simulations d'expériences sur les trois bases de données 2D bien connus des images de visages FERET, XM2VTS et ORL. Nous avons qualitativement et quantitativement comparé notre approche avec les méthodes classiques.

Abstract

This work is part of the theme of face recognition. It is a question of automatically deciding the identity of a person according to the texture of his face. Our goal is to develop a simple and effective face authentication system, extracting features from the image is a crucial step in the process of facial recognition. Our goal is to overcome the weaknesses of conventional 2D and 3D recognition systems and improve the recognition rate obtained by the different techniques used in this field. In the first part of our thesis, we proposed four 2D face recognition systems using two methods of extraction of characteristics based on the local binary pattern (LBP) namely: histograms of local binary patterns (LBPH) + Principal Component Analysis (PCA), histogram local multiscale binary pattern (MLBPH) + Principal Component Analysis (PCA).

The proposed face recognition algorithm is based on the representation of local facial regions using the discrete cosine transform (DCT), principal component analysis (PCA). Local representation provides robustness to variations in appearance in local areas caused by partial occlusion of the face or facial expression, while the use of frequency information provides robustness to illumination changes.

We are also introducing an automatic recognition system for electronic components. This system is based on the digital image of Nyquist curves drawn by transfer functions of their equivalent impedances.

In the classification stage we used Neural Networks (RNs), Near-Neighbor K-Plus (KPPV) and Machine Supported Vectors (SVM). To test our methods and evaluate their performance, we performed a series of simulations of experiments on the three well-known 2D databases of FERET, XM2VTS and ENT face images. We qualitatively and quantitatively compared our approach with classical methods. We have shown that the use of local binary pattern, for 2D and 3D face recognition, improves the efficiency of the results, compared to the results of standard techniques.

Dédicaces

A mon père, à ma mère ‘

A mes chères frères et soeurs

À la famille BARRAH ‘

A tout mes amis(es)

Remerciements

Je tiens tout particulièrement à remercier mes directeurs de thèse **Abdessamad MALAOUI** Professeur à la Faculté Polydisciplinaire de Béni Mellal et **Said SAFI** Professeur à la Faculté Polydisciplinaire de Béni Mellal, de m'avoir dirigé durant ces années de thèse, d'avoir accompli leur part de travail avec compétence et de m'avoir laissé réaliser celle qui m'incombait sans aucune contrainte. Travailler avec vous a été très enrichissant, merci de m'avoir encouragé tout au long de mes années de thèse.

Je tiens à adresser mes plus sincères remerciements à **Bouzi MANAUTE** Professeur à la Faculté Polydisciplinaire de Béni Mellal, **xxxxxxxxxxxxx** Professeur à la Faculté Polydisciplinaire de Béni Mellal et **xxxxxxxxxxxxxxxxxxxxx** Professeur à la Faculté des Sciences Juridiques, Economiques et Sociales Souissi-Rabat pour avoir accepté de juger mon manuscrit et de participer au jury de ma soutenance, ainsi qu'à **xxxxxxxxxxxxxxxxxxxxx** Professeur à la Faculté des Sciences et Techniques de Béni Mellal qui m'a fait un grand honneur de bien vouloir présider ce jury.

J'adresse mes vifs remerciements à Monsieur **Rachid LBIBB**, Directeur du Laboratoire de Développement Durable (L2D), Faculté des Sciences et Techniques de Sultane Moulay Slimane, pour m'avoir accueilli au sein du laboratoire et pour son aide précieuse. J'aimerais remercier tous les collègues du Laboratoire, ainsi que les professeurs du département d'informatique de la Faculté des Sciences et Techniques.

Merci à toutes celles et à tous ceux qui ont participé, par leurs relectures et leurs conseils, à la rédaction de ce manuscrit. Un merci spécial à **Yacine BARRAH**, **Rachid AHDID**, **Rachid AKILI** et **Abdessamad Akili** pour être mes vrais amis, votre soutien m'est très utile tout au long des quatre dernières années.

Enfin, je tiens à remercier mes parents **Lhoucein BARRAH** et **Zohra OUMLOUSSE**, mon frère **Yacine BARRAH** et soeurs **Mouna BARRAH** et **Houria BARRAH** pour leurs soutien inconditionnel, sans qui ce travail n'aurait jamais pu voir le jour.

Table des matières

Liste des travaux scientifiques	i
Conférence internationales	i
Résumé	ii
Abstract	iii
Dédicaces	iv
Remerciements	v
Abréviations	1
Introduction générale	2
0.1 Contexte et motivation	2
0.2 Problématique et Contributions	2
0.3 Organisation de la thèse	4
1 Etat de l'art de la reconnaissance du visage	6
1.1 Introduction	6
1.2 Méthodes globales	7
1.3 Approches locales	8
1.3.1 Méthode basée sur le modèle binaire local (LBP)	8
1.3.2 Méthode basée sur Local Phase Quantification (LPQ)	12
1.3.3 Méthodes basées sur les ondelettes de Gabor	14
1.4 Autres méthodes	18
1.5 Conclusion	19
2 Bases de données et méthodes d'apprentissage	20
2.1 Bases de données FERET	20
2.2 Base de données FRGC 2.0 avec EEB (Environnement d'Expérimentation Biométrique)	21
2.3 Base de données XM2VTS avec protocole Lausanne	22

2.4	Mesures de performance	24
2.4.1	Mesures de performance dans l'identification du visage	24
2.4.2	Mesures de performance dans la vérification des visages	24
2.5	Base de données Tifinagh	25
2.5.1	Création de la base de données	26
2.6	Base de données des composants électroniques	27
2.6.1	Circuit électrique	27
2.6.2	Diagramme de Nyquist	27
2.7	Méthodes de réduction de dimension	29
2.7.1	L'analyse en composantes principales	30
2.7.2	Transformée en cosinus discret	32
2.7.3	Décomposition de la valeur singulière (SVD)	34
2.8	Apprentissage et classification	35
2.8.1	Les réseaux de neurones artificiels	35
2.8.2	Support vector regression	37
2.9	Conclusion	38
3	Reconnaissance de visage basée sur l'apparence locale	40
3.1	Méthodes de représentation de l'apparence locale	41
3.1.1	La transformée cosinus discrète	41
3.1.2	Analyse en composantes principales (ACP)	42
3.1.3	Reconnaissance faciale basée sur l'apparence locale utilisant la transformée cosinus discrète	45
3.2	Classification avec KPPV	46
3.3	Expériences	46
3.3.1	Bases de données de référence	46
3.4	Conclusion	47
4	Résultats et simulations	48
4.1	Introduction	48
4.2	Authotification de visage 2D	48
4.2.1	Problème d'identification de visage: Base de données FERET	48
4.2.2	Simulation sur la base de données XM2VTS	51
4.2.3	Simulation sur la base de données FRGC 2.0	53
4.3	Conclusion	54
4.4	Reconnaissance faciale avec ELBP	55
4.4.1	Résultats sur la base de données AR	55
4.4.2	Résultats obtenus sur la base de données FERET	59
4.4.3	Résultats sur la base de données SCface	60
4.4.4	Conclusion	62
4.5	LPQ comme descripteur d'extraction de caractéristiques du visage	63
4.5.1	Résultats sur la base de données FERET	64
4.5.2	Résultats sur la base de données FRGC 2.0	66

4.5.3	Conclusion	68
4.6	Vérification de visage 3D	69
4.6.1	Aperçu sur l'approche proposée	69
4.6.2	Les descripteurs locaux	70
4.6.3	LBP basée sur les patches (Patch base LBPs)	72
4.6.4	Méthode de 3DLBP	75
4.6.5	Extraction des caractéristiques d'histogrammes	77
4.7	Résultats de simulation et discussion	77
4.8	Conclusion	85
	Conclusion générale	86

Abréviations

Notation

ACP : Analyse en Composante Principale

ADL : Analyse Discriminante Linéaire

LTP : Motif Ternaire local

RN : Réseaux de Neurones

KPPV : K-Plus Proche Voisin

SVM : Vecteurs à Support Machines

LBP : modèle binaire local

ELBP: Elliptical Local Binary Pattern

LPQ : Local Phase Quantification

PLS : Partial Least Squares

EPOEM : elliptiques des grandeurs de bord orientées

STFT : la transformée de Fourier de courte durée

FR : Reconnaissance Facial

LGBPHS : Histogramme de modèle binaire Gabor local séquence

Introduction générale

0.1 Contexte et motivation

De nos jours, la reconnaissance faciale (FR) est un secteur de recherche attrayant et plein de défis. Elle peut être utilisée pour des applications de sécurité (contrôle d'accès aux zones autorisées, ordinateurs, aéroports, etc.), des dispositifs de surveillance dans les espaces publics (stades, gares, centres commerciaux). Il existe traditionnellement deux manières d'identifier un individu. La première méthode est basée sur une connaissance à priori "knowledge-based" de la personne telle que, par exemple, la connaissance de son code PIN qui permet d'activer un téléphone portable. La seconde méthode est basée sur la possession d'un objet "token-based". Il s'agit d'une pièce d'identité, d'une clef, d'un badge, etc. Ces deux modes d'identification peuvent être utilisés de manière complémentaire afin d'obtenir une sécurité accrue comme dans le cas de la carte bleue. Cependant, elles ont chacune leurs faiblesses. Les caractéristiques biométriques sont une solution alternative aux anciens moyens de vérification d'identité. L'avantage de ces caractéristiques biométriques est d'être universelle. C'est-à-dire elles sont présentes chez toutes les personnes à identifier. D'autre part elles sont mesurables et uniques : deux personnes ne peuvent pas posséder exactement les mêmes caractéristiques. Elles sont aussi permanentes ce qui signifie qu'elles ne varient pas ou peu au cours du temps.

0.2 Problématique et Contributions

Malgré toutes les avancées dans le domaine de la reconnaissance de visages, aucun système de reconnaissance fiable n'a encore pu voir le jour, Bien que certains systèmes proposés aient montré une grande efficacité, ils sont souvent limités par les changements de l'environnement, comme le changement de l'éclairage, ainsi que par les variations propres aux visages, comme les expressions faciales. Cette variabilité a rendu l'automatisation de la reconnaissance faciale une tâche de grande complexité.

La complexité de la reconnaissance automatique des visages est due à deux difficultés majeures. La première difficulté est la ressemblance structurelle entre les visages. Deux visages différents sont très proches en termes de structure, car composées des même caractéristiques (yeux, bouche, nez, etc) dont la localisation et la forme varient légèrement.

La deuxième difficulté provient des changements d'apparence d'un même visage acquis dans des conditions d'acquisition différentes. Ces différences sont dues, généralement, à des facteurs environnementaux comme : les conditions d'éclairage, les caractéristiques des capteurs et aussi leur positionnement par rapport aux visages lors de l'acquisition, les modifications propres du visage telles que les expressions, les variations de poids et de l'âge.

Plusieurs méthodes ont été développées pour la reconnaissance de visage 2D. Cependant, elles présentent un certain nombre de limitations liées à l'orientation du visage ou à la pose, à l'éclairage, à l'expression faciale, aux occultations, etc. Des techniques de reconnaissance 3D ont été proposées comme une solution alternative pour résoudre les problèmes cités ci-dessus.

L'extraction des caractéristiques du visage à partir de l'image est une étape primordiale dans le processus de reconnaissance faciale. L'algorithme d'extraction doit être efficace et doit fournir, sans aucune ambiguïté, les caractéristiques du visage dans l'image. C'est dans ce contexte qu'intervient cette thèse. Nous avons amélioré de nouveaux algorithmes alliés à des stratégies originales pour surmonter les résultats de la reconnaissance du visage.

La première contribution de cette thèse porte sur la reconnaissance faciale 2D. Son objectif est de surmonter les faiblesses des systèmes de reconnaissance 2D sensibles aux changements d'éclairage, d'orientation, d'expressions faciales. En raison de sa simplicité et de son efficacité pour la reconnaissance 2D du visage, le descripteur LBP 3D (modèle Binaire locaux) a été utilisé dans de nombreux travaux récents de reconnaissance 3D. Dans ce travail de thèse, nous nous utilisons ce descripteur utilisé pour des images de profondeur pour la représentation de la forme 3D du visage. Nous présentons ainsi dans cette section son principe et les méthodes de reconnaissance 3D de visages basées sur ce descripteur.

Dans la première partie contient quatre méthodes d'extraction de caractéristiques (MLBPH+ACP, LBPH+ACP, ELBP(h,v) et LBP). Dans l'étape d'extraction des caractéristiques, l'image de visage est divisée en plusieurs régions à partir desquelles les distributions d'entités LBP sont extraites et concaténées dans un vecteur de caractéristiques amélioré à utiliser comme descripteur de visage.

Le score global de similarité est obtenu en fusionnant les scores de similarité des descripteurs régionaux. Cette méthode a été mise en œuvre et comparée à d'autres repères bien connus dans la vérification des visages. Les résultats obtenus sur la base de données XM2VTS montrent clairement que (MSLBPH + LDA + Réseau de neurone) surpasse les autres prétendants à l'état de l'art. Les stratégies pour résoudre le problème de l'illumination peuvent être résumées comme suit: convertir l'image de visage en une forme plus canonique dans laquelle les variations d'éclairage sont supprimées, la normalisation photométrique et le descripteur de texture robuste (tels que les filtres Gabor et LBP)

peut aider à cet égard.

La normalisation est utilisée pour réduire les effets de l'illumination avant d'appliquer la méthode de reconnaissance de visage LBPH. Les techniques ont été testées sur les bases de données FERET, XM2VTS et FRGC 2.0.

La seconde contribution de cette thèse consiste à mettre un système de reconnaissance automatique des composantes électroniques. Le système de reconnaissance est fondé sur l'image numérique des courbes de Nyquist tirées par des fonctions de transfert de leurs impédances équivalentes.

la dernière contribution que nous avons développé et la mise en place d'un système de reconnaissance de caractères Tifinagh avec des données manquantes. Nous avons proposé une approche basée sur l'utilisation de la corrélation entre le bloc perdu et son voisin. Ce système comprend trois parties principales: prétraitement, extraction et reconnaissance des caractéristiques.

multi-échelle proposé avec la méthode de normalisation offre une solution simple et robuste au problème d'illumination et à la présence d'erreurs de localisation pour la reconnaissance faciale. Les résultats montrent également que le système proposé peut atteindre des performances comparables aux systèmes de pointe.

0.3 Organisation de la thèse

Cette thèse est composée de deux grandes parties. La première partie présente l'état de l'art ainsi que les concepts et notations requis pour la compréhension des systèmes développés. Cette partie est divisée en deux chapitres. La deuxième partie de ce rapport est constituée de deux chapitres qui présentent l'ensemble de nos contributions. Les chapitres de ce manuscrit sont répartis comme suivant:

Le premier chapitre présente l'état de l'art sur la reconnaissance automatique de visage. En effet, ce chapitre décrit le principe, l'architecture d'un système de reconnaissance de visage et les différentes techniques utilisées dans ce domaine.

Le second chapitre: présente les bases de données d'images de visage que nous avons utilisées dans le cadre de cette thèse. En effet, il décrit le processus de création de cette base, la collecte des données, la numérisation, ainsi que les opérations de prétraitement effectuées sur les images des visages. il présente ensuite le mode de stockage et d'étiquetage des images de la base, en terminant par présenter les différentes techniques de projection d'espace.

Le troisième chapitre dans un premier temps décrit les résultats expérimentaux obtenus suite aux tests effectués sur les bases de données (FERET, XM2VTS et FRGC 2.0). Seconds temps, il décrit nos résultats expérimentaux sur les bases de données du

caractère Tifinaghe. Enfin, il présente une évaluation d'un nouveau système de vérification de visage 3D basé sur les histogrammes de quatre descripteurs locaux multi blocs : LBP[?], TPLBP[?], FPLBP[?], 3DLBP[?] et LPQ[?]. Il donne une analyse détaillée de ces algorithmes et comparé leurs performances pour construire un système automatique efficace de vérification 3D de visage.

Dans le quatrième chapitre, nous nous intéressons à l'exploitation des informations sur des bases de données des circuits électronique pour l'authentification. Dans un premier temps, nous décrivons le dispositif de création que nous avons choisi, Ensuite, nous détaillons les différentes approches de reconnaissance faciale des caractère tfinagh avec des données manquante.

Enfin, la conclusion générale dresse un bilan de ce travail. Elle présente une synthèse des résultats de nos expérimentations et met en évidence les apports de notre travail. Ainsi, elle présente un aperçu sur nos contributions et donnera quelques perspectives sur les travaux futurs.

CHAPITRE

1

Etat de l'art de la reconnaissance du visage

1.1 Introduction

Un système de reconnaissance de visages est composé, généralement, de quatre étapes: l'acquisition, le prétraitement, l'extraction de descripteurs, et enfin la classification pour la prise de décision (identité de la personne). Ce chapitre est destiné à donner un état de l'art de l'extraction des caractéristiques et les informations importantes à partir de l'image du visage afin de fournir une signature biométrique. Nous insisterons surtout sur la place de la reconnaissance faciale parmi les autres techniques biométriques, car elle constitue l'objectif de cette thèse. Enfin, nous exposerons les difficultés majeures liées à la reconnaissance de visage, et qui font toujours l'objet de recherche par la communauté scientifique.

Le but de ce chapitre est de donner un panorama des méthodes les plus utilisées en reconnaissance 2D et 3D de visages. Tout d'abord, une brève présentation des méthodes les plus populaires, utilisées en reconnaissance faciale, est proposée. Deux de ces méthodes les plus connues pour l'extraction des caractéristiques faciales, étape indispensable dans les systèmes de reconnaissance de visages, sont décrites et enfin les bases de données de visages et les protocoles d'évaluation qui ont été utilisés dans ce travail sont présentés.

1.2 Méthodes globales

Le principe des méthodes globales est d'utiliser de visage au complet comme source d'information, et ce sans segmentation de ses parties. En effet, une image en niveau de gris de taille 112 x 92 est représentée par un vecteur de taille 10304, les couleurs ne sont généralement pas utilisées dans ses méthodes ce qui simplifie un grand nombre d'opérations.

Parmi ses méthodes, on cite les visages propres (propreFaces) qui sont une extension de la transformation de Karhunen-Loève (également connue sous le nom d'Analyse en composantes principales - ACP). À soutenu que chaque image de visage d'un ensemble donné peut être représentée sous la forme d'une combinaison linéaire de vecteurs propres orthogonaux de base calculés par ACP sur l'ensemble d'images lui-même. Dans la méthode Eigenfaces [1], les images d'apprentissage sont d'abord réajustées de leurs intensités, sous forme de matrices entières 2-D de même taille (M ligne x N colonne), aux vecteurs colonnes de longueur M x N. Ces vecteurs sont alors normalisés pour avoir une norme unitaire et une moyenne soustraite pour avoir un zéro. À partir des vecteurs normalisés, l'algorithme ACP est utilisé pour trouver le vecteur propre principal, correspondant aux plus grandes valeurs propres, qui sont utilisées comme un ensemble de graines à représenter pour toutes les autres images de visage via une opération de projection. Cette méthode est appelée Eigenfaces, car ces vecteurs propres peuvent être reconstruits et visualisés sous forme d'images faciales. Les visage propre [1] peuvent fonctionner raisonnablement bien avec des images de bonne qualité capturées dans des conditions strictement identiques de lumière, de pose, d'expressions faciales et s'il n'y a ni variation de vieillissement ni occlusion. À l'inverse, sa performance est considérablement dégradée (par exemple, elle ne propose que 4,7 % de taux de reconnaissance moyenne sur la base de données SCface [2]) et ne peut évidemment pas être applicable en réalité.

Lorsque l'ensemble d'apprentissages a plus d'une image par sujet, les Eigenfaces [1] n'utilisent pas ces informations disponibles pour améliorer la précision du système puisque l'ACP est une technique d'apprentissage non supervisée. Motivé par cela, Fisherfaces [1] a été proposé en utilisant l'algorithme d'apprentissage d'Analyse Discriminante Linéaire (ADL) [1] pour maximiser les variations extraclasses entre les images appartenant à différentes personnes tout en minimisant les variations intraclasse entre celles d'une même personne. En raison du fait intraclasse induites par des facteurs stimulants tels que l'illumination, la pose de la tête et les changements d'expression sont presque toujours plus grands que les variations extraclasses qui proviennent des différences d'identités faciales [4]. Ce qui permet de faire des images d'une même personne extrêmement différentes, l'utilisation de Fisherfaces pour réduire cette déficience est précieuse et conduit à une plus grande précision que les Eigenfaces [3][6]. Mais au contraire, Fisherfaces ne peut être appliqué que lorsque l'ensemble d'entraînement a plus d'une image par personne. Bien que cette condition préalable ne soit pas toujours satisfaite, elle peut être considérée comme une faiblesse de la méthode.

Il existe aussi d'autres approches globales qui étendent les Eigenfaces et les Fisherfaces,

telles que la 2D ACP [8] et la 2D LDA [9], mais leurs performances sont bien inférieures à celles des méthodes locales qui seront présentées dans le paragraphe suivant.

1.3 Approches locales

Bien que les approches globales qui utilisent toute la surface du visage comme source d'information, celles basées sur les caractéristiques s'appuient sur des traits faciaux locaux extraits de composants locaux tels que les yeux, le nez, la bouche, etc. Dans cette section, les méthodes d'ondelettes Gabor pures (aucune stratégie de fusion n'est utilisée) sont d'abord étudiées. Ces méthodes exploitent le fait que les coefficients d'ondelettes de Gabor codent à la fois la forme du visage et les caractéristiques d'apparence locale. Contrairement à celles-ci, LBP et LPQ, sont deux descripteurs élémentaires basés sur l'intensité qui suivent et capturent les caractéristiques de micro-apparence des images de visage via leurs propres opérateurs. L'utilisation de plusieurs LBP ou LPQ avec des paramètres de base différents, créent de nouvelles méthodes à plusieurs échelles. Plus efficacement, les stratégies de fusion sont utilisées dans les méthodes multi-résolution / multi-échelle pour combiner des méthodes élémentaires.

1.3.1 Méthode basée sur le modèle binaire local (LBP)

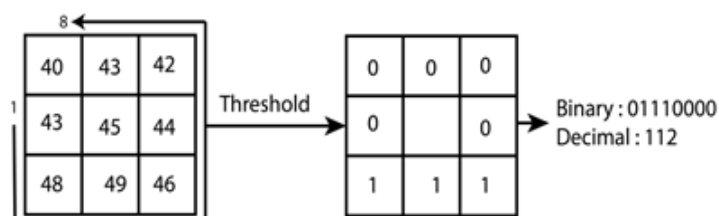


Figure 1.1: Schéma de codage LBP.

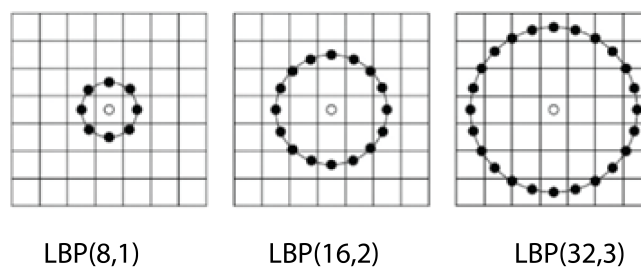


Figure 1.2: Modèles LBP.

Initialement conçu pour le problème de classification des textures en tant que descripteur de texture, les modèles binaires locaux (LBP) [11] est rapidement devenu l'un des caractéristiques les plus populaires de la littérature de reconnaissance de visage. Dans le LBP d'origine [11], chaque pixel d'une image d'entrée est affecté d'un nombre décimal (appelé étiquette LBP) qui est calculé par seuillage binaire son niveau de gris avec ses voisins P situés de façon dispersée sur un cercle de rayon r centré au pixel lui-même. Une interpolation bilinéaire est exploitée pour calculer les valeurs de pixels voisins si elles ne sont pas au centre d'un pixel. Ce schéma de codage est appelé opérateur LBP et noté $LBP(P, r)$ (pour plus de détails, voir les figures 1.1 et 1.2). L'application d'un opérateur LBP sur chaque pixel d'une image de visage produit une image LBP (figure 1.3 montre quelques échantillons) qui contient des informations très importantes pour la reconnaissance de visages.

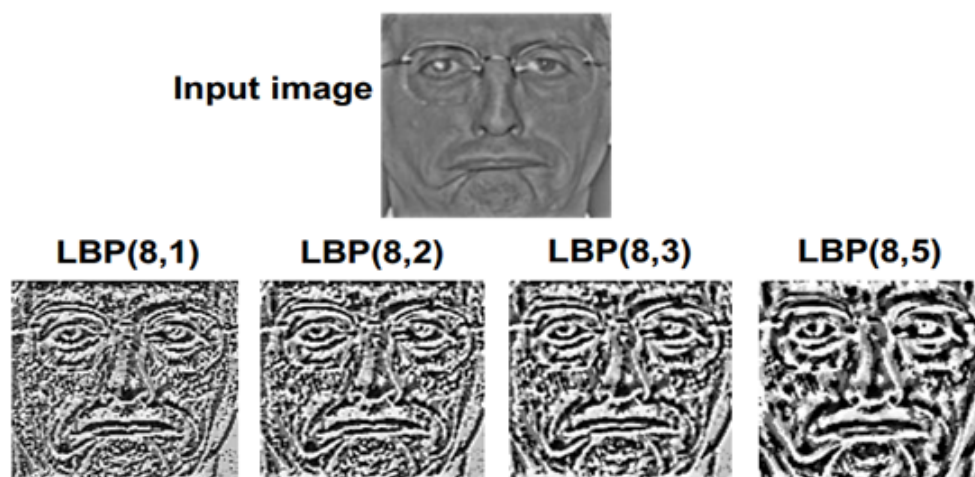


Figure 1.3: Une image de visage et ses LBP.

L'image LBP obtenue par l'opérateur LBP est alors divisée uniformément en sous-régions rectangulaires non chevauchées pour calculer leurs histogrammes. Le vecteur caractéristique LBP de l'image de visage donnée est construit en concaténant ces histogrammes séquences. Le vecteur LBP intègre des informations spatiales utiles (répartition des caractéristiques faciales dans différentes sous-régions).

De manière, un vecteur LBP est une représentation de $2^P - bloc$, mais les études statistiques des étiquettes LBP sur différents types d'images ont révélé que certains bloc sont plus fréquents que les autres. Ces casiers principaux, dont les formes binaires n'ont plus que 2 transitions de bit de 1 à 0 et vice versa, sont appelés modèles uniformes et sont utilisés pour réduire la taille des vecteurs de caractéristiques LBP [11]. Cette opération de compression se traduit par des vecteurs de caractéristiques plus courts et permet la classification plus rapide, mais au contraire elle provoque un léger décretement dans la performance de précision. Toutes ces étapes sont représentées à la Fig. 1.4. Il est intéressant de noter que les descriptions LBP sont généralement formées avec $P = 8$ et par conséquent,

ils sont 256-bins vecteurs caractéristiques. Ahonen et al. [11] utilisait la méthode LBP pour extraire les micro caractéristiques des images faciales, puis ils ont utilisé le modèle de correspondance pour la classification et ils ont obtenu des résultats très prometteurs. D'autres applications de la LBP (liées à la reconnaissance faciale) incluent la détection des visages [12], la reconnaissance de l'expression faciale, l'estimation de l'âge [14], la classification selon le sexe [15], etc. Mais après tout, la méthode LBP a été appliquée avec plus de succès à la reconnaissance faciale. Les mérites de LBP sont le calcul simple, la petite taille du vecteur caractéristique (en comparaison avec les méthodes à base d'ondelettes de Gabor) et la robustesse aux variations d'éclairage [13].

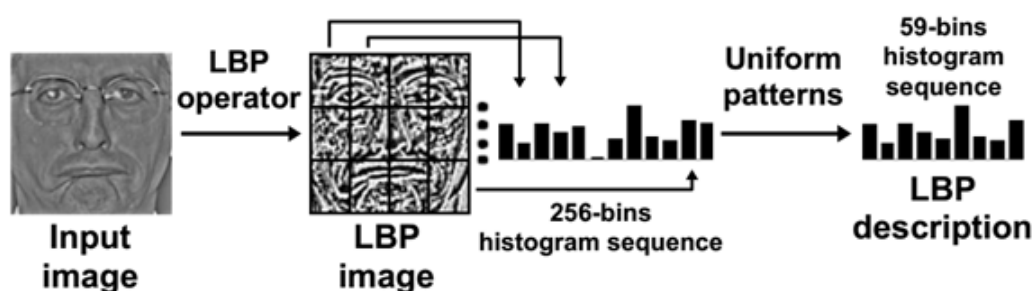


Figure 1.4: Modèle LBP.

Suite à sa première application réussie [16], de nombreuses variantes de LBP ont été proposées ces dernières années pour la reconnaissance faciale. Un LBP stimulante a été introduit par Zhang et al. Dans [17]. Pour une image de visage, on calcule plus de 7000 distances carrées de motifs LBP générées par la sous-fenêtre de décalage et de mise à l'échelle sur l'image donnée.

Heikkilä et al. [24] ont prouvé que la ILBP est plus efficace. À comparer les paires symétriques centrales de pixels pour former le LBP central symétrique (CS-LBP). CS-LBP capture à la fois les caractéristiques micros et les caractéristiques de dégradé des images de visage. La taille du vecteur caractéristique CS-LBP est la moitié de la taille du vecteur caractéristique LBP en utilisant le même motif circulaire. Cette technique a été développée par Choi et al. [22] lorsque des paires de pixels symétriques dans différentes orientations et différents rayons ont été comparées pour constituer la représentation des paires symétriques de pixels circulaires (CCS-POP). De cette façon, CCS-POP capture des informations de bord local en pixels et obtient une précision supérieure à celle de LBP en combinant avec PLS [23] pour la réduction de dimensionnalité.

Basé sur un opérateur de seuil ternaire, Tan et al. [24] ont proposé une nouvelle variante LBP appelée Motif Ternaire local (LTP) en utilisant deux vecteurs LBP pour construire un LTP. LTP a été vérifié pour être plus efficace que LBP contre l'illumination et bruit. L'inconvénient de cette approche est deux fois plus lent en vitesse et plus grande en taille du vecteur caractéristique. De plus, l'idée de la LTP a été étendue par Ren et al. [25]

avec une nouvelle variante appelée Relaxed LTP (RLTP).

Les auteurs ont utilisé quatre marqueurs LBP pour coder chaque pixel d'image et ensuite ils ont été accumulés dans un histogramme LBP d'un vecteur RLTP. RLTP a été utilisée dans le traitement pour apporter l'amélioration à LTP avec bruit de l'image. Le concept d'application de modèles elliptiques multiples dans la LBP a été utilisée par Liao et al. [26] dans leur méthode allongée LBP. En utilisant des facteurs pondérés pour six régions de l'image du visage et quatre motifs elliptiques différents (dans quatre directions), LBP allongée a été argumentée pour coder l'information anisotrope de l'image. Tout en obtenant des résultats plus élevés que la LBP. La LBP allongée présentait un inconvénient, car la taille de son vecteur caractéristique était quatre fois plus longue que celle de la LBP.

Plutôt que de creuser dans les ensembles fixes et prédéfinis de pixels voisins situés sur (Cercles, ellipses, paires symétriques) ou en algorithmes de seuil (binaires, ternaires ou détendus) comme dans les variantes LBP ci-dessus, Maturana et al. [27] ont exploité une technique d'apprentissage supervisé pour rechercher la plupart des quartiers discriminants lors du calcul de l'étiquette LBP pour un pixel d'image. Ceci est basé sur la maximisation d'un critère de séparabilité de type Fisher. Bien que la méthode, nommée comme des modèles binaires locaux discriminants (DLBP), a obtenu des résultats prometteurs, son coût de calcul pour le processus d'apprentissage est une faiblesse réelle. Dans [28], un descripteur appelé modèles de magnitudes de bord orienté (POEM) a été développé en appliquant plusieurs LBP sur des images d'amplitude accumulées. Les auteurs ont ensuite combiné le POEM avec des modèles d'orientations dominantes (PDO) [29] et ils ont obtenu des meilleurs résultats.

À l'origine, LBP est conçu pour le problème de classification des textures et il s'avère qu'il présente des propriétés souhaitables pour être une représentation faciale efficace en reconnaissance faciale. Alors que d'autres variantes de LBP ont essayé d'améliorer sa puissance en utilisant des techniques d'apprentissage (comme dans [30][31]) ou des mécanismes différents dans l'étape de seuillage (par exemple MB-LPB, CS-LBP et LTP), ils tendent à laisser derrière le fait que leur objectif principal est pour le problème de Reconnaissance de Visage (RF).

Pour ce but, tout type d'inspiration d'une variante LBP doit être basé sur des aspects qui conduisent évidemment à des résultats plus précis. Guidé par cette règle, Modèles binaires locaux elliptiques (ELBP), une de nos propositions dans ce travail, est une variante LBP qu'est émanée des observations des images de visage suivantes :

- Les composants faciaux cruciaux, les yeux et la bouche sont naturellement elliptiques. De plus, le visage humain contient des structures plus horizontales, qui jouent un rôle important dans la mémorisation et la reconnaissance des visages [32][33], que les visages verticaux. Ainsi, les motifs elliptiques horizontaux sont plus efficaces et plus pertinents que les modèles circulaires.

- En combinant les informations horizontales et verticales, les performances de précision sont améliorées [33]. Ainsi, au lieu d'utiliser une seule description ELBP horizontale, nous proposons de la fusionner avec sa contrepartie verticale pour améliorer la discrimination de la représentation obtenue.

Les détails de l'ELBP ainsi que les preuves à prouver pour son efficacité sont décrits au chapitre 3 de ce document.

En outre, en utilisant le ELBP comme description primitive, nous constituons des méthodes elliptiques des grandeurs de bord orientées (EPOEM) et modèles locaux de gradients (LPOG), pour les appliquer sur des images basées sur le gradient.

1.3.2 Méthode basée sur Local Phase Quantification (LPQ)

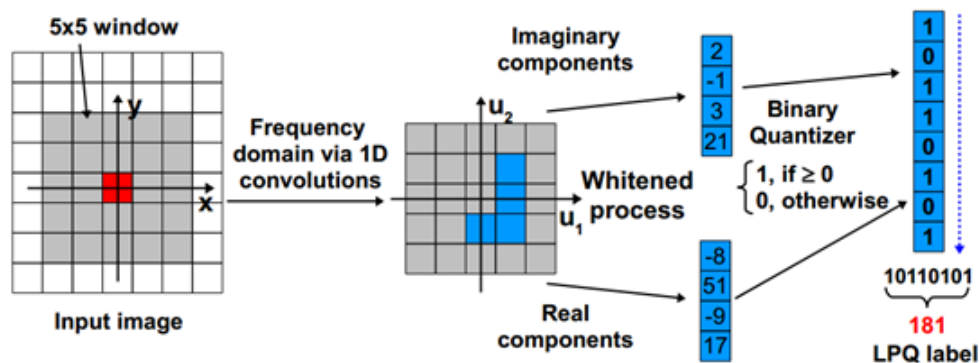


Figure 1.5: Schéma de codage LPQ.

Très récemment, la Quantification de Phase Locale (LPQ), un descripteur de texture tolérant le flou [34], a été étudiée plus en détail pour la reconnaissance de visage. Dans le problème de classification des textures, le LPQ gagne de meilleures performances que les méthodes Gabor basées sur les ondelettes et les méthodes LBP, en particulier lorsqu'on travaille avec des images floues. Ayant des expériences avec le cas de LBP (également initialement développé pour la classification des textures, mais a rapidement gagné sa meilleure notoriété comme une méthode d'extraction de caractéristiques en reconnaissance de visage), de nombreux chercheurs ont en outre étudié l'utilisation de LPQ pour RF. Tout en étant prouvé être flou insensible [34], LPQ été en outre signalée comme un descripteur qui donne des bons résultats en cas de reconnaissance des images floues [35].

L'opérateur LPQ appliquée sur un pixel d'image se fait en utilisant la transformée de Fourier de courte durée (STFT) sur une fenêtre de taille $M \times M$, dont le centre est l'image elle-même, avec quatre fréquences scalaires. Quatre composantes imaginaires et quatre composantes réelles sont ensuite blanchies sur la base d'un paramètre avant d'être

quantifiées binaires pour obtenir l'étiquette LPQ pour le pixel donné. Le procédé d'application d'un opérateur LPQ sur un pixel d'image est illustré sur la Fig. 1.5.

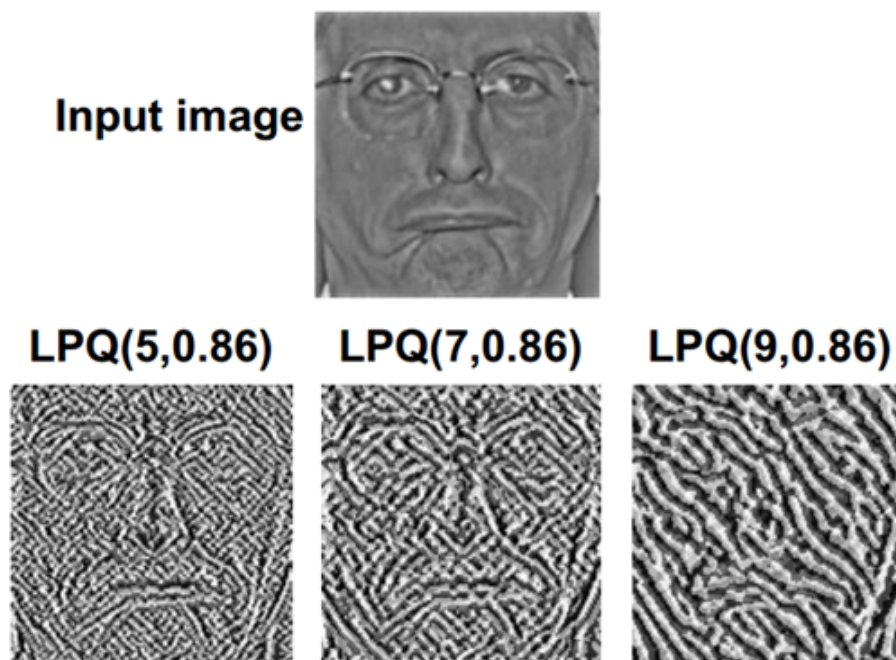


Figure 1.6: Une image de visage et ses LPQ.

Après avoir employé un opérateur LPQ pour produire une image LPQ (la figure 1.6 montre des exemples) à partir d'une image de face d'entrée, Ahonen et al. [35] ont exploité la même technique dans [16] pour la construction de la représentation de la face LPQ en concaténant des séquences histogrammes d'images LPQ sous-régions rectangulaires. Mais contrairement au vecteur de caractéristique LBP [16], dont la taille peut être réduite efficacement en employant des motifs uniformes, chaque vecteur LPQ est une description de 256 bits. Par conséquent, chaque vecteur LPQ est environ quatre fois plus long qu'un vecteur LBP avec 8 pixels voisins (représentation de 256 bins), l'opérateur LBP le plus utilisé, lors de l'utilisation des mêmes sous-régions divisées.

Pour autant que nous sachions dans le domaine RV, il n'y a pas eu beaucoup de variantes de LPQ qui sont venues après sa première apparition [35]. Dans LPQ, l'information de magnitude n'est pas utilisée. En tenant compte à la fois des caractéristiques de magnitudes et de phases des images obtenues à partir de STFT, Lei et al. [36] a présenté une nouvelle méthode appelée Local Frequency Descriptor (LFD), qui pourrait être considérée comme une variante LPQ, pour une RF de faible résolution. La même technique de codage que dans l'opérateur LBP est appliquée sur l'image de grandeur alors qu'une qualification binaire est utilisée sur l'image de phase pour générer deux images codées. Le vecteur de caractéristique d'Analyse Discriminante Linéaire (ADL) est ensuite construit en concaténant les séquences d'histogrammes des sous-régions de ces images. Cette méth-

ode apporte des performances plus élevées que le modèle binaire local (LBP) et LPQ en cas de faible résolution.

Comme on le verra dans [35], LPQ, lorsqu'elle est utilisée dans un modèle de reconnaissance de visage basée sur le modèle binaire, est plus robuste que LBP pour traiter les images de flou, d'illumination et d'expression faciale, mais en littérature RF, LPQ n'a pas reçu la réputation méritée et a été généralement éclipsé par LBP et ses variantes. Dans cette dissertation, nous montrerons que LPQ est plus efficace que LBP et ses variantes contre tous les problèmes difficiles et quand il s'engage avec WACP, il surpasse beaucoup d'autres systèmes de Reconnaissance de visage de pointe.

En outre, motivé par son efficacité face aux défis en matière de RF et dans les composantes du filtre monogénique et les images en gradient, nous exploitons LPQ pour constituer deux nouvelles méthodes d'extraction de traits faciaux: Patch LPQ basé sur les compositions monogéniques (PLPQMC) et locales Patterns of Gradients (LPOG).

1.3.3 Méthodes basées sur les ondelettes de Gabor

L'ondelette de Gabor est une fonction à moyenne nulle, oscillante et définie sur un support compact. Dans le cas 1D, elle est définie par les expressions :

$$\psi^g(x) = \exp\left(-\frac{x^2}{2}\right)\exp(jw_0x) \quad \psi^g(x) = \exp\left(-\frac{x^2}{2}\right)\sin(w_0x) \quad (1.1)$$

Sous ses formes respectivement complexe et réelle, L'ondelette de Gabor ne peut pas être utilisée simplement pour établir une transformation, du fait qu'elle ne permet pas de constituer des atomes orthogonaux et que les ondelettes duales ne disposent pas d'une expression analytique. De ce fait, elle n'a été utilisée que pour la modélisation paramétrique avec une recherche itérative de la solution. En revanche, elle réalise un excellent compromis entre les descriptions spatiale et fréquentielle de la fonction modélisée.

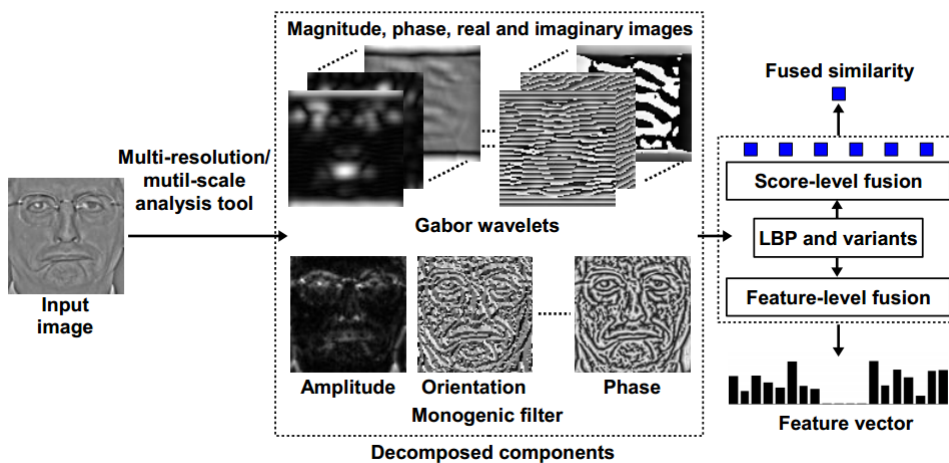


Figure 1.7: Schéma général d'extraction de fonctionnalités multi-résolution / multi-échelle.

L'ondelette de Gabor est d'abord employée pour décomposer une image de visage en plusieurs composantes sous la forme des images indépendantes. Le nombre d'images dépend de la méthodologie dans laquelle elles sont utilisées, par exemple les données d'ondelette de Gabor génèrent généralement 40 parties complexes à 5 échelles et 8 orientations pour englober suffisamment d'informations d'une image d'entrée, mais nous pouvons générer 160 Gab ou images wavelet (40 images réelles, 40 images imaginaires, 40 images de magnitude, 40 images de phase) au total de ces 40 parties complexes. Ensuite, LBP et ses variantes sont appliquées sur ces images composantes pour extraire des caractéristiques faciales utiles pour RF et aboutir à des descriptions différentes, chacune pour une image. Pour combiner toutes ces descriptions distinctes, il existe deux stratégies fusionnées: la fusion au niveau du score et la fusion au niveau des fonctionnalités. Par des méthodes de fusion au niveau du score, les similitudes fusionnées entre les images tests et les images de galerie sont calculées sur la base de différents scores et sont utilisées pour déterminer les identités des images de test. Dans les algorithmes de fusion de niveau de fonctionnalité, les vecteurs de caractéristiques globales sont obtenus en incorporant toutes les descriptions de l'étape précédente. Par conséquent, ces vecteurs sont de grande dimension avec beaucoup d'information redondante et doivent être projetés dans un sous-espace avant de procéder à l'étape de classification de manière efficace.

Depuis ses débuts en littérature RF [16], de nombreux chercheurs ont tenté de combiner LBP avec des ondelettes de Gabor en employant un modèle multi-résolution / multi-échelle tel que mentionné dessus pour améliorer la performance de reconnaissance. Histogramme de modèle binaire Gabor local séquence (LGBPHS) [37], ensemble de la FDA par morceaux (EPFDA) basés sur l'espace Histogrammes des modèles binaires locaux de Gabor [38], histogramme des modèles de phase de Gabor (HGPP) [39], en fusionnant les modèles locaux de magnitude et de phase de Gabor (FLPGMP) [41] et surface de Gabor feature (GSF) [42] sont les plus représentatives méthodes. Suite à la stratégie de fusion au niveau des fonctionnalités, le vecteur LGBPHS est formé en utilisant l'opérateur LBP sur 40 images de magnitude de Gabor.

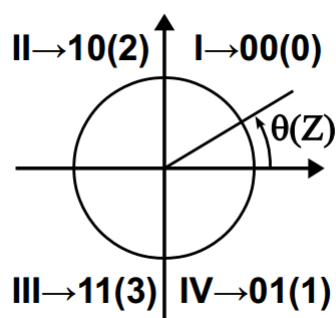


Figure 1.8: Schéma de codage des bits.

Les auteurs [37] ont ensuite utilisé un modèle LBP avec la distance carrée pour la clas-

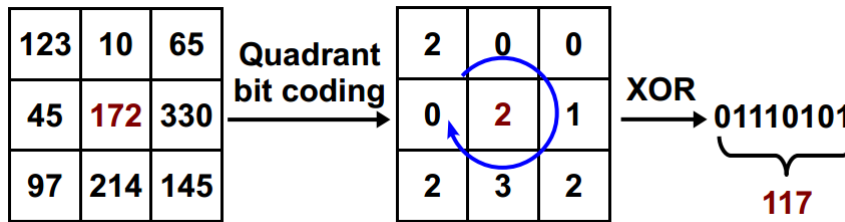


Figure 1.9: Schéma de codage de LXP.

sification et ont obtenu de bons résultats sur les bases de données FERET [43] et AR [44]. Pour éviter la FDA directe sur des vecteurs de caractéristiques de très grande taille, EPFDA [45] partitionne chaque image en petits blocs qui sont ensuite divisés en sous-blocs où les dépoteurs Gabor sont appliqués pour produire un segment d'entité de chaque bloc. L'ensemble FDA est un processus effectué sur ces segments pour construire des sous-espaces FDA dans lesquels chaque image de face est constituée comme une séquence de segments d'entités projetées. Une règle de somme pour combiner des classificateurs individuels sur des segments d'entités projetées est utilisée dans l'étape de classification. Dans HGPP [46], un schéma de codage de bits de quadrants a d'abord été proposé pour assigner chaque valeur de phase par un nombre de 2 bits de 0 à 3 (pour plus de détails, voir la figure 1.9). Ensuite, pour extraire des caractéristiques des images de phase Gabor, un descripteur LBP similaire nommée motif XOR local (LXP) (voir la figure 1.9 pour plus de détails) a été introduit en appliquant l'opérateur XOR sur les codes de bits de quadrants. Puis LXP a été utilisé sur des parties réelles et imaginaires de la phase Gabor images (les auteurs ont utilisé totalement 90 images) pour coder les modèles de phase Gabor globale et locale. Les résultats élevés de HGPP montrent que l'information de phase de Gabor joue également un rôle important dans RF. Tan et al. [47] ont proposé une méthode de fusion au niveau des fonctionnalités pour fusionner les LBP avec les caractéristiques d'ondelettes de Gabor. Ils ont d'abord utilisé la ACP pour réduire les représentations des ondelettes LBP et Gabor puis appliqué les vecteurs communs discriminants du noyau [48] pour projeter des vecteurs de caractéristiques fusionnés dans le sous-espace discriminant pour poursuivre la tâche de classification. Dans FLPMP [49], la définition de LXP et LBP ont été utilisées pour exploiter la magnitude des deux ondelettes de Gabor et l'information de phase, respectivement. Les vecteurs d'amplitude et de phase obtenus ont ensuite été introduits dans une procédure FDA (BFDA) basée sur des blocs pour raccourcir leurs longueurs et éliminer les informations inutiles qu'ils portaient. Les deux stratégies de fusion au niveau du score (par une règle de somme) et au niveau de l'élément ont été évaluées sur des vecteurs projetés [49], et à la fin, le premier a surpassé celui-ci en fournissant d'excellents résultats sur FERET et d'autres bases de données faciales. La méthode GSF [50] utilise LBP sur des cartes combinées d'images de magnitude de Gabor et leurs dérivées 1, 2. L'EPFDA [45] a été utilisé pour réduire les longueurs des vecteurs de caractéristiques GSF et les scores pondérés basés sur les distances de cosinus ont été calculés pour la phase de classification. GSF [50] a obtenu des résultats de pointe sur la

base de données FERET.

Récemment, certaines nouvelles méthodes à base d'ondelettes de Gabor ont été proposées et ont obtenu des résultats très significatifs dans [51], Hussain et al. Proposent une nouvelle méthode d'extraction de caractéristiques appelées Quantification Locale Quantique (LQP) en utilisant la quantification vectorielle et la table de consultation pour construire la description faciale sur les images d'ondelettes de Gabor. Ayant des avantages par rapport à de nombreuses méthodes basées sur les ondelettes LBP et Gabor, LQP donne d'excellentes performances pour l'identification et la vérification des faces, lors de l'incorporation avec ACP et cosine métrique. Les caractéristiques statistiques locales (FSL), une nouvelle représentation faciale, ont été proposées par Yang et al. [52]. Les auteurs utilisent une technique multi-partition pour améliorer l'invariance de SLF à l'erreur d'enregistrement d'image d'abord. Ensuite, un modèle de représentation basé sur le noyau est adopté pour exploiter complètement les caractéristiques discriminantes intégrées dans SLF. Une structure RF appelé SLF-based robust kernel representation (SLFRKR) est également proposée. Des expériences approfondies montrent que le SLF-RKR (utilisant le SLF basé sur l'amplitude de Gabor) est robuste aux occlusions et gagne des résultats supérieurs en comparaison avec des systèmes à la pointe de la technologie, sauf lorsqu'il doit faire face à des variations de pose. Un autre roman Gabor wavelets basé méthode d'extraction de caractéristiques appelées Mesures ordinales de Gabor (MOG) est introduit dans [53]. Les mesures ordinales, qui reflètent l'information de la relation de commande entre des variables multiples (intensités ou valeurs d'entités), sont utilisées pour coder la caractéristique faciale à partir de l'amplitude, de l'orientation, des images réelles et imaginaires de 40 composantes d'ondelettes de Gabor (90 images au total). Chaque vecteur de caractéristiques GOM est affiné par une stratégie de partition basée sur les blocs pour avoir des dimensions de 5760. L'algorithme ADL est utilisé pour réduire encore les tailles des vecteurs GOM et une fusion des distances cosinus de la règle de somme est exploitée dans la classification. Les résultats fournis par GOM sont très étonnants, mais pas plus élevés que ceux de SLF-RKR [52]. En outre, un inconvénient de GOM est qu'il est relativement lent, en passant environ 700ms pour le traitement d'une image de visage. Le coût de cette approche demeure un problème non résolu. Pour les systèmes en temps réel, contexte de surveillance vidéo, où le calcul est un objectif primordial, une question doit être résolue.

Une autre observation importante des meilleurs résultats des travaux cités ci-dessus est que: Une méthode d'extraction de traits uniques qui fonctionne uniquement sur des images d'intensité (telles que LBP, LPQ et leurs variantes), même avec les paramètres accordés et à son meilleur, n'est pas suffisant pour répondre à l'exigence d'une haute précision et un système fiable. Par une sorte de manière, une représentation faciale robuste doit contenir des dispositifs utiles qui sont subtilement extraits des composants multi-résolution / multi-échelle. Dans le même temps, avoir un temps de calcul élevé.

1.4 Autres méthodes

Certains autres descripteurs locaux, tels que la SIFT [54] et les histogrammes de gradients orientés (HOG) [55], ont été couramment utilisés dans de nombreuses applications du monde réel en raison de leurs calculs efficaces, de leur résistance aux occlusions partielles et de leur relative insensibilité aux changements de point de vue. Bien que le SIFT et le HOG aient été évidemment prouvés comme deux des meilleures méthodes pour encoder l'information de bord ou de forme locale, il n'y a pas beaucoup de contributions dans la construction de systèmes RF robustes.

Selon [60], les performances des fonctionnalités de HOG sur la base de données FERET sont pires que celle des ondelettes LBP et Gabor. En résumé, ces preuves signifient que les EIPD et HOG ne peut pas ouvrir la voie à une extraction robuste qu'ont attendue à cette thèse.

Comme les images de visage capturées dans le contexte de la vidéosurveillance sont de faible résolution alors que les images de galerie sont souvent de haute qualité, il est préférable d'avoir des images de test avec une meilleure résolution. Pour ce faire, des techniques de super-résolution sont utilisées pour produire des images à haute résolution avant d'effectuer l'extraction de caractéristiques pour améliorer la précision du système. Nous avons étudié trois méthodes bien connues de super-résolution, qui utilisent différents algorithmes pour construire une image à haute résolution à partir d'une ou plusieurs images d'entrée faible résolution:

- Interpolation bicubique: le pixel d'image de sortie est une moyenne pondérée de pixels en le quartier 4×4 le plus proche.
- Méthode basée sur la régression [61]: Ici, l'idée de base est d'apprendre une cartographie à partir d'images faible résolution d'entrée pour cibler les images haute résolution en fonction de paires d'images d'exemple en utilisant la régression de crête du noyau. Pour supprimer les effets de flou et de sonnerie autour des bords forts introduits par la régression, un modèle qui prend en compte la propriété de discontinuité des images est utilisé pour le post-traitement.

Une fois que les images de sortie de ces méthodes sont générées, nous les utilisons avec Eigenfaces [62] et LBP [16] sur la base de données SCACE pour évaluer si certaines améliorations de précision sont atteintes ou non. Avec les Eigenfaces, les TR ne sont pas améliorés. L'amélioration globale du TR à la LBP est négligeable, soit environ 0,3%. Comme on le verra dans [63], même en combinant une méthode de super-résolution avec la relation entre une image haute résolution et une de sa version faible résolution dans l'ensemble d'apprentissages pour construire une meilleure cartographie des images de test LR, les TR sur SCface sont très faibles (TR n'est que de 20,2%). Ces résultats soulignent que les techniques de super résolution ne sont pas la bonne solution, du moins à l'heure actuelle, pour traiter les images de test LR acquis par les caméras de surveillance. En outre, un inconvénient des algorithmes super résolution est qu'ils viennent avec un coût de calcul significatif pour la production d'images HR. En tant que question ouverte, nous

pensons que la bonne façon de traiter la RF à basse résolution peut dégrader la résolution des images de la galerie des RH au lieu d'améliorer celle des images de la sonde LR.

1.5 Conclusion

Ce chapitre présente un aperçu actualisé de l'état de l'art des principales approches d'extraction de traits de visage avec leurs méthodes les plus représentatives. Commencer par Eigenfaces et l'Analyse Discriminante Linéaire (ADL), les deux algorithmes globaux les plus populaires, et le reste du chapitre est consacré de deux méthodes locales d'extraction de caractéristiques. À partir des méthodes d'ondelettes Gabor pures, de LBP et de plusieurs de ses variantes, des méthodes basées sur LPQ, des méthodes multi-échelle / multi-résolution basée sur une approche simple et quelques transformations multi-résolutions / multi-échelles bien connues comme les ondelettes de Gabor et le filtre monogénique, des méthodes basées sur la représentation, toutes sont décrites en mettant l'accent sur leurs idées clés, leurs avantages et leurs limitations. En outre, d'autres descripteurs locaux et des approches basées sur la résolution d'image sont également considérés.

Tout en analysant les avantages et les inconvénients de ces méthodes, nous présentons également les principaux concepts de nos approches proposées dans cette thèse:

- Utilisation d'échantillons elliptiques pour capturer des micro textures à partir d'images de visage pour former Elliptical Local Binary Pattern (ELBP). Combiner les ELBP horizontaux et verticaux dans la construction d'une représentation faciale avec un ensemble de fonctionnalités plus riches pour performance de reconnaissance.
- Appliquer l'opérateur LPQ sur les composants monogéniques pour construire une description de visage multi-résolution / multiscale.
- Utilisation de ELBP sur des images d'amplitude de bord orientées pour former un motif elliptique de la méthode d'extraction de caractéristique de bordure orientée.
- Intégrer deux types de modèles locaux, ELBP et LPQ, directement sur le gradient d'images pour tenir compte des caractéristiques significatives des caractéristiques locales et les avantages des images de gradient sur l'intensité brute.

CHAPITRE

2

Bases de données et méthodes d'apprentissage

Les systèmes de reconnaissance faciale sont très difficiles à comparer, car leurs tests doivent être effectués sur un grand nombre d'échantillons dans des conditions diverses, représentant des scénarios en termes de variations dans la taille de la base de données de différents modèles, le capteur utilisé, les conditions d'observation, l'éclairage et l'arrière-plan. Par conséquent, les bases de données publiques à grande échelle avec un protocole bien défini peuvent aider à atteindre ces objectifs. Ce chapitre présente les bases de données, y compris leurs protocoles communs et leur cadre d'évaluation, utilisés pour caractériser et évaluer les méthodes de reconnaissance faciale développées et étudiées dans cette thèse. Ensuite, nous terminerons le chapitre par une synthèse des méthodes de réduction de dimension et à la fin de ce chapitre un résumé est donné pour récapituler les informations utiles et nécessaires relatives aux différentes bases de données qu'on manipule.

2.1 Bases de données FERET

La base de données Feret [64] a été recueillie à l'université de George Masson et au laboratoire de recherche de l'armée américaine. Cette base de données et un ensemble étendu de données de performance réalisés sur cette base de données est disponible pour une gamme d'algorithmes de recherche et de systèmes commerciaux de reconnaissance faciale. Les images sont capturées en échelle de gris à la résolution 256 par 384. La base de données contient 14,126 images dont 3 816 sont des images frontales. Cette base de données est divisée en un ensemble de galeries et quatre ensembles de sondes. Les images d'échantillon sont présentées à la figure 2.1.

Le cadre d'évaluation publique à source ouverte [65] décrit à la section 2.4.1 a été utilisé pour tester et évaluer la performance de nos méthodes avec d'autres. Toutes nos expériences ont appliqué l'ensemble d'entraînement standard CSU pour estimer les paramètres de la méthode.

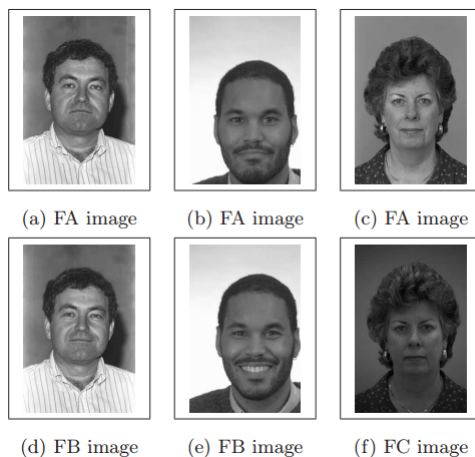


Figure 2.1: Exemples d'images de la base de données Feret.

2.2 Base de données FRGC 2.0 avec EEB (Environnement d'Expérimentation Biométrique)

La reconnaissance de visage Grand Challenge Version 2.0 (FRGC 2.0) [66] est une grande base de données contenant des images qui fixes les couleurs de haute résolution (2 272 par 1 704 pixels), des images 3D et des images multiples d'un sujet au total 50 000 enregistrements. Les données pour les expériences FRGC sont divisées en un ensemble d'entraînement et d'essais. Les données de l'ensemble de formation ont été recueillies au cours de l'année scolaire 2002-2003. L'ensemble de formation contient 12 776 images fixes de 222 sujets, où 6 389 images sont collectées dans un environnement contrôlé et les autres sont acquises dans un environnement incontrôlé. Les images contrôlées capturées en studio sont prises en pleine pose frontale sous deux conditions lumineuses et avec deux expressions faciales. Les images de face frontale non contrôlées sont prises dans l'illumination variable, comme les couloirs, les oreillettes ou à l'extérieur et avec deux expressions faciales. Certaines des images non contrôlées sont hors sujet. Les données du jeu de test contiennent 24 042 images fixes de 466 sujets, dont 222 sujets sont communs à l'ensemble d'entraînement, mais leurs images ne sont pas partagées avec l'ensemble d'entraînement.

Il y a six expériences et notre travail est uniquement focalisé sur les expériences 1 et 4. L'expérience 1 est conçue pour mesurer la performance de reconnaissance automatique de visage à partir d'images frontales prises sous éclairage contrôlé. Dans cette expérience,

16 028 images provenant de 466 sujets dans l'environnement contrôlé sont utilisées pour établir une matrice de confusion de similarité de $16,028 \times 16,028$. L'expérience 4 est conçue pour mesurer la performance de reconnaissance sur des images fixes de face frontales contrôlées ou non contrôlées. Dans cette expérience, l'ensemble cible se compose de 16 028 images fixes contrôlées et l'ensemble de requêtes contient 8 014 images fixes non contrôlées. Par conséquent, la dimension de la matrice de confusion est de $16,028 \times 8,014$. Les images d'exemple sont présentées à la figure 2.2.

L'environnement de test est appelé l'environnement d'expérimentation biométrique (BEE) qui constitue l'infrastructure du test de fournisseur de reconnaissance faciale (FRVT 2006). Il permet à l'expérimentateur de se concentrer sur l'expérience en simplifiant la gestion des données de test, la configuration de l'expérience et le traitement des résultats. Dans notre travail, le format binaire de la matrice de confusion de similitude est fourni comme entrée de BEE. Les taux de vérification mentionnés, dans la section 2.4.2, sont rapportés, correspondant au masque 1, 2 et 3, où le masque 1 se concentre sur les images capturées en un semestre, le masque 2 dans une année et le masque 3 entre les semestres.

2.3 Base de données XM2VTS avec protocole Lausanne

La base de données multimodale étendue M2VTS (XM2VTS) [67] inclut la couleur images, données audio, séquences vidéo et modèle 3D. Dans notre travail, nous utilisons images frontales capturées à la résolution 720 par 576 pixels dans un environnement contrôlé. La capture de données est conçue pour présenter un éclairage minimal et que la source de variation au sein de la base de données n'est due qu'aux différences entre les sujets.

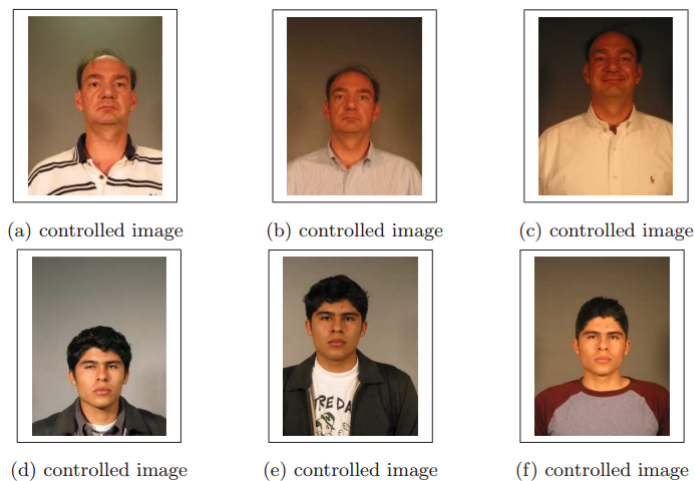


Figure 2.2: Exemples d'images de la base de données FRGC 2.0.

La base de données contient 2360 images de visage frontal de 295 sujets, capturés pendant plus de 4 sessions à intervalles d'un mois. Les tests de vérification des visages sont

effectués en utilisant le protocole de Lausanne qui divise la base de données en formation, évaluation et test ensemble. Il existe deux configurations qui diffèrent par l'attribution de plans particuliers de formation et d'évaluation. L'ensemble de formation est utilisé pour construire des clients sous l'approche de l'apprentissage supervisé. Dans la configuration I, les images client pour la formation et l'évaluation sont acquises à partir des trois premières sessions.

Outre l'ensemble standard, la base de données XM2VTS contient également un ensemble d'images illumination variable appelée Darkened set. Chaque sujet a quatre images supplémentaires avec d'un côté (deux images ont été éclairées de la gauche et deux de la droite). Ces 1 080 images (270 sujets fois 4 images) sont utilisées pour évaluer la performance du système sous des changements d'éclairage. Dans l'expérience avec illumination, les ensembles d'entraînement et d'évaluation proviennent d'images bien éclairées tandis que l'ensemble de tests est remplacé par l'ensemble Darkened. Par conséquent, la tâche de cette expérience est très difficile parce que le modèle de formation est acquis en l'absence de connaissances sur les caractéristiques des données d'image noircies. La figure 2.3 donne quelques exemples d'images.

La décision d'acceptation ou de rejet est basée sur la relation d'une mesure de la similarité entre la galerie et la moyenne des images de formation du client avec notre travail. Le seuil est sélectionné au point d'erreur égal, EER, à dont le taux de faux rejet est égal au taux d'acceptation faux sur l'ensemble d'évaluation. La définition du taux de fausse acceptation et du taux de faux rejet sera donnée en section 2.4. Le taux d'erreur total, TER, déclaré pour les résultats de la base de données XM2VTS est défini comme la somme du taux de faux rejet et du taux de fausse acceptation.



Figure 2.3: Exemples d'images de base de données XM2VTS.

2.4 Mesures de performance

2.4.1 Mesures de performance dans l'identification du visage

Le cadre d'évaluation à code source ouvert a été utilisé pour tester et évaluer la performance de nos méthodes avec d'autres. Dans notre travail, le taux de Rang 1 pour chaque ensemble de sondes et deux mesures statistiques sont utilisées pour comparer les performances des méthodes. Ces mesures statistiques, à savoir le taux de reconnaissance Rang 1 et la probabilité que l'algorithme surpasse un autre, sont évalués en utilisant un ensemble d'images de sonde et un ensemble d'images de galerie. Dans ce test statistique, une paire d'images de sonde-galerie pour chaque sujet est tirée des 12 paires d'images correspondantes dans chaque image une expérimentation de 160 sujets et chaque sujet a 4 images. Afin de déduire la qualité de la généralisation à une plus grande population des sujets, une permutation, générant une distribution d'échantillonnage du taux de reconnaissance pour différents rangs calculer de façon répétée le taux de reconnaissance à partir de différents ensembles de 10 000 essais, est utilisé. La moyenne du taux de reconnaissance au rang 1 défini dans [84] est la moyenne du taux de reconnaissance au rang 1 au total 1000 essais. Pour estimer la probabilité que l'algorithme produise un autre, $P(\text{Alg1} > \text{Alg 2})$ dans Rang 1 sur un total de 1000 essais [68], la différence signée entre le taux de reconnaissance de Alg1 et Alg 2 est calculée dans chaque essai. On détermine $P(\text{Alg1} > \text{Alg 2})$ en additionnant les probabilités des différences supérieures à 0. Il existe une différence significative entre Alg1 et Alg 2 en termes de reconnaissance si $P(\text{Alg1} > \text{Alg 2})$ est supérieur ou égal à 0.95. Sinon, la performance des deux algorithmes est considérée comme similaire.

2.4.2 Mesures de performance dans la vérification des visages

Les systèmes de vérification font deux types d'erreur:

- Prendre des mesures biométriques de deux personnes différentes pour être de la même personne, à savoir False Acceptation (FA).
- Confondre deux mesures biométriques de la même personne avec deux personnes différentes, à savoir False Rejection (FR). La performance est mesurée en termes de faux taux d'acceptation (FAR) et de faux taux de rejet (FRR), définis comme suit:

Nombre d'accès

$$FAR = \frac{\text{Nombre de FAs}}{\text{Nombre total des vrai accès}} \quad (2.1)$$

$$FRR = \frac{\text{Nombre de FRs}}{\text{Nombre total des vrai accès}} \quad (2.2)$$

Il existe un compromis entre FAR et FRR dans chaque système de vérification, car les FAR et FRR sont une fonction du seuil (T). Pour une valeur donnée du seuil (T), il existe une paire de FAR(T) et FRR(T). Ils peuvent être tracés l'un contre l'autre sous la forme d'une courbe connue sous le nom de Receiver Operating Characteristic (ROC)

pour exprimer le comportement. Des FAR et FRR. Dans les expériences XM2VTS, le seuil est généralement choisi.

L'évaluation définie à $FAR = FRR$, appelée taux d'erreur égal (EER). Il est ensuite appliqué à l'ensemble test pour obtenir FAR et FRR et par conséquent somme des deux pour obtenir le Total Taux d'Erreur (TER). En comparant le TER à d'autres systèmes, nos systèmes peuvent être évalués. D'autre part, le taux de vérification (c'est-à-dire $1 - FRR$) à 0,1% FAR est généralement utilisé pour représenter la précision du système dans des expériences FRGC.

2.5 Base de données Tifinagh

La langue amazighe (berbère) est parlée dans des régions du Maroc, en Algérie, Tunisie, Libye et Siwa (une oasis égyptienne); il est également parlée par de nombreuses autres communautés dans certaines parties du Niger et du Mali. En termes linguistiques, la langue se caractérise par prolifération des dialectes dus à des facteurs historiques, géographiques et sociolinguistiques. Au Maroc, le terme berbère (amazigh) englobe les trois principales variantes marocaines: le Tarifite, Tamazighte et Tachelhite. Une partie importante de populations parle Amzigh. Tous les Marocains sont concernés avec cet alphabet. L'établissement du "Institut Royal de la Culture Amazighe"(IRCAM) a mené une action majeure pour standardiser la langue amazighe. Dans la même foulée, et depuis 2003, la langue amazighe a été intégrée dans le système éducatif marocain. Il est enseigné dans les classes de l'enseignement primaire des différentes écoles marocaines.

Les alphabets amazighs, appelés «Tifinagh-IRCAM», adoptés par l'institut royal de la culture amazighe, ont été officiellement reconnus comme appartenant à la base multilingue prévue par L'Organisation internationale de Normalisation (ISO) [70]. La figure 2.4 représente le répertoire de Tifinagh qui est reconnu et utilisé au Maroc avec leurs correspondants en latin personnages.



Figure 2.4: Caractères de Tifinagh adoptés par l'IRCAM.

2.5.1 Création de la base de données

Pour construire notre base de données, nous devons analyser trente-trois caractères. Première partie du système de reconnaissance de caractères tifinagh qui couvre quatre fonctions pour produire une version nettoyée de l'image d'origine afin qu'elle puisse être utilisée directement et efficacement par les composants d'extraction de caractéristiques de l'OCR. Ces fonctions sont: numériser le texte et le numériser en une image numérique et le nettoyer, convertir l'image scalaire en image binaire, normaliser le texte et reconstruire les données manquantes dans l'images. données.

Les figures 2.5 et 2.6 représentent le processus d'étapes de reconstruction de notre base de données et le processus de reconstruction des données.

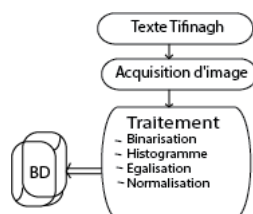


Figure 2.5: Principales étapes de la construction de notre base de données.

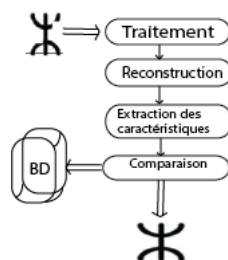


Figure 2.6: Système de reconnaissance des caractères.

Dans cette thèse, nous présentons un algorithme de reconstruction à partir de scans 2D incomplets pour les caractères Tifinagh. Cet algorithme est basé sur l'utilisation de la corrélation entre le bloc perdu et ses voisins. Ce système proposé comprend trois parties principales: pré-traitement, extraction et reconnaissance des caractéristiques. Dans la première étape, nous construisons une base de données de caractères Tifinagh. Dans la deuxième étape, nous appliquerons «algorithme d'analyse de forme». Dans la partie de classification, nous utiliserons les Réseaux de Neurones(RN). Les résultats de la simulation démontrent que la méthode proposée donne de bons résultats.

2.6 Base de données des composants électroniques

2.6.1 Circuit électrique

Les capteurs électroniques, en particulier ceux basés sur les cellules à semi-conducteurs, connaissent un grand intérêt dans divers domaines comme la télécommunication, l'instrumentation, la médecine, et autres. La détermination des modèles électriques précis de ces composants permet d'optimiser le temps, le coût et l'énergie de la fabrication de ces éléments, et facilite ses intégrations aux logiciels de simulation. Dans la littérature on trouve plusieurs méthodes analytiques ou numériques [71] qui peuvent réaliser cette tâche, mais chaque méthode reste limitée selon des spécifications bien déterminées. Dans le paragraphe suivant nous présentons une nouvelle méthode basée sur la reconnaissance de la forme de la fonction de transfert d'un tel circuit électronique. Pour atteindre notre système de reconnaissance, il est nécessaire de construire une base de données de circuits simples dont les fonctions de transfert sont connues.

2.6.2 Diagramme de Nyquist

Le diagramme de Nyquist est la représentation la plus immédiate de la fonction complexe de réponse en fréquence: rapporté dans le plan, la partie réelle le long de l'axe horizontal et la partie imaginaire le long de l'axe vertical [71][72][73][74][75]. Après avoir sélectionné les circuits de notre base de données, nous devons tracer leurs diagrammes de Nyquist pour cela il est nécessaire de déterminer la fonction de transfert qui est équivalente à l'étude d'impédance de circuit électronique.

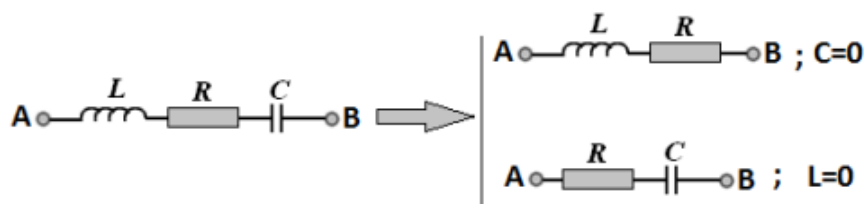


Figure 2.7: Modèle série RLC .

Pour construire notre base de données nous devons analyser électroniquement les 17 circuits: d'abord, nous déterminons leurs fonctions de transfert (impédances équivalentes) et traçons leur diagrammes de Nyquist. Deuxièmement, nous traiterons l'image des courbes de Nyquist. Cette étude informatique repose sur les étapes de prétraitement suivantes: binarisation, égalisation d'histogramme, standardisation et sauvegarde des images dans une extension sélectionnée (jpg, png, bmp, ...). Le schéma de la figure 2.8 résume les étapes à suivre pour atteindre notre base de données:

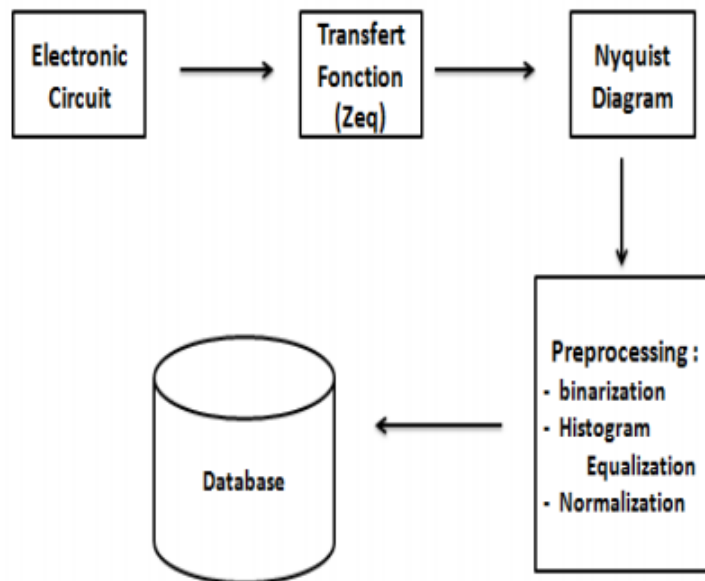


Figure 2.8: Étapes principales de la construction de notre base de données.

Cette base de données contient 170 images de 17 diagrammes de Nyquist de circuits. Pour chaque circuit, nous avons 10 images différentes de diagrammes de Nyquist binarisés et normalisés à une résolution de 283 x 390 pixels. Dans notre simulation, et 85 images Nyquist ont été utilisées pour tester l'identification performance.

Toutes les images de diagrammes de Nyquist de 17 circuits électroniques construisent notre base de données de composants électroniques sont montrés à la figure 2.9:

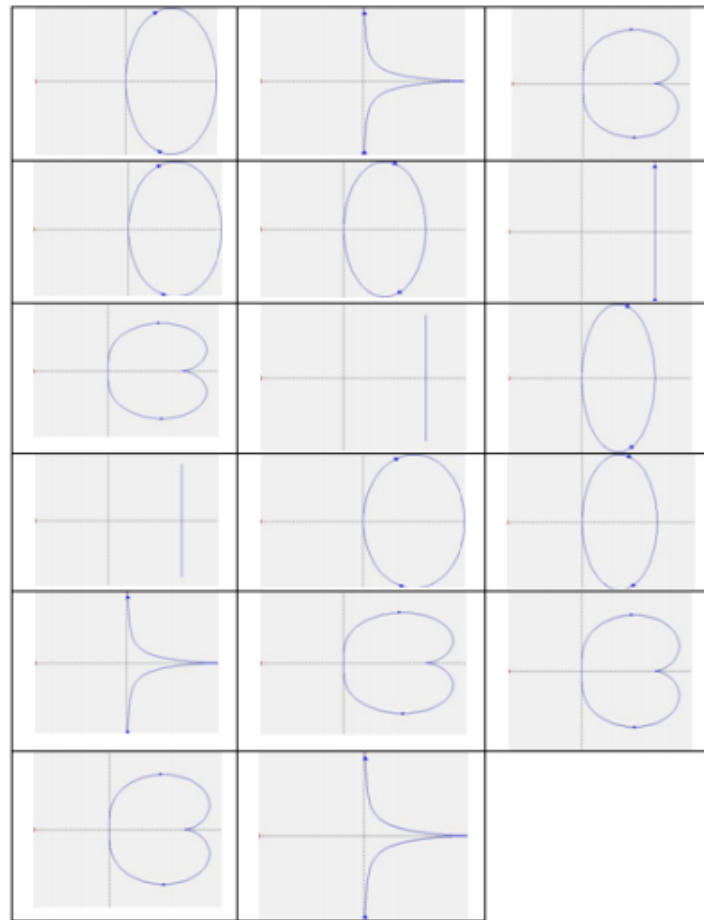


Figure 2.9: Les diagrammes de Nyquist de 17 circuits électroniques construisent notre Base de données des composants électroniques.

2.7 Méthodes de réduction de dimension

Tout système de reconnaissance biométrique comporte une phase très importante basée sur la réduction d'espace. Pour cela nous consacrons cette partie pour présenter les différentes techniques de projection d'espace.

Une image du visage est un signal à 2 dimensions, acquis par un capteur digital. Ce capteur codera la couleur ou l'intensité des différents points de l'image dans une matrice de pixels à deux dimensions dans l'espace des images, nous devons spécifier une valeur pour chaque pixel de cette image. Le nombre de points constituant cet espace devient rapidement très grand, même pour les images de petite dimension. Cette dimensionnalité pose un certain nombre de problèmes pour les algorithmes de reconnaissance, qui se basent sur cette représentation de l'image, à savoir :

Dans un contexte de la reconnaissance de visages, travailler dans un grand espace pose

un problème de complexité de calcul.

Pour les méthodes paramétriques, le nombre de paramètres à estimer peut rapidement dépasser le nombre d'échantillons d'apprentissage, ce qui pénalise l'estimation.

Pour les méthodes non paramétriques, le nombre d'exemples nécessaires afin de représenter efficacement la distribution des données peut être insuffisant.

En 1994, Ruderman [82] a démontré que les images naturelles possèdent une grande redondance statistique. En 1996, Penev [93] a démontré que dans le cas précis des images normalisées des visages, cette redondance statistique est devenue importante.

Les méthodes de réduction de dimensionnalité (comme l'ACP) appliquées aux données 3D, montre des résultats moins convaincants que leur application aux images 2D ; elles sont souvent utilisées en tant que méthodes de référence lors des études comparatives [89].

2.7.1 L'analyse en composantes principales

L'Analyse en Composantes principales (ACP) [88], aussi connue sous le nom de transformée de Karhunen-Loève [90] est une méthode très utilisée en statistique. Introduite par Pearson [91] puis plus tard par Hotelling [92], sa principale idée est de réduire la dimension d'un jeu de données tout en gardant un maximum d'informations. Cela est réalisé grâce à une projection qui maximise la variance tout en minimisant l'erreur quadratique moyenne de la reconstruction. Pour plus de détails, voir [93]. Pour la dérivation Hotelling définit l'ACP comme une projection orthogonale maximisant la variance dans l'espace projeté. Étant donné n échantillons $x_i \in R^D$ et $u \in R^D$ tel que :

$$\|u\| = u^T u = 1 \quad (2.3)$$

Soit U un vecteur orthonormal de projection. Un échantillon x_i est projeté sur u par : $a_i = u^T x_i$. La variance de l'échantillon peut donc être estimée :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.4)$$

Où \bar{x} est la moyenne des projetés des échantillons de la base :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{d'où} \quad \bar{a} = u^T \bar{x} \quad (2.5)$$

Ainsi la variance du projeté est donnée par :

$$\begin{aligned} S^2 &= \frac{1}{n-1} \sum_{i=1}^n (a_i - \bar{a})^2 = \frac{1}{n-1} \sum_{i=1}^n (u^T x_i - u^T \bar{x})^2, \\ &= u^T C u \end{aligned} \quad (2.6)$$

Où

$$C \in R^{D \times D} = \frac{1}{n-1} \sum_{i=1}^n ((x_i - \bar{x})(x_i - \bar{x})^T) \quad (2.7)$$

Est la matrice de covariance de $X = [x_1; \dots; x_n] \in R^{D \times n}$. Le problème de maximisation de la variance dans l'espace projeté peut donc s'écrire : $\max u^T C u$ avec $u^T u = 1$. Le calcul de la solution optimale peut être réalisé grâce au multiplicateur de Lagrange :

$$f(u, \lambda) = u^T C u + \lambda(1 - u^T u) \quad (2.8)$$

Par dérivation partielle selon u :

$$\frac{\partial f(u, \lambda)}{\partial u} = 2Cu - 2\lambda u = 0 \quad (2.9)$$

On obtient :

$$Cu = \lambda u \quad (2.10)$$

Ainsi, le maximum pour le multiplicateur de Lagrange est obtenu si λ est une valeur propre et u un vecteur propre de C . Ainsi la variance décrite par le vecteur de projection u est donnée par λ . D'autres méthodes de dérivation de l'ACP sont données dans [99]. Pour une vue probabiliste de la dérivation de l'ACP, voir [88].

Calcul de l'ACP pour la mise en œuvre de méthodes : il est supposé que le jeu de données d'entraînement est disponible en entier. Ainsi nous avons un ensemble de n observations $x_i \in R^D$ organisés sous forme matricielle $X = [x_1; \dots; x_n] \in R^{D \times n}$. L'estimation de la base de projection de l'ACP revient donc à estimer les éléments propres de la matrice de covariance C de X . Le calcul requiert d'abord l'échantillon moyen :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.11)$$

Puis les échantillons sont normalisés par rapport à la moyenne x_{moy} :

$$\hat{x}_i = \bar{x}_i = x_i - \bar{x}, \quad i = 1, 2, 3, \dots, n \quad (2.12)$$

Pour former la nouvelle matrice $\hat{x} = [\hat{x}_1; \dots; \hat{x}_n]$, la matrice de covariance $C \in R^{D \times D}$ est ensuite calculée par :

$$C = \frac{1}{n-1} \hat{x} \hat{x}^T \quad (2.13)$$

La recherche des éléments propres de C conduit à l'obtention de la base de vecteurs propres $u_i \in R^D$, pour lesquels, à chacun d'eux, est associée une valeur propre λ_i . Généralement triés par ordre décroissant de valeur propre associée, les premiers vecteurs propres forment alors une base dans laquelle la plupart de l'information du jeu de données d'entraînement est gardée.

La dimension de la matrice de covariance dépend de la dimension D des vecteurs du jeu de données, qui peut être relativement grande pour certains types de données (typiquement des images). La méthode décrite plus haut devient alors difficile à appliquer, essentiellement à cause de la recherche des éléments propres de la matrice de covariance C . En effet, pour des images de taille 100*100 par exemple, la matrice de covariance C à inverser est de taille 10000*10000. Cependant, il est connu que pour toute matrice X , les produits matriciels XX^T et $X^T X$ partagent les mêmes valeurs propres différentes de zéro.

Ainsi, le calcul des éléments propres de $C = XX^T$, $C \in R^{D \times D}$ peut se ramener au calcul des éléments propres de la matrice $M \in R^{n \times n}$ où $M = X^T X$. Soit e_i les vecteurs propres de M associés aux valeurs propres δ_i . On a donc :

$$X^T X e_i = \delta_i e_i \quad (2.14)$$

En multipliant à gauche par X les deux côtés de l'équation, on obtient ainsi :

$$X(X^T X e_i) = X(\delta_i e_i) \quad (2.15)$$

$$XX^T(X e_i) = \delta_i(X e_i) \quad (2.16)$$

On voit donc que $X e_i$ est vecteur propre de XX^T et que δ_i est la valeur propre associée, d'où

$$u_i = X e_i \quad \text{et} \quad \lambda_i = \delta_i$$

La matrice M étant beaucoup plus petite que la matrice C (typiquement, on passe d'une complexité de l'ordre de la dimension des échantillons à une complexité de l'ordre du nombre d'échantillons d'apprentissage), les calculs sont donc plus efficaces.

Des variantes de l'ACP ont été proposées. Ainsi plusieurs méthodes ont été proposées pour extraire des axes principaux robustes notamment au bruit contenu dans les images d'apprentissage [97], ou des méthodes basées sur une formulation Espérance- Maximisation de l'ACP [94]. Dans le cas où les données d'apprentissage ne sont pas toutes disponibles au départ (cas de vidéos par exemple), des versions incrémentales de l'ACP ont été mises au point[96]. Des méthodes combinant l'aspect incrémental et robuste ont également été proposées dans [95].

2.7.2 Transformée en cosinus discret

La DCT(Discrete Cosine Transform) est une technique populaire dans l'imagerie et la compression vidéo, qui transforme les signaux dans la représentation spatiale en une représentation de fréquence.

Le 2Dimension-DCT en avant [93], [92] d'une image de bloc $M \times N$ est défini comme suit:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) * \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (2.17)$$

La transformation inverse est définie comme suit:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) * \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (2.18)$$

où

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{M}} & \text{si } u = 0 \\ \sqrt{\frac{2}{M}} & \text{si } 1 < u < M - 1 \end{cases}$$

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & \text{si } v = 0 \\ \sqrt{\frac{2}{N}} & \text{si } 1 < v < N - 1 \end{cases}$$

Où x et y sont des coordonnées spatiales dans le bloc d'image, et u et v sont des coordonnées dans le bloc de coefficients DCT. La figure 2.10 montre les propriétés des coefficients DCT dans les blocs $M \times N$ avec le motif zigzag utilisé par la compression JPEG pour traiter les coefficients DCT. Bien que l'énergie totale reste là même dans les blocs $M \times N$, la distribution d'énergie change avec la plupart des énergies compactées par rapport aux coefficients de basse fréquence. Le coefficient DC est représenté par $C(0,0)$ dans l'avant 2D.

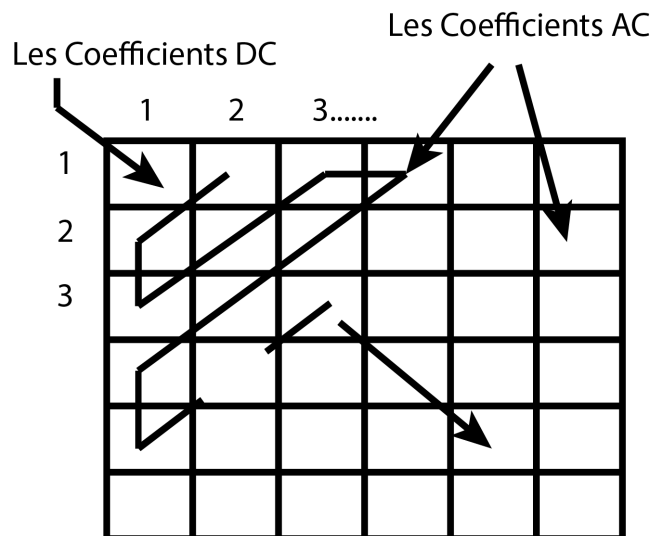


Figure 2.10: Blocs des coefficients DCT et leur sélection en zigzag.

L'équation DCT est simplifiée pour:

$$C(0,0) = \frac{1}{M*N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \quad (2.19)$$

Le coefficient de DC, situé dans le coin supérieur gauche, détient la majeure partie de l'énergie de l'image et représente, la moyenne proportionnelle des blocs $M \times N$. Les coefficients restants ($M \times N$) désignent les changements d'intensité parmi les images de bloc et sont appelés coefficients AC. Le DCT est effectué sur l'image entière obtenue après le traitement des images du visage d'entrée par égalisation d'histogramme.

2.7.3 Décomposition de la valeur singulière (SVD)

La décomposition de la valeur singulière (SVD) représente un sujet important dans l'algèbre linéaire. SVD possède de nombreuses valeurs pratiques et théoriques. La caractéristique spéciale de SVD est qu'elle peut être effectuée sur n'importe quelle matrice réelle (m, n). Disons que nous avons une matrice A avec m ligne et n colonnes, avec rang r et $r \leq n \leq m$. Ensuite, la matrice A peut être factorisé en trois matrices:

$$A = USV^T \quad (2.20)$$

$$\begin{array}{c} \boxed{A} \\ m \times n \end{array} = \begin{array}{c} \boxed{U} \\ m \times m \end{array} \begin{array}{c} \boxed{S} \\ m \times n \end{array} \begin{array}{c} \boxed{V^T} \\ n \times n \end{array}$$

Figure 2.11: une illustration de factorisation de A à USV^T

Où U est une Matrice orthogonale de taille $m * m$.

$$U = [U_1, U_2, \dots, U_r, U_{r+1}, \dots, U_m]$$

Les vecteurs de colonne U_i , pour $i = 1, 2, \dots, m$, forment un ensemble orthonormé:

$$U_i^T U_j = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (2.21)$$

Et la matrice V est une matrice orthogonale de taille $n * n$.

$$V = [V_1, V_2, \dots, V_r, V_{r+1}, \dots, V_n]$$

Les vecteurs de colonne V_i pour $i = 1, 2, \dots, n$ forment un ensemble orthonormé:

$$V_i^T V_j = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (2.22)$$

La matrice S est une matrice diagonale de taille $m * n$ avec des valeurs singulières (VS) sur la diagonale. La matrice S peut être affichée dans la matrice suivante:

$$S = \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 & 0 \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 & 0 \cdots & 0 \\ \vdots & 0 & \cdots & 0 & 0 & 0 \cdots & 0 \\ \vdots & 0 & \cdots & \sigma_r & 0 & 0 \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{r+1} & 0 \cdots & 0 \\ \vdots & 0 & \cdots & 0 & 0 & 0 \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \cdots & \sigma_n \end{bmatrix} \quad (2.23)$$

Pour $i = 1, 2, \dots, n$, σ_i sont appelés Valeurs Singulières (VS) de la matrice A . On peut prouver que

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0 \text{ et } \sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$$

Les V_i et U_i sont appelés vecteurs singuliers droites et gauches de A [16].

2.8 Apprentissage et classification

2.8.1 Les réseaux de neurones artificiels

Les réseaux de neurones artificiels, en anglais Artificial Neural Networks (ANN), constituent la première classe de méthodes qui nous intéressera dans cette thèse. Leur nom est dû aux neurones du cerveau humain. En effet, c'est la description des processus mentaux faite par les neurobiologistes qui est à l'origine des modèles théoriques des ANN. Ils sont aussi appelés réseaux connexionnistes ou réseaux neuromimétiques. Deux éléments caractérisent ces méthodes, une organisation appelée réseau comprenant un certain nombre d'automates aux fonctionnalités relativement simples appelés neurones. L'information se propage dans les réseaux sur des connexions pondérées par des paramètres souvent appelés poids. Le deuxième élément est l'algorithme d'apprentissage dont l'objectif est de faire évoluer les poids du réseau de neurones de manière à obtenir un comportement global intéressant.

Dans cette section, nous passerons rapidement en revue sur les principes élémentaires des réseaux de neurones. Nous établirons aussi un lien entre les méthodes statistiques rencontrées dans 2.2 et ces méthodes connexionnistes.

Principes des réseaux de neurones artificiels

Le neurone formel est l'entité de base qui compose un réseau de neurones. Ce paragraphe présente le modèle de neurones le plus souvent utilisé et qui sera considéré dans le reste

de ce travail. Mathématiquement un neurone peut être représenté par une fonction à plusieurs variables avec une sortie unique. La figure 2.12 représente le schéma classique d'un neurone formel. Les valeurs x_1, \dots, x_n représentent les entrées du neurone i . Les valeurs w_{ij} représentent les poids associés à chaque entrée x_j avec $j \in 1, \dots, n$.

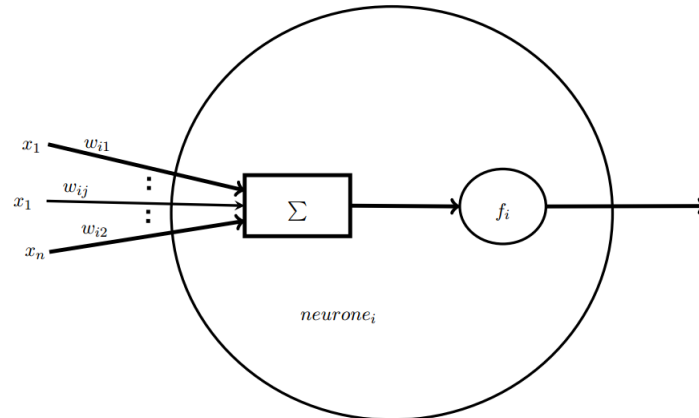


Figure 2.12: Architecture d'un neurone formel

Un neurone peut être également caractérisé par quatre aspects :

- La nature des entrées qui peut être soit binaire soit réelle ; dans notre cas ce sont toujours des réels.
- La fonction de sommation des entrées qui peut être linéaire ou non linéaire. En pratique, ainsi que dans nos expérimentations, c'est une combinaison pondérée par les poids w_{ij} des entrées qui est utilisée.
- La nature de la fonction de transfert utilisée (linéaire ou non-linéaire). En pratique, on trouve des fonctions à seuil (équation 2.24) ou des fonctions sigmoïdes (équation 2.25).
- La nature de la sortie (binaire ou réelle). Dans notre cas, elle sera toujours réelle.

fonction sigmoïde

$$f(x) = \frac{1}{1+e^{-ax}} \quad (2.24)$$

fonction gaussienne

$$f(x) = e^{-\frac{ax^2}{2}} \quad (2.25)$$

Nous pouvons remarquer que la fonction de transfert sigmoïde possède un régime quasi-linéaire au voisinage de 0 et nonlinéaire ailleurs. Cette particularité offre au réseau la possibilité de s'adapter aussi bien aux problèmes linéaires qu'aux problèmes non linéaires.

Les connexions entre les neurones

Les connexions permettent de construire le réseau de neurones en liant les différentes entités ensemble. La plus simple et la plus classique des connexions sont la connexion unidirectionnelle pondérée. Il s'agit d'un arc unidirectionnel pondéré par un poids w_{ij} qui relie le neurone j au neurone i et qui correspond au poids utilisé dans la fonction de sommation. Ces connexions simples sont représentées dans la figure 2.13.

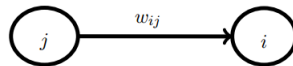


Figure 2.13: Connexion simple entre deux neurones

2.8.2 Support vector regression

Les Support Vector Regression (SVR) sont une adaptation du problème de la classification au problème de la régression de l'algorithme des Machines à Supports Vecteurs, en anglais Support Vector Machines (SVM). Nous ne rentrerons pas dans ce qui suit dans les détails théoriques de ces méthodes, des informations plus détaillées peuvent être trouvées dans [83].

Présentation générale des SVM

Les SVM (ainsi que les SVR) sont une classe d'algorithmes d'apprentissage supervisé. Les SVM sont basés sur les mêmes principes d'apprentissage que les réseaux de neurones. Toutefois, ils possèdent l'avantage d'être plus simples à configurer.

Le modèle du classificateur est construit à partir d'un ensemble d'apprentissages de N exemples étiquetés (x_i, y_i) avec $x_i \in \mathbb{R}^p$ et $y_i \in [-1; 1]$ selon la classe (p représente la dimension des vecteurs d'entrée ou encore le nombre de caractéristiques dans les exemples d'entrées). Il est à remarquer que la version originale des SVM traite uniquement le cas binaire à deux classes, mais des extensions existent pour la multi-classe. L'apprentissage permet de construire la fonction f de décision appelée aussi hyperplan séparateur.

$$f(x) = \text{signe}(\langle w, x \rangle + b) \quad (2.26)$$

Avec $w \in \mathbb{R}^p$ et b les paramètres permettant de déterminer dans quelle partie de l'hyperplan se trouvent l'exemple x . La figure 2.14 présente le cas linéairement séparable, où la marge Δ est définie par la distance minimale entre les deux points des différentes classes.

Support Vector Regression

Le principe des SVM est repris et adapté par les SVR pour modéliser un problème de régression. Le but est d'approximer un ensemble de données (x_i, y_i) par une fonction f .

$$f(x) = \langle \omega, x \rangle + b \quad (2.27)$$

telle que:

$$|f(x) - y_i| \leq \epsilon \quad (2.28)$$

Avec $i \in [1, \dots, N]$.

L'idée consiste à minimiser le terme ω tout en étant sous la contrainte de ne pas dépasser un taux d'erreur ϵ . De point de vue graphique, ceci revient à trouver une zone du plan qui contient tous les exemples x_i de largeur 2ϵ appelée tube (voir figure 2.14). Si on considère la minimisation de $\|\omega\|^2$, on obtient le problème quadratique suivante :
minimiser $\|\omega\|^2$

$$y_i - \langle w, x_i \rangle - b \leq \epsilon \quad \langle w, x_i \rangle + b - y_i \leq \epsilon \quad (2.29)$$

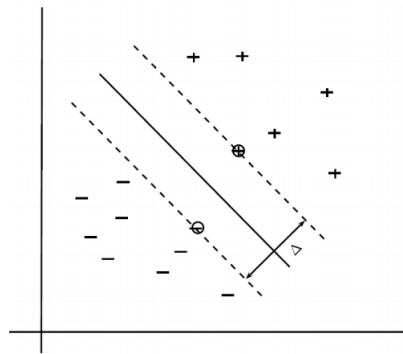


Figure 2.14: Hyperplan linéaire séparant, avec la marge Δ , les points positifs des points négatifs

2.9 Conclusion

Dans ce chapitre, trois bases de données bien connues sont introduites pour évaluer nos systèmes. XM2VTS et FRGC 2.0 sont utilisés pour les tâches de vérification tandis que FERET est utilisé pour les tâches d'identification. La différence entre le protocole de Lausanne dans XM2VTS et le protocole FRGC2.0 est que le protocole de Lausanne est pour la vérification de l'ensemble fermé et FRGC 2.0 pour les expériences semi-ouvertes. Dans la vérification semi-ouverte, certains sujets de la galerie ou de la sonde ne figurent jamais dans l'ensemble d'entraînement. Cependant, lors de la vérification de l'ensemble, les sujets en galerie sont dans les ensembles de formation et d'évaluation. En d'autres termes, la vérification semi-ouverte-set est plus difficile que la vérification fermée pour les méthodes d'apprentissage supervisées. Les performances de nos systèmes, utilisant ces bases de données, seront discutées dans les chapitres 3 et 4.

Dans la troisième partie de ce chapitre, nous avons décrits un système automatique de

reconnaissance de circuit électronique fondé sur le diagramme de Niquist. Nous avons des solutions pour la quasi-totalité de ses étapes, y compris les étapes de la reconnaissance, le prétraitement, la réduction de la dimension et la classification, en adoptant des techniques rationnelles pour chaque étape, à l'exception de l'extraction de la caractéristique. Par ces derniers, nous nous concentrons ensuite sur l'objectif de concevoir des méthodes robustes d'extraction des courbes, car ils représentent la partie la plus cruciale d'un système de reconnaissance. Le contenu du chapitre trois abordera ce but, par étape, par nos propositions de nouvelles descriptions faciales.

CHAPITRE

3

Reconnaissance de visage basée sur l'apparence locale

La reconnaissance de visage basé sur l'apparence locale est un algorithme générique, pratique, et le plus robuste de reconnaissance de visage qui utilise les représentations des régions locales de visage et les associe au niveau des fonctionnalités qui assure la conservation des relations spatiales. Les idées sous-jacentes pour préférer une approche basée sur l'apparence locale sur une approche holistique basée sur l'apparence sont les suivantes :

- Dans une approche holistique fondée sur l'apparence, une modification d'un local peut affecter l'ensemble de la représentation des entités, alors que dans la reconnaissance de visage basée sur l'apparence, elle affecte uniquement les fonctionnalités qui sont extraites du bloc correspondant tandis que les fonctionnalités qui sont extraites des autres blocs restent inchangés.
- Un algorithme local basé sur l'apparence peut faciliter la pondération des régions locales. Il peut mettre plus de poids dans les régions qui se révèlent être plus discriminantes. En outre, cela peut également améliorer la robustesse contre l'occlusion, en donnant moins de poids aux régions où une occlusion est détectée.

Un diagramme indiquant l'extraction de caractéristiques par représentation d'apparence locale est présentée à la figure 3.1. Dans l'approche, une image de visage détectée et alignée est d'abord divisée en régions locales. Par exemple, dans la figure 3.1, il est fait sans considération de toutes les régions saillantes, telles que les yeux. Ensuite, une transformation peut être utilisée pour représenter les régions faciales locales. Ensuite, les coefficients de représentation extraits de chaque bloc sont combinés afin de fournir le vecteur caractéristique qui représente l'image de face entière.

Dans ce chapitre, d'abord, les approches de représentation qui peuvent être utilisées pour les régions faciales locales sont examinées. Selon les résultats expérimentaux [30], La transformée cosinus discrète (DCT) est la méthode de représentation optimale. Par conséquent, dans la deuxième section visage local basé sur l'apparence Reconnaissance par le DCT est expliquée en détail. Dans la dernière section, une description du méthode de classification plus proche voisin, utilisé comme classificateur dans l'algorithme de reconnaissance faciale.

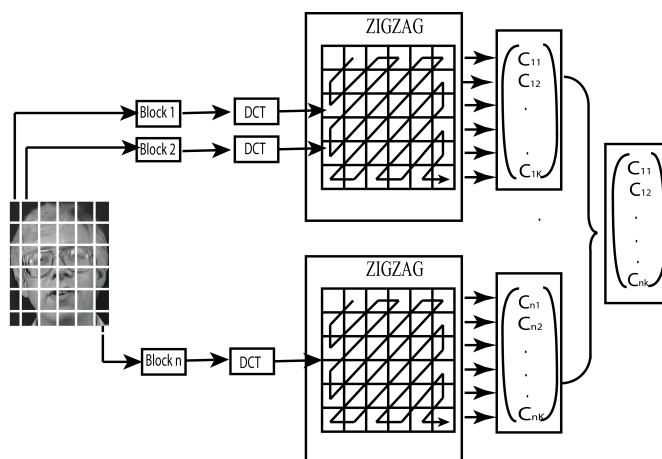


Figure 3.1: Représentation d'apparence locale.

3.1 Méthodes de représentation de l'apparence locale

Dans les sous-sections suivantes, les transformations d'image qui peuvent être utilisées pour représenter l'apparence locale du visage sont brièvement expliquées.

3.1.1 La transformée cosinus discrète

La transformée cosinus discrète (DCT) est un outil d'analyse de signal bien connu utilisé en compression grâce à sa puissance de représentation compacte, qui est supérieure à celle des autres transformations indépendantes d'entrée largement utilisées, par ex. Transformée de Fourier discrète et transformation de Walsh-Hadamard. Bien que La transformation de Karhunen-Loeve (KLT) est connue pour être la transformation optimale en termes d'information, sa nature dépendante des données rend l'utilisation inapplicable de certaines tâches pratiques, ce qui en fait un outil important pour la représentation à la fois en termes d'emballage de l'information et en termes de complexité de calcul en raison de sa nature indépendante des données.

La transformée cosinus discrète 2-D d'un bloc d'image $m \times m$ est définie comme

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{m-1} \sum_{y=0}^{m-1} (f(x, y) * \text{Cos}\left[\frac{(2x+1)\pi u}{2m}\right] \text{Cos}\left[\frac{\pi(2y+1)v}{2m}\right]) \quad (3.1)$$

$\alpha(u)$ et $\alpha(v)$ les facteur d'orthogonalité de la transformée.

$f(x, y)$ la valeur du pixel de l'image a transformer à la position x et y .

ou X et Y , indiquant le pixel

La transformation inverse est définie comme:

$$f(x, y) = \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} \alpha(u)\alpha(v)C(u, v) * \text{Cos}\left[\frac{(2x+1)\pi u}{2m}\right]\text{Cos}\left[\frac{(2y+1)\pi v}{2m}\right] \quad (3.2)$$

pour $u, v = 0, 1, \dots, m - 1$ où

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{m}} & \text{if } u = 0 \\ \sqrt{\frac{2}{m}}, & \text{if } 1 \leq u \leq m - 1 \end{cases} \quad (3.3)$$

Les fonctions de base de DCT sont illustrées à la figure 3.2. Comme on peut le voir à partir de la partie supérieure des fonctions de base et également de l'équation 3.1, la composante $(0, 0)$ représente la valeur d'intensité moyenne de l'image, qui est directement affectée par variations d'éclairage. A partir de la figure, on peut également remarquer que le $(0, 1)$ Et $(1, 0)$ représentent les changements d'intensité verticale et horizontale, respectivement.

Un échantillon de sortie DCT est représenté à la figure 3.3. Comme on peut l'observer, les coefficients qui représentent un degré plus élevé de la capacité de représentation sont situés dans le bloc supérieur gauche de la matrice. Pour construire le vecteur des coefficients DCT 2D, les coefficients sont ordonnés en utilisant le modèle de balayage en zigzag (voir Figure 3.4). De cette façon, les coefficients qui contiennent plus d'information sont préservés lorsque le vecteur est tronqué.

3.1.2 Analyse en composantes principales (ACP)

L'analyse en composantes principales (ACP) est l'une des techniques de réduction de la dimensionnalité les plus connues. Il s'agit d'une méthode non supervisée et elle essaye de trouver le meilleur sous-espace de représentation qui minimise l'erreur de reconstruction. La méthode a été largement utilisée pour la reconnaissance holistique des visages, par exemple pour les eigenfaces[25]. Il peut également être utilisée pour représenter les régions faciales locales [35].

Dans l'analyse en composantes principales, d'abord, le sous-espace qui représente l'entrée des données sont tirées des échantillons d'entraînement disponibles. Soit B_i un $m \times n$ bloc d'image de résolution ou n et m , la largeur et la hauteur de l'image en pixel. En concaténant les lignes ou les colonnes, le bloc d'image bidimensionnel peut être converti en un $N = m \times n$ vecteur dimensionnel y_i . Soit $y_{i,1}, y_{i,2}, y_{i,3}, \dots, Y_{i,K}$ l'ensemble des i^{th} blocs à partir de K images. La moyenne du i^{me} bloc, $y_{i,m}$, est calculée comme suit:

$$y_{i,m} = \frac{1}{K} \sum_{k=1}^K y_{i,k} \quad (3.4)$$

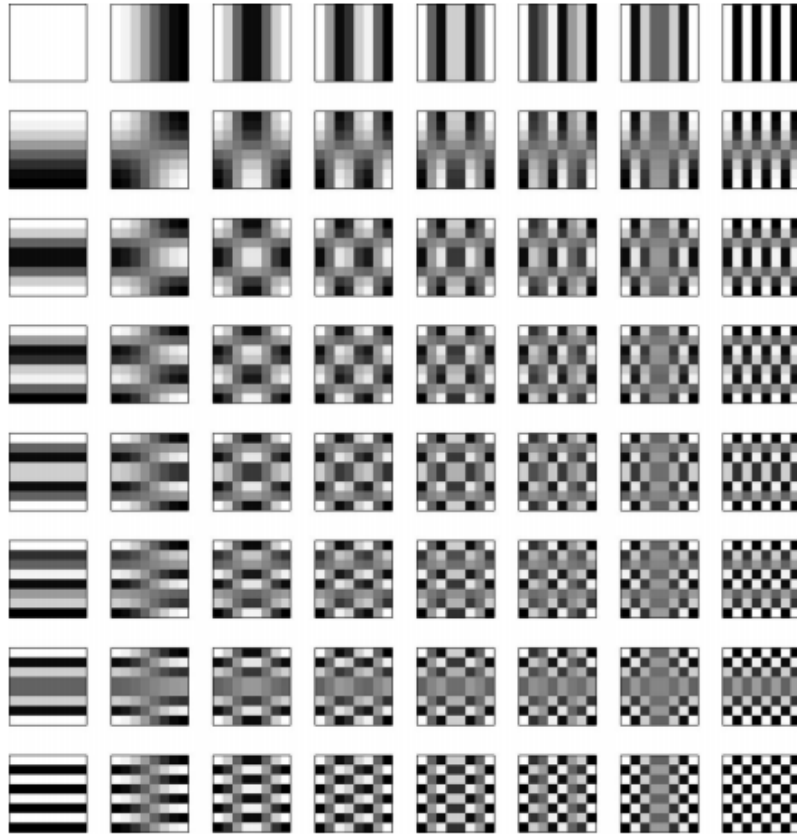


Figure 3.2: Fonctions de base DCT pour $m = 8$.

La différence entre le bloc image d'entraînement et le bloc image de visage moyenne est :

$$\bar{y}_{i,m} = y_{i,k} - y_{i,m} \quad (3.5)$$

La matrice de covariance $N \times N$, C , peut être calculée comme:

$$C = FF^T \quad (3.6)$$

Où F est une matrice $N \times K$, contenant les $\bar{y}_{i,k}$, dans ses colonnes. Si la dimension du bloc d'image est inférieur au nombre d'images d'apprentissage, $N \leq K$, puis Les vecteurs propres de C peuvent être calculés directement. Si la dimension de l'image Bloc est supérieur au nombre d'images d'apprentissage, $K \leq N$, puis premier $K \times K$ La matrice L est calculée

$$L = F^T F \quad (3.7)$$

Les vecteurs propres v_L et les valeurs propres λ_L de L peuvent être calculés comme:

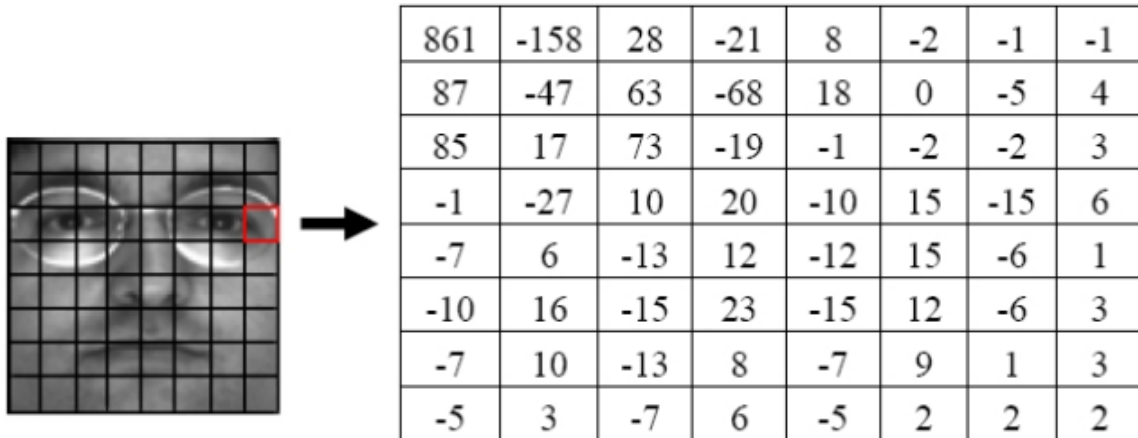


Figure 3.3: Exemple de sortie DCT d'un bloc d'image de visage.

$$Lv_L = \lambda_L v_L \quad (3.8)$$

En remplaçant L par l'équation 3.7

$$F^T F v_L = \lambda_L v_L \quad (3.9)$$

Multiplier les deux côtés par F

$$F F^T F v_L = \lambda_L F v_L \quad (3.10)$$

remplaçant $F F^T$ par C

$$C F v_L = \lambda_L F v_L \quad (3.11)$$

Le bloc d'image de face peut être représenté dans cet espace sans perte d'information, par la somme pondérée de ces vecteurs propres. Cependant, généralement, un petit nombre de vecteurs propres suffit à représenter correctement les blocs d'image de face. La première M vecteurs propres, $M \inf N$, correspondant aux plus hautes M premières valeurs propres, sont choisis pour construire le sous-espace. Les blocs d'image de face sont représentés avec des vecteurs de caractéristique M -dimensionnels r_k en les projetant sur le sous-espace.

$$r_k = V^T \bar{y}_k \quad (3.12)$$

où V est la matrice dimensionnelle $N \times M$ qui contient les vecteurs M propres dans ses colonnes.

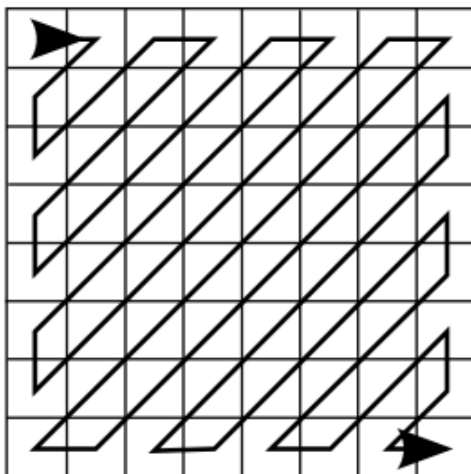


Figure 3.4: Modèle de balayage en zigzag.

3.1.3 Reconnaissance faciale basée sur l'apparence locale utilisant la transformée cosinus discrète

Dans le cadre de l'approche locale de reconnaissance faciale fondée sur l'apparence, l'image de visage détectée et enregistrée est divisée en régions locales. Ensuite, sur chaque local, la DCT est effectuée. Les coefficients DCT obtenus sont ordonnés en utilisant le balayage en zigzag. A partir des coefficients ordonnés, selon la stratégie de sélection d'une caractéristique, M d'entre eux sont sélectionnés et normalisés. Un vecteur caractéristique local M -dimensionnel. Enfin, les vecteurs de caractéristique basés sur DCT extraites de chaque bloc sont concaténées pour construire le vecteur caractéristique global. La classification est faite à l'aide d'un classificateur voisin le plus proche. Comme on peut le voir il ya plusieurs paramètres impliqués dans une telle approche. Ceux-ci sont:

- Normalisation des caractéristiques: Comme les coefficients DCT ont une plage d'amplitude différente et que les blocs locaux avec des niveaux de luminosité différents conduisent aux coefficients DCT avec des niveaux de valeur différents, il est important, contribution de chaque coefficient et de chaque bloc à la classification.
- Taille du bloc: L'application de la DCT sur les grandes régions locales offre une représentativité plus compacte, mais elle fournit une représentation statistique médiocre de la région. Une meilleure représentation statistique peut être fournie en utilisant de petites régions locales, mais cette fois la représentation devient moins compacte.
- Partitionnement d'image: L'image du visage peut être partitionnée en considérant certaines régions saillantes, telles que les yeux, ou en mettant une grille rectangulaire compte tenu de toute région saillante.

- Sélection des caractéristiques: Différents types de variations d'apparence faciale peuvent être Manipulés en utilisant différentes bandes de fréquences. Par conséquent, une sélection automatique de bandes de fréquence est requise pour déterminer les Bande de fréquence à utiliser pour classer l'image de test.

Dans les chapitres suivants, ces points seront expliqués et analysés en détail

3.2 Classification avec KPPV

En raison de la distribution non paramétrique des données faciales et du fait du petit échantillon de la taille disponible pour la formation, la méthode de classification la plus utilisée pour la reconnaissance faciale. Il s'agit d'un algorithme de classification paresseux facile et efficace, où il n'y a pas de travail en cours d'entraînement, comme l'estimation de la densité, et tout le travail est effectué pendant les tests. La classification est faite en comparant un échantillon d'essai avec tous les échantillons de formation dans la base de données et en trouvant l'échantillon de formation qui a la distance la plus proche. Plusieurs métriques de distance peuvent être utilisées pour la classification du plus proche voisin, telle que $L1$, $L2$ les normes et la corrélation normalisée. Le choix de la distance métrique est très importante et la performance de classification peut changer selon la métrique de distance utilisée.

3.3 Expériences

Dans ce chapitre, les paramètres de l'approche proposée locale de reconnaissance de visage à base d'apparence (LAFR) sont analysés au moyen d'expériences approfondies. Le chapitre commence par la description des bases de données de référence et des configurations expérimentales.

Les sections suivantes présentent une évaluation détaillée des effets de la normalisation des caractéristiques de paramètres, la métrique de distance, la taille des blocs, la bande de fréquence, la stratégie de partitionnement d'image et la méthode de représentation sur la performance de classification. Enfin, dans la dernière section, la robustesse de l'algorithme proposé contre la compression est testée.

3.3.1 Bases de données de référence

En raison de l'intérêt énorme pour la recherche de reconnaissance faciale, beaucoup de bases de données de visage ont été publiquement disponibles pour évaluer de manière comparative la performance de reconnaissance de visage. Quatre d'entre elles ont été choisies pour tester la robustesse de l'algorithme de reconnaissance de visage proposé contre l'apparence du visage, les variations provoquées par l'occlusion de la face partielle, l'expression, l'illumination, l'intervalle de temps, et des conditions incontrôlées. Les bases de données utilisées sont:

- La base de données de la grande reconnaissance des visages (FRGC) [45]
- La base de données de face AR [46].
- La base de données Yale face B [47].
- La base de données d'extension de Yale B [48].

Un algorithme local basé sur l'apparence peut faciliter la pondération des régions locales. Il peut mettre plus de poids dans les régions qui se révèlent être plus discriminantes. En outre, cela peut également améliorer la robustesse contre l'occlusion, en donnant moins de poids aux régions où une occlusion est détectée.

3.4 Conclusion

Dans ce chapitre, nous avons proposé un aperçu des ensembles de données et des configurations théorique et pratique de modèle binaire local (LBP), les diverses méthodes de LBP dans l'analyse de texture et classification, l'analyse de visage et la reconnaissance, et les autres applications de détection sont réexaminées. Le principe de méthode de LBP est brièvement discuté, qui analyse principalement l'opération de seuil, le modèle et la rotation l'uniforme modèle invariant dans la méthode de LBP. Deuxièmement, l'analyse de texture et classification de la méthode de LBP, l'analyse de visage et la reconnaissance de la méthode de LBP et les autres applications de détection de la méthode de LBP est particulier peigné et commenté. Enfin, les problèmes importants actuels de la méthode de LBP sont analysés et l'avenir pour la méthode de LBP est fait remarquer.[23]

CHAPITRE

4

Résultats et simulations

4.1 Introduction

Le modèle binaire local a été appliqué dans de nombreuses applications, mais la plupart des systèmes LBP ne fonctionnent que dans un seul canal d'image d'une seule résolution. Par conséquent, l'opérateur original LBP a les limitations suivantes dans ses applications. Tout d'abord, les caractéristiques calculées dans une seule échelle capturant la structure d'image uniquement à une résolution particulière peuvent ne pas nécessairement détecter les caractéristiques de texture dominantes, et elles peuvent également ne pas être robustes à la traduction et à la rotation de l'image. Deuxièmement, les fonctionnalités extraites dans un seul canal, tel que l'espace d'images en niveaux de gris, ne capturent que les informations d'intensité mono-chromatique qui peuvent limiter la performance de reconnaissance.

Dans ce chapitre, des nouvelles représentations sont proposées pour étendre le LBP afin de fournir un outil pour l'analyse des faces. Les méthodes LBP résultantes apportent une contribution à la PCA et à diverses méthodes de fusion de classificateurs pour la reconnaissance faciale. La configuration expérimentale est ensuite introduite ainsi que résultats obtenus sur la base de données XM2 VTS, FERET et FRGC 2.0.

4.2 Authotification de visage 2D

4.2.1 Problème d'identification de visage: Base de données FERET

Cette expérience a permis d'appliquer l'ensemble de formation standard CSU pour estimer les paramètres des méthodes d'apprentissage supervisées. Dans ce test, le taux de

reconnaissance au rang 1 et deux mesures statistiques sont utilisées pour comparer les performances des méthodes. Les mesures sont le taux de reconnaissance moyen avec un intervalle de confiance de 95% et la probabilité que l'algorithme surpasse un autre. La probabilité est notée $P(Alg1 \geq Alg2)$ et elle est calculée en permutant les ensembles de galeries et de sondes. Les résultats obtenus avec PCA, DCT et SVD sont présentés et comparés entre eux.

Le résultat de la méthode des histogrammes régionaux $LBP_{8,2}^{u2}$ avec les mesures de similarité, $LBP_{8,2}^{u2}$ histogrammes régionaux projetés sur l'espace PCA pour la corrélation normalisée (LBPH + PCA) et notre méthode proposée qui est une combinaison des méthodes (MLBPH + PCA + SVM) avec différentes régions $k \times k$ sont représentés graphiquement à la figure 4.1. Les résultats montrent que la mesure du log vraisemblance ($LBP_{KL} + SVM$) est clairement pire lorsque k est supérieur à 2. Il est très difficile de dire que l'une des mesures de similarité de l'histogramme, c'est-à-dire $LBP_{HI} + SVM$, $LBP_{Chi} + LBP_{JS} + SVM$, à l'exception de la mesure log vraisemblance, serait la meilleure, mais l'intersection de l'histogramme se comporte légèrement mieux que d'autres en termes de taux de reconnaissance et de temps de calcul.

En comparant le taux de reconnaissance moyen avec les méthodes de similarité dans LBPH et LBPH avec PCA, il est clair que l'application de PCA à la représentation engendrée par des histogrammes régionaux de modèles uniformes améliore les performances. Comme prévu pour les méthodes fondées sur l'histogramme LBP, la vitesse de reconnaissance moyenne est réduite à mesure que la taille de la fenêtre augmente en raison de la perte de l'information spatiale, mais pour notre méthode, le taux de reconnaissance moyen est robuste pour une large plage de valeurs de k ($16 \geq k > 3$) régions. Par exemple, le taux de reconnaissance moyen avec $k = 3$ est 72,1%, alors que pour $k = 11$ est 91,4%. En présence des inexactitudes de localisation de visage, la performance du processus. En d'autres termes, la modification du nombre de régions, k , affecte uniquement la longueur du vecteur de caractéristiques et le temps de calcul.

Dans le tableau 4.1, le paramètre k des méthodes LBP est optimisé en maximisant le taux moyen de reconnaissance. LBP avec PCA mettent clairement en évidence les autres dans tous les tests statistiques et tous les ensembles de probes, sauf pour DUP2. Comme le montre le tableau 4.1. En comparant MLBPH et LBPH avec PCA, la précision n'est pas significativement différente, mais MLBPH + PCA + SVM est légèrement plus faible que (MLBPH + PCA + SVM) > (LBPH + PCA + SVM). Les résultats des méthodes MLBPH sur l'ensemble F_c sont entre 24% et 58%. L'une des raisons pour lesquelles la mauvaise performance des méthodes MLBPH sur l'ensemble F_c est que la dimensionnalité du descripteur de visage dans MLBPH est élevée et la seconde est la façon dont les motifs sont regroupés pour former un histogramme. Enfin, l'avantage principal d'utiliser LBP, c'est que toute transformation monotonique ne peut être exploitée.

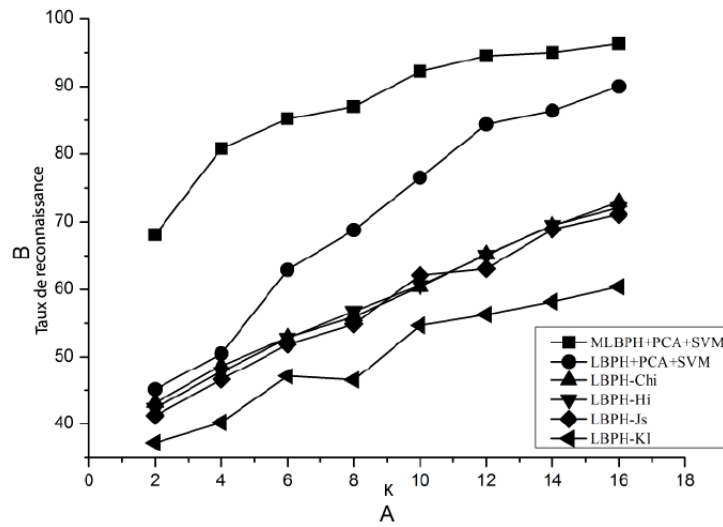


Figure 4.1: Taux moyen de reconnaissance de six méthodes LBP par rapport au nombre ($k \times k$) de régions.

Tableau 4.1: Comparaisons sur les ensembles de sondes et la vitesse de reconnaissance moyenne sur la base de données FERET.

	K	Fb	Fc	Dup1	Dup2	Lower	Mean	Upper
MLBPH+PCA+SVM	13	0.989	0.577	0.717	0.487	0.838	0.879	0.919
$MLBPH_{Chi} + SVM$	13	0.928	0.294	0.630	0.513	0.725	0.774	0.829
$MLBPH_{Hi} + SVM$	13	0.956	0.325	0.641	0.530	0.731	0.779	0.825
$MLBPH_{Kl} + SVM$	13	0.895	0.247	0.605	0.487	0.700	0.752	0.800
$MLBPH_{Js} + SVM$	13	0.920	0.284	0.625	0.509	0.725	0.771	0.819
LBPH+PCA+SVM	13	0.979	0.686	0.711	0.483	0.819	0.861	0.900
$LBPH_{Chi} + SVM$	13	0.961	0.526	0.641	0.487	0.744	0.792	0.838
$LBPH_{Hi} + SVM$	13	0.967	0.552	0.651	0.513	0.750	0.797	0.844
$LBPH_{Kl} + SVM$	16	0.943	0.541	0.625	0.449	0.706	0.754	0.800
$LBPH_{Js} + SVM$	13	0.959	0.521	0.637	0.479	0.744	0.79	0.838

4.2.2 Simulation sur la base de données XM2VTS

Dans cet essai, les histogrammes régionaux $LBP_{8,2}^{u2}$ monochromes projetés sur PCA avec corrélation normalisée (LBPH + PCA + SVM), $LBP_{8,2}^{u2}$ opposé histogrammes régionaux projetés sur PCA avec corrélation normalisée (Opp-LBPH + PCA + SVM) et notre méthode proposée (MSLBPH + PCA + SVM) avec un nombre différent de régions ($k \times k$) sont représentées graphiquement sur la figure 4.2. Il est clair que la performance de ces méthodes en utilisant la taille de région maximale possible, $k = 1$, est pire que les autres, ce qui signifie que plus la taille de la région (plus grand k) est importante, plus la performance peut être meilleure. En comparant notre méthode proposée (MSLBPH + PCA + SVM) avec d'autres, il s'ensuit que la combinaison des caractéristiques d'histogramme régional LBP monochrome et opposée améliore le résultat où le meilleur taux d'erreur total dans l'ensemble de tests de mon LBPH + PCA + SVM à $k = 9$ est de 1,291%, Opp LBPH + PCA + SVM à $k = 16$ est 1,141% et notre méthode proposée (MSLBP).

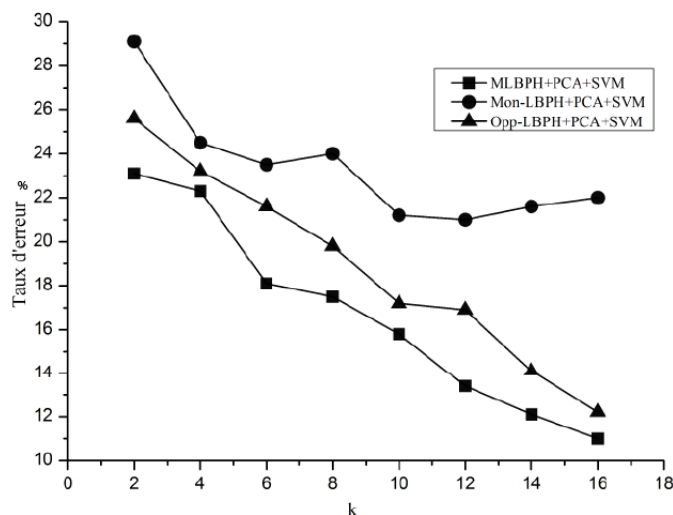


Figure 4.2: Taux d'erreur total sur l'ensemble de tests dans la configuration I pour les méthodes LBP à trois couleurs contre les régions $k \times k$.

Comparaison avec d'autres méthodes sur le XM2VTS.

Dans le protocole pour la base de données XM2 VTS, l'erreur totale, TER, qui est la somme du taux de faux rejet et du taux de fausses acceptations, est utilisée pour rapporter la performance des méthodes. Dans cette expérience, nous mettons en œuvre et comparons la méthode monochrome $LBP_{8,2}^{u2}$ méthode (mon LBPH + PCA + SVM), Adversaire $LBP_{8,2}$ méthode (Opp LBPH + PCA + SVM) et notre approche proposée (MSLBPH, MLBPH et LBPH + PCA) avec différentes méthodes de fusion et différentes

méthodes de similarité d'histogramme mentionnées dans la section 5.3 avec le classificateur Adaboost pour LBPH [85] (LBPH + Adaboost). Rodriguez a constaté que le taux d'erreur total de LBPH-Adaboost donnant 7,88 sur l'ensemble-teste est pire que celui de LBPH Chi + SVM, à savoir 6,8. Néanmoins, nous avons constaté que le taux d'erreur de LBPH-Adaboost peut être réduit à 5,263 si 300 histogrammes régionaux (caractéristiques) sont utilisés. Le tableau 4.2 présente les résultats comparatifs des méthodes.

Tableau 4.2: Taux d'erreur totale.

	K	Eval Set	Test set
MSLBPH+PCA+SVM	8	0.978	0.533
<i>MSLBPH_{Chi} + SVM</i>	13	0.934	0.221
<i>MSLBPH_{Hi} + SVM</i>	13	0.987	0.353
<i>MSLBPH_{Kl} + SVM</i>	13	0.830	0.245
<i>MSLBPH_{Js} + SVM</i>	13	0,930	0.281
Mon-MLBPH+PCA+SVM	8	0.984	0.565
<i>opp - MLBPH_{Chi} + SVM</i>	13	0.934	0.225
MLBPH+PCA+SVM	8	0.976	0.545
<i>MLBPH_{Chi} + SVM</i>	13	0.938	0.276
<i>MLBPH_{Hi} + SVM</i>	13	0.945	0.304
<i>MLBPH_{Kl} + SVM</i>	13	0.848	0.245
<i>MLBPH_{Js} + SVM</i>	13	0,870	0.293

La différence entre MLBPH et MSLBPH est que MLBPH est un descripteur de texture multirésolution régionale et MSLBPH est un descripteur de texture de couleur régionale capturant l'information de texture dans chaque spectre et dans le spectre croisé. Dans cette expérience, le taux d'erreur total fourni par les méthodes MSLBPH est meilleur que celui obtenu par les méthodes MLBPH, comme le montre le tableau 4.2. Ces résultats indiquent que la texture de la couleur et/ou l'information de spectre croisé est essentielle pour la reconnaissance de visage pour obtenir une grande précision. En comparant les résultats de fusion montrés à la figure 4.2, les taux d'erreurs totaux produits par la méthode de la machine à vecteur de support (SVM), en particulier $k \sup 11$. Autrement dit, l'information spatiale locale exploitée par nos méthodes proposées est très importante.

En comparant le taux d'erreur total des méthodes de similarité d'histogramme et la méthode de similarité dans l'espace LDA, il est clair que l'application de la PCA à la représentation améliore la performance. Pour les méthodes LBPH, le taux d'erreur total

est augmenté ainsi que la taille de la fenêtre augmente (c'est-à-dire k décroît) en raison de la perte de l'information spatiale. Cependant, pour notre méthode MLBPH, le taux d'erreur total est robuste pour une large gamme de $16 > k > 2$ régions. Contrairement aux autres méthodes, notre approche basée sur MSLBPH obtient meilleures performances et le meilleur résultat global est obtenu avec le classificateur de fusion SVM, pour lequel le taux d'erreur total sur l'évaluation et le test pour la configuration I est de 0,29 et de 0,28 respectivement.

4.2.3 Simulation sur la base de données FRGC 2.0

En FRGC2.0, le taux de vérification à 0,1 de taux d'acceptation fausse (FAR) atteint par nos méthodes proposées (MLBPH + PCA + SVM et MSLBPH + PCA + SVM) avec un nombre différent de $k \times k$ Régions. Pour la méthode SVM, le taux de vérification est évalué pour $k > 6$ par ce que la performance montre que la précision s'améliore lorsque k devient plus grand. En comparaison avec la méthode MLBPH et MSLBPH, plus la taille de la région est petite, plus la performance MSLBPH est élevée. De plus, en comparant les méthodes de fusion, la performance de SVM est meilleure que celle de la règle de somme. Ces deux observations sont cohérentes avec les résultats présentés. Le Tableau 4.3 présente les résultats comparatifs des méthodes ci-dessus avec l'histogramme LBP en bloc à échelles multiples (MBLBPH) + Adaboost [49], LBP Chi [49], LBP + AdaBoost [49] et les méthodes de référence. Il montre que les méthodes basées sur le LBP fonctionnent beaucoup mieux que la ligne de base. Par rapport aux méthodes de LBP, nos versions avancées, telles que MLBPH, MSLBPH et MBLBPH, sont supérieures, car ces descripteurs captent des informations plus discriminatives. La performance de la méthode MBLBPH est similaire à la méthode MLBPH. La principale différence entre MBLBP et MLBPH est que le premier est basé sur le calcul des valeurs moyennes des sous-régions. Par conséquent, MBLBP capture les informations locales, mais il est difficile d'extraire l'énergie de contraste dans de petites régions sur de grandes distances, alors que notre méthode peut capturer les informations locales et non locales. Il a été prouvé que l'information non locale était importante pour la détection des visages dans des conditions d'éclairages différentes, comme cela a été expliqué dans la section 3.1. Cependant, nos méthodes proposées ne sont pas meilleures que MBLBPH + Ada Boost. Une des raisons possibles est la façon dont les codes LBP sont regroupés en histogramme. Notre méthode proposée exploite le concept de motif uniforme pour grouper des motifs binaires. Cependant, les motifs binaires locaux uniformes LBP_{u2} ne sont pas les motifs principaux pour représenter l'image de face lorsque le rayon de l'opérateur LBP est supérieur à 4. En appliquant l'approche LBP^{u2} , la dimension de l'histogramme (caractéristiques) est réduite, mais l'information captée par l'opérateur LBP de grands rayons est perdue. Par conséquent, la performance est dégradée, en particulier lorsque les échantillons de visage sont capturés dans des conditions d'illumination incontrôlée et également utilisée pour la formation.

Tableau 4.3: Taux de vérification en% à 0,1% FAR pour différentes méthodes sur FRGC 2.0

	K	visage1	visage2	visage3
MSLBPH+PCA+SVM	11	97.90	94.53	94.04
MSLBPH+PCA+SVM	16	93.43	93.34	92.83
MSLBPH+PCA+SVM	7	95.24	92.32	94.51
<i>MLBPH + AdaBoost</i>	13	98.35	98.82	95.24
<i>LBP_{Chi} + SUM</i>	13	84.17	80.35	76.67
<i>LBPH + AdaBoost</i>	13	82.72	78.53	74.78

4.3 Conclusion

Dans cette section, deux descripteurs discriminatifs ont été proposés pour la reconnaissance faciale. Le premier descripteur contenant les informations d'une analyse multi-résolution de l'image de face, appelée MLBPH + PCA + SVM, a été montré pour fournir un système très robuste qui est relativement insensible aux erreurs de localisation, car il bénéficie de la multi-résolution informations recueillies à partir de l'histogramme régional. Le système a été mis en œuvre pour faciliter l'identification et la vérification des faces en utilisant les bases de données FERET, XM2 VTS et FRGC 2.0, puis évalué à l'aide de leurs protocoles standard. Dans l'identification faciale effectuée sur la base de données FERET, les résultats expérimentaux obtenus montrent que le taux de reconnaissance moyen de 88% avec un intervalle de confiance de 95%, délivré par notre méthode surpassent les autres prétendants à la pointe de la technique. En particulier, notre système a obtenu le meilleur résultat global de 97.90% de taux de reconnaissance dans l'expérience impliquant diverses sondes d'expression faciale set (FB set) tout en fournissant des résultats comparatifs à d'autres méthodes basées sur LBP pour d'autres séries de sondes. Les méthodes MLBPH ont été testées en mode vérification sur les bases de données XM2 VTS et FRGC 2.0, notre système MLBPH + PCA + SVM a obtenu le huitième meilleur résultat, soit $TER = 1,66\%$, parmi les 28 concurrents, tandis que le MLBPH + PCA + SUM a obtenu le neuvième meilleur résultat, $TER = 2,04\%$. Dans la configuration II, MLBPH + PCA + SVM a obtenu le sixième meilleur résultat, $TER = 1,13\%$ parmi les 22 concurrents et MLBPH + PCA + SVM a obtenu le quatrième meilleur résultat, $TER = 0,88\%$. En FRGC2.0 Experiment 1, MLBPH + PCA a atteint le taux de vérification de 94,73% dans Mask, 93,21% en Mask et 93,76% en Mask à 0,1% FAR, ce qui nettement surperforme d'autres méthodes traditionnelles LBP. Ces résultats montrent que la méthode (MLBPH + PCA + SVM) peut non seulement être comparable aux méthodes de référence de pointe sur des faces annotées manuellement, mais elle est également plus robuste en présence d'erreurs de localisation.

Le second descripteur discriminatif proposé exploite l'information de texture de couleur

de l'image de face. Le descripteur est formé en projetant l'image de face locale acquise par les opérateurs LBP multispectrale, dans l'espace PCA. Le score global de similarité est obtenu en fusionnant les scores de similarité des descripteurs régionaux. Cette méthode a été mise en œuvre et comparée à d'autres benchmarks bien connus dans la vérification de visage à l'aide des bases de données XM2VTS et FRGC 2.0. Dans la configuration I de XM2VTS, les résultats expérimentaux ont montré que MSLBPH + PCA + SVM a obtenu le meilleur résultat, TER = 0,28%, soit trois fois mieux que le meilleur résultat (c'est-à-dire TER = 0,96% [58]). Dans la configuration II de XM2VTS, MSLBPH + PCA + SVM a également obtenu le meilleur résultat, TER = 0,21%, soit deux fois mieux que le résultat le plus connu (c'est-à-dire TER = 0,51% [58]) rapporté dans les littératures. Ces résultats montrent clairement que MSLBPH + PCA.

Bien que MBLBPH soit similaire à MLBPH, nos méthodes proposées ne sont pas meilleures que MBLBPH + AdaBoost. Une des raisons possibles est la façon dont nous regroupons les codes LBP en histogramme. Notre méthode proposée se concentre sur des modèles uniformes comme base de regroupement de motifs binaires. Cependant, les motifs binaires locaux uniformes LBP^{u2} ne sont pas les modèles dominants pour représenter l'image de face lorsque le rayon de l'opérateur LBP est supérieur à 4. En appliquant l'approche LBP^{u2} , la dimension de l'histogramme (caractéristique) est réduite, mais l'information véhiculée par le grand rayon de l'opérateur LBP est perdue. Par conséquent, la performance est dégradée, en particulier lorsque les échantillons de visage capturés dans des conditions d'illumination non contrôlées sont également utilisés pour la formation. Ce problème sera étudié plus en détail dans le chapitre suivant.

4.4 Reconnaissance faciale avec ELBP

Dans l'application de l'ELBP pour la reconnaissance faciale, nous utilisons une méthode avancée qui utilise la fonction de distance cosinus négative pour la classification et WPCA pour la réduction de la dimension, avec la structure d'appariement des modèles, nous utilisons LBP, ELBP (h) (moyen seulement ELBP horizontal est utilisé), ELBP (h + v) (signifie une paire symétrique à la fois horizontale et ELBPs verticale sont utilisés) pour indiquer la méthode d'extraction de caractéristiques correspondante tandis que leurs méthodes équivalentes de WPCA sont nommées en ajoutant le mot "WPCA" comme suffixe. Les taux de reconnaissance obtenus sur les bases de données AR, FERET et SCface sont comparés à d'autres systèmes FR de l'état de l'art.

4.4.1 Résultats sur la base de données AR

Il convient de rappeler que les ensembles de sondes de la base de données AR, numérotés de 1 à 12 dans le tableau 4.4, sont constitués d'images capturées sous diverses conditions d'expression faciale (Smile, Angers, Scream), des changements d'éclairage (Neutre + Lumière droite, neutre + lumière des deux côtés), et occlusions (lunettes de soleil, lunettes

de soleil + lumière de gauche, lunettes de soleil + lumière droite, écharpe, tard + lumière gauche, Scard + lumière droite).

Les résultats des méthodes basées sur le LBP et ELBP sont fournis dans le tableau 4.5 tandis que la comparaison entre le WPCA/ ELBP (h + v) et d'autres systèmes bien connus sont présentés dans le tableau 4.4. Pour clarifier le but, nous visualisons les RR des systèmes basés sur LBP et ELBP comme diagrammes de dispersion dans la Fig. 4.4.

Tableau 4.4: Rang-1 TR (%) comparaison entre les méthodes basées sur la LBP et l'ELBP sur AR base de données

	Testé/Méthode	Ensemble de sonde								Avg
		1	2	3	4	5	6	7	8	
Exp1	LBP	100	100	74.8	98	98	100	81.5	53.2	83.675
	LBP WPCA	100	98	72.3	96	97	97	81.5	65.7	85.625
	ELBP(h)	100	100	76.6	100	100	100	84.9	59.1	86.275
	ELBP(h) WPCA	100	100	80.1	100	100	100	88.6	70.7	91.45
	ELBP(h + v)	100	100	83	100	100	100	89.3	55.2	87.25
	ELBP(h + v) WPCA	100	98	87.5	100	100	100	91.3	68.2	93.1
Exp2	LBP	100	100	75,8	98	98	100	84,5	53,2	84,425
	LBP WPCA	100	98	79,3	96	97	98	84,5	70,7	88,125
	ELBP(h)	100	100	73,6	100	100	100	86,9	60,1	86,275
	ELBP(h) WPCA	100	100	81,1	100	98	100	87,6	70,7	91,45
	ELBP(h+v)	100	98	84	100	100	100	90,3	56,7	87,375
	ELBP(h+v) WPCA	100	98	87,5	100	98	100	92,3	68,2	92,85

D'après les résultats du tableau 4.4 et des figures 4.3 et 4.5, nous pouvons conclure que:

- Les résultats des méthodes basées sur le LBP et ELBP sont fournis dans le tableau 4.1 tandis que la comparaison entre le WPCA/ ELBP (h + v) et d'autres systèmes bien connus sont présentés dans le tableau 4.4. Pour clarifier le but, nous visualisons les RR des systèmes basés sur LBP et ELBP comme diagrammes de dispersion dans la Fig. 4.4.
- La fusion dans le domaine caractéristique des ELBP horizontaux et verticaux dans l'extraction des caractéristiques (ELBP (h + v)) donne de meilleures performances que l'utilisation d'ELBP (ELBP (h)). Encore une fois, cette amélioration est constante, car elle est réalisée dans toutes les expériences avec des méthodes de matrice de concordance et de FR basées sur WPCA. Nous pensons que cela provient de caractéristiques de texture horizontale et verticale utiles qui sont extraites par la

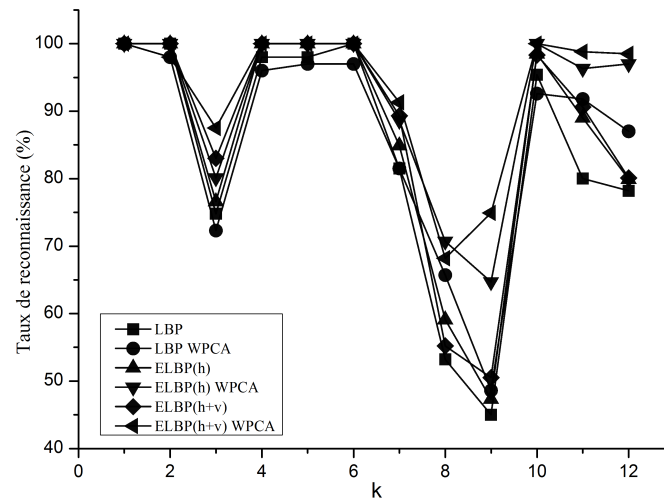


Figure 4.3: Expérience 1 (Taux de reconnaissance ELBPH(h v) contre cinq méthodes de normalisation différentes en fonction de k.)

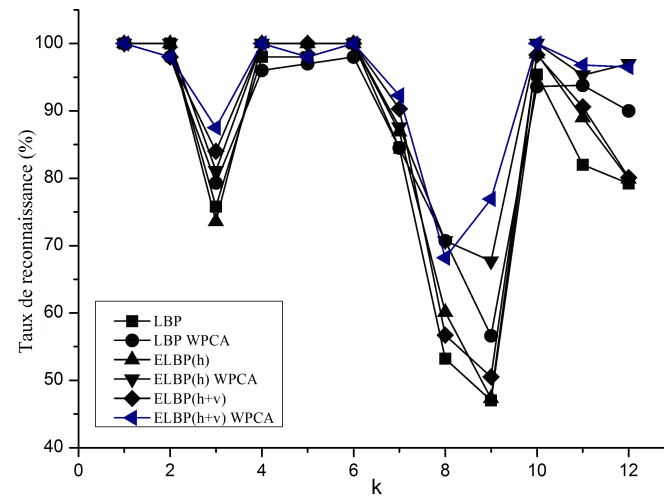


Figure 4.4: Expérience 2 (Taux de reconnaissance ELBPH(h v) contre cinq méthodes de normalisation différentes en fonction de k.)

combinaison de la paire symétrique d'ELBPS. Ces résultats indiquent également que la méthode ELBP (h + v) est forte contre les expressions faciales, les variations d'éclairage et les occlusions.

- Entre les modèles d'appariement des modèles et les cadres basés sur le WPCA, ce dernier est nettement plus efficace que le précédent. Dans tous les ensembles de sondes de quatre expériences et avec toutes les méthodes d'extraction de caractéristiques, l'utilisation de WPCA apporte des taux de reconnaissance plus élevés que la méthode d'appariement direct en utilisant la distance Chi square. Comme nous l'avons mentionné, cette amélioration est obtenue puisque WPCA produit des vecteurs caractéristiques plus discriminants en éliminant les informations redondantes de leurs formes d'origine. Il est important de noter que ces améliorations confirment que si une méthode d'extraction de caractéristiques est plus puissante qu'une autre dans le cadre d'appariement des modèles, les résultats correspondants obtenus avec le paradigme basé sur WPCA seront plus élevés que ceux de la méthode comparée.

Bien que les expériences sur la base de données AR d'autres méthodes dans la littérature utilisée habituellement quelques ensembles de sondes seulement, leurs résultats étaient généralement compatibles avec ceux de notre Exp 1, Exp 2. Par conséquent, nous rapportons que les résultats obtenus étaient communément compatibles avec ceux de notre Exp 1, Exp 2. Par conséquent, nous rapportons dans le tableau 4.4 les résultats de comparaisons les plus représentatifs entre ELBP (h + v) WPCA et d'autres systèmes.

Tableau 4.5: Rang-1 RRs (%) comparaison avec d'autres systèmes contemporains sur la base de données AR utilisant la même méthode d'évaluation

	1	2	3	7	10	
Méthode	Sourire	Colère	Crier	lunettes	Écharpe	Classes ¹
Exp 1						
S-LNMF[90]	96	50	49	84	87	100
LGBP[135]	89	65	70	80	98	50
IRF[141]	82.4	77	50	87	91	120
String Face[21]	87.5	87.5	25.9	88.0	96.0	100
Sparse Coding[129]	79.2	65.1	72.2	94.2	91.2	100
DMMA[78]	99	93	69	76	76	100
Notre méthode	100	100	88.5	92.2	100	100

Il se dégage du tableau 4.5 que notre système est comparable à d'autres méthodes à la pointe de la technique. ELBP(h + v) WPCA est le seul système qui a des taux de reconnaissance sur les ensembles de sondes Smile, Angers et Scarf dans Exp 1. alors que sur les ensembles de sondes Scream et Sun Grasse d'Exp 1, notre méthode surpasse également d'autres rivaux avec des taux de reconnaissance plus élevés. Ces résultats sont plus intéressants lorsque l'on considère qu'ils sont obtenus avec un plus grand nombre

d'images de galerie / sondent par rapport à la plupart des autres méthodes, car la tâche de reconnaissance devient plus difficile quand il y a plus de sujets impliqués.

4.4.2 Résultats obtenus sur la base de données FERET

Cette section donne les RR de rang 1 des systèmes fondés sur ELBP par rapport à LBP et d'autres wavelets de Gabor basés sur la base de données FERET. Nous rapportons les résultats comparatifs du protocole standard dans le tableau 4.5.

Tableau 4.6: Rang-1 RRs (%) comparaison avec d'autres résultats de l'état de l'art sur la base de données FERET en utilisant l'évaluation standard Protocole

Method	Fb	Fc	Dup1	Dup2	Average
LBP	96.8	92.5	71.1	68.6	82.25
ELBP(h)	97	95	71.8	70.2	83.5
ELBP(h + v)	97.9	96	73	71.3	84.55
LBP WPCA	98.7	99.0	83.9	78.2	92.1
GEWC [29]	96.3	99.5	78.8	77.8	89.3
LGBPHS [135]	98.0	97.0	74.0	71.0	87.8
DMMA [78]	98.1	98.5	81.6	83.2	91.6
LGBP [88]	98.1	98.9	83.8	81.6	92.1
FGLBP [113]	98.0	98.0	90.0	85.0	94.2
ELBP(h) WPCA	99.3	99.0	87.7	83.8	94.2
CHG [22]	97.5	98.5	85.6	84.6	92.6
DLBP [82]	99.0	99.0	86.0	85.5	93.6
ELBP(h+v) WPCA	99.5	100	90.1	88.9	94.62

Les résultats de la comparaison dans le tableau 4.6 confirment que l'ELBP horizontal est plus robuste que le LBP dans l'extraction de caractéristiques micros faciales (en utilisant des méthodes de couplage de modèles et de WPCA), en particulier dans l'ensemble de sondes Dup 2. il est évident que l'utilisation des ELBP horizontaux et verticaux apporte une amélioration très impressionnante des taux de reconnaissance par rapport au LBP original et au ELBP horizontal unique (les cas d'amélioration les plus significatifs sont dans l'état de vieillissement. Une fois, avec chaque méthode utilisée pour l'extraction de caractéristiques (LBP, ELBP (h) et ELBP (h + v)), le cadre fondé sur WPCA fournit des précisions plus élevées que le modèle correspondant à un jeu de sonde Fc d'ELBP (h + v) WPCA illustre l'efficacité de l'ELBP sous des variations d'éclairage.

4.4.3 Résultats sur la base de données SCface

Cette section donne des résultats de méthodes fondées sur WPCA utilisant LBP, ELBP (h) et ELBP (h + v) pour l'extraction de caractéristiques. Puisque les expériences sont évaluées avec des méthodes fondées sur WPCA, nous nommons chaque méthode en fonction de l'algorithme d'extraction de caractéristiques qu'il emploie.

Les résultats de la reconnaissance du tableau 3.6 montrent que le cadre WPCA ELBP (h + v) surpasse d'autres systèmes à la pointe de la technique, en particulier par rapport à la PCA de base [42] (notre résultat moyen dans l'expérience DayTime est environ neuf fois plus élevé que celui signalé dans [42]). Ces résultats (tableau 3.6) prouvent également que le descripteur ELBP horizontal est plus robuste que le LBP dans l'extraction des caractéristiques micros faciaux (dans les conditions de temps de jour et de nuit à trois distances) et encore (comme dans les évaluations sur AR et FERET présentées dans les sections précédentes), la combinaison de l'ELBP horizontal et vertical apporte la meilleure performance. Au mieux de nos connaissances, nos résultats sur la base de données SCface sont les premiers résultats complets et les plus élevés rapportés dans la littérature jusqu'à présent.

Il est clair que les résultats sur la base de données SCface ne sont beaucoup plus bas que les taux de reconnaissance sur la base de données AR et sur la base FERET. La très faible résolution (petite taille et très mauvaise qualité) des images de sonde dans la base de données SCface est la cause de ces résultats. En plus de cela, cela est compréhensible puisque ELBP est un descripteur élémentaire. Pour obtenir une performance de reconnaissance plus élevée, des représentations faciales plus puissantes sont nécessaires.

Paramètres de ELBP

Le LBP [3] original pour la reconnaissance faciale utilise les opérateurs $LBP_{8,1}$ et $LBP_{8,2}$ sur les sous-régions 7×7 des images d'entrée (résolution 128×128) pour obtenir les meilleures performances. Nos meilleurs résultats sur la base de données AR utilisent $LBP_{8,5}$ (sous-régions 9×9), $ELBP_{8,5,3}$ et $ELBP_{8,3,5}$ (sous-régions 9×9). Les sous-régions $LBP_{8,5}$ (sous-régions 9×9), $ELBP_{8,5,3}$ et $ELBP_{8,3,5}$ (sous-régions 9×9) sont utilisées avec la base de données FERET. Sur la base de données SCface, les $LBP_{8,3}$ (sous-régions 6×6), $ELBP_{8,3,5}$ et $ELBP_{8,5,3}$ (sous-régions 6×6) donnent les taux de reconnaissance les plus élevés. Toutes ces informations sur les paramètres d'ELBP indiquent que le meilleur rapport entre le rayon horizontal et le rayon vertical d'ELBP est de 1,67. Cependant, nous pensons que les rayons des échantillons d'ellipse utilisés par ELBP dépendent probablement du type d'image qu'elle traite (dans ce chapitre, ELBP ne fonctionne que sur l'image d'intensité) et quand elle est utilisé sur des images d'amplitudes de bord orientées accumulées. En outre, le nombre de sous-régions est définitivement compté sur la taille de l'image et il décide de la longueur du vecteur caractéristique résultant.

Avec chaque méthode d'extraction de caractéristiques présentée dans cette dissertation, nous essayons d'abord de trouver et d'assigner certains paramètres de base à des valeurs

fixe, par exemple, les rayons d'ELBP dans cette partie, par des expériences empiriques. Ensuite, ces valeurs sont utilisées uniformément dans 3 bases de données (AR, FERET et SCface) pour signaler leurs résultats correspondants. D'autres paramètres dépendants (sur la taille de l'image ou le type d'image utilisé) peuvent varier avec les explications appropriées. De cette façon, nous souhaitons montrer et mettre en évidence la vraie fiabilité de chaque méthode, qui découle constamment de son efficacité inhérente, plutôt que de présenter les meilleures RR obtenues après avoir effectué une recherche épuisée pour les valeurs des paramètres d'ajustement d'une base de données spécifique ou expérience.

Coût computationnel

Toutes nos expériences sont implémentées par le MATLAB R2010b sur un PC avec 2.10 GHz Intel Core(TM) 2 Duo CPU et une RAM de 2 Go.

Tableau 4.7: Temps de calcul de ELBP par rapport à d'autres méthodes d'extraction de caractéristiques

Méthode	Taille de l'image	Temps(s)	Temps d'extraction (ms)	Image/s
ELBP(h)	128 x 128	4.32	2.97	260
ELBP(h+v)	128 x 128	6.78	5.04	160
Monogenic [128]	150 x 130	19.81	16.56	60
Gabor wavelets	80 x 88	96.23	80.46	12

Lorsque l'on envisage de déployer un système FR dans des situations réelles, son efficacité (haute précision) n'est pas la seule condition préalable. Dans de nombreux scénarios, en particulier la reconnaissance des visages, la vitesse de traitement, joue un rôle essentiel. Pour étudier le coût de calcul des méthodes d'extraction de caractéristiques présentées dans cette thèse, nous avons procédé à quelques tests de référence en exécutant son implémentation Matlab sur l'ensemble de Fa (1196 images) de la base de données FERET et en comparant les métriques résultantes (temps total requis, et sa vitesse, qui est mesurée en nombre d'images par seconde) avec celles de l'étape initiale des ondelettes de Gabor.

Pour être juste, nous avons utilisé des images de résolution de 80 x 88 pour générer des images d'ondelettes de Gabor comme dans [135] et 150 x 130 images de pixels avec des algorithmes fondés sur monogénique [128] alors que toutes les autres méthodes ont été effectuées sur 128 x 128. Plus précisément, les mêmes paramètres que sur les images FERET frontales ont été appliqués avec ELBP (h), ELBP (h + v) alors que les méthodes monogéniques [128] utilisent leurs paramètres par défaut.

Toutes les expériences ci-dessus ont été effectuées sur une machine de pc portable hp (intel(R) Core(TM)2Duo CPU T6570 @ 2.10 GHz, 2.10 GHz, RAM 2 Go) qui a été installé

avec Windows 7 32 bit et Matlab 2010 32 bits environnement de programmation. Bien que la machine ait un processeur central multi core (2 coeurs), toutes les implémentations testées ne sont pas parallèles. Nous avons évalué chaque point de référence 100 fois et indiqué les résultats moyens du tableau 4.8.

On peut voir à partir du tableau 3.7 qu'ELBP (h) est très rapide lorsqu'il peut finir jusqu'à 260 images par seconde (environ 3,7 millisecondes pour une image). De plus, ELBP (h + v) est rapide, car il ne nécessite que 6,1 millisecondes pour traiter une image. L'étape initiale de la méthode à base d'ondelettes de Gabor est environ 13 fois plus lente qu'Elbp (h + v) (bien que la taille de l'image dans le calcul des ondelettes de Gabor soit beaucoup plus petite) alors que celle du filtre monogénique est environ 2,7 fois plus lente qu'ELBP (v). Avec sa vitesse rapide, ELBP peut-être utilisé comme descripteur élémentaire pour former des représentations faciales multirésolutions / multiéchelles avancées.

4.4.4 Conclusion

Cette partie de la dissertation introduit une nouvelle variante de l'opérateur LBP dite ELBP. Nous utilisons des motifs d'ellipse horizontale et verticale pour former le descripteur de face ELBP pour l'extraction de caractéristiques. Ensuite, les images ELBP sont divisées en sous-régions rectangulaires pour construire leurs séquences d'histogramme ELBP. Le vecteur d'entités ELBP est généré en concaténant les séquences d'histogrammes des sous-régions. Dans l'étape de réduction de dimension, nous utilisons WPCA pour une meilleure performance de reconnaissance.

Les évaluations expérimentales sur AR et les bases de données FERET montrent que l'Elbp est plus efficace que le LBP en ce qui concerne les caractéristiques micros faciales et l'Elbp peut bien fonctionner dans diverses conditions telles que l'occlusion partielle, les expressions faciales, les variations de temps et de pose. De plus, la performance de reconnaissance sur la base de données SCface prouve l'efficacité de l'ELBP pour le problème de la reconnaissance faciale dans le contexte de la vidéosurveillance. Le LBP original est populaire pour sa robustesse à la rotation parce qu'il emploie des modèles circulaires. Bien que nos résultats dans cette section démontrent les avantages de l'ELBP sur LBP pour la reconnaissance faciale, nous ne suggérons pas qu'Elbp est robuste contre la rotation. De plus, le code ELBP est une caractéristique orientée qui contient des caractéristiques faciales en orientation horizontale, la direction principale de l'information de visage, et également l'information de direction verticale. Cela fait que la description ELBP à une puissance discriminative plus forte et, par conséquent, gagne des performances plus élevées que LBP. Sans aucun doute, nous croyons fortement que l'ELBP peut obtenir de meilleurs résultats dans les domaines de recherche liés à la reconnaissance faciale, où LBP a été appliquée.

4.5 LPQ comme descripteur d'extraction de caractéristiques du visage

En utilisant LPQ pour l'extraction de caractéristiques, deux cadres FR, le modèle d'appariement et le WPCA base qui emploie la fonction de distance de cosinus négatif pour la classification et WPCA pour la réduction de dimension (pour plus de détails, voir chapitre II) sont formées et évaluées sur les bases de données AR, FERET et SCface. Ces systèmes sont nommés LPQ pour le cas d'appariement de modèle et comme LPP WPCA quand ils se réfèrent au cadre fondé sur WPCA dans les tableaux de comparaison dans la section suivante.

Expériences sur la base données AR

Les résultats complets des cadres de FR fondés sur LPQ sur la base de données AR sont rapportés et comparés à ceux d'ELBP ($h + v$) dans le tableau 4.8 tandis que les comparaisons entre LPQ-WPCA et d'autres méthodes de pointe sont indiquées au tableau 3.8. De plus, les RR des systèmes fondés sur LPQ et ELBP ($h + v$) sont représentées sur les Fig. 3.5 et Fig. 3.6 pour des comparaisons plus visuelles.

Il se dégage du tableau 3.8 et des Fig. 3.5 et 3.6 que dans le cadre d'appariement des modèles, LPQ surpasse de manière significative ELBP ($h + v$), en particulier sur les ensembles de sondes difficiles, tels que 8, 9, 11, 12 (écharpe et lunettes de soleil avec des variations d'éclairage) de quatre expériences et 3 dans les expériences 3 et 4. Taux de reconnaissance moyens, dans les expériences 1 et 2, la LPQ est inférieure à 4% et 5% supérieure à ELBP ($h + v$), respectivement, alors que les numéros des 3 et 4 ans sont encore plus impressionnants: 9% et 10%. Le cadre LPQ WPCA a également des résultats globaux plus élevés que le WPCA ELBP ($h + v$) dans toutes les expériences. Pour une comparaison plus détaillée, dans les expériences 3 et 4, où il y a présence de variations temporelles, sur les ensembles de sondes de lunettes de soleil (numéros 10, 11 et 12), LPP WPCA a des taux de reconnaissance inférieurs, mais pas beaucoup ($h + v$) WPCA. En résumé, nous concluons que LPQ est plus robuste qu'Elbp ($h + v$), ELBP (h) et LBP [3] méthodes d'extraction de caractéristiques lors de l'adaptation aux expressions faciales, l'illumination, les variations de temps d'intervalle et les occlusions. Une autre conclusion à partir des résultats très élevés de LPQ WPCA (toutes les 4 expériences) sur la sonde fixe les nombres 4, 5 et 6 (expression neutre avec variations d'éclairage) est que LPQ est robuste aux variations d'éclairage.

Tableau 4.8: Rang-1 RR (%) comparaison entre les méthodes ELBP (h + v) et LPQ Sur la base de données AR

	Test/Méthode	Probe set									Avg
		1	2	3	4	5	6	7	8	9	
Exp2	ELBP(h + v)	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
	LPQ	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
	ELBP(h + v) WPCA	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
	LPQ WPCA	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
Exp2	ELBP(h + v)	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
	LPQ	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
	ELBP(h + v) WPCA	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1
	LPQ WPCA	100	100	74.4	100	100	97.8	76.9	55.2	45.4	38.1

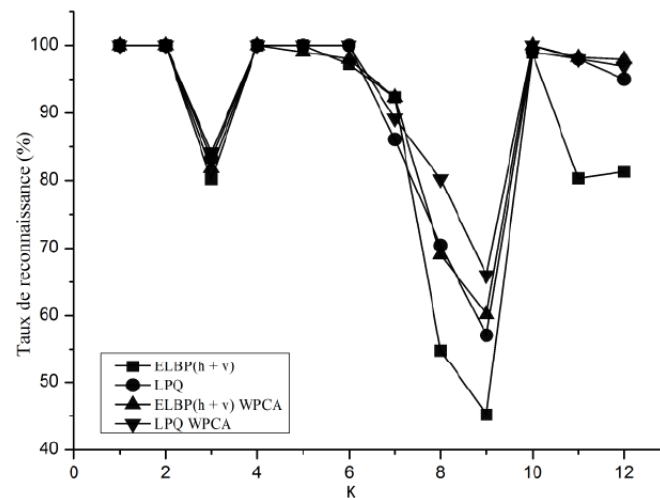


Figure 4.5: Expérience 1 exactitude des systèmes LPQ et ELBP(h + v) sur la base de données AR .

4.5.1 Résultats sur la base de données FERET

Les résultats des systèmes FR fondés sur la LPQ sur la base de données FERET (en utilisant le protocole standard) par rapport aux concurrents ELBP et d'autres concurrents à la pointe de la technique sont présentés au tableau 3.9. Ces résultats confirment que la LPQ est plus puissante que LBP, ELBP (h) et ELBP (h + v) contre les expressions

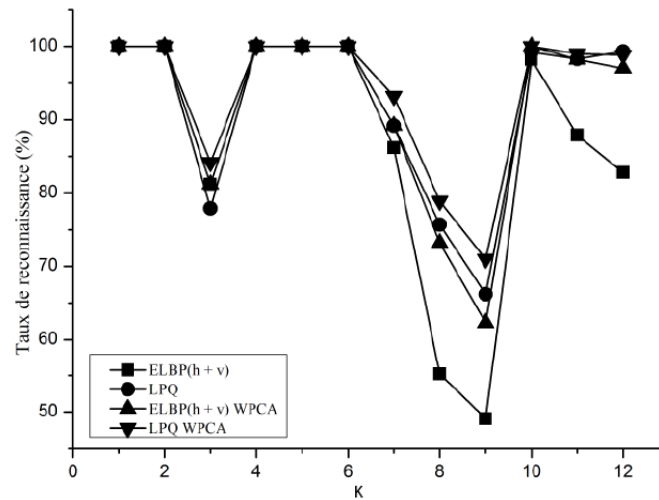


Figure 4.6: Expérience 2 exactitudes des systèmes LPQ et ELBP($h + v$) sur la base de données FERET.

Tableau 4.9: Rang-1 RRs (%) comparaison du système basé sur LPQ avec d'autres résultats de l'état de l'art sur la base de données FERET

Method	Fb	Fc	Dup1	Dup2	Average
ELBP($h + v$)	96.0	N/A^2	49.0	84.0	87.0
LGBPHS [135]	Smile	Anger	Scream	Glasses	Scarf
HMBP [130]	96.0	N/A^2	49.0	84.0	87.0
GEWC [29]	96.0	N/A^2	49.0	84.0	87.0
LPQ	96.0	N/A^2	49.0	84.0	87.0
HGPP [173]	Smile	Anger	Scream	Glasses	Scarf
DMMA [78]	96.0	N/A^2	49.0	84.0	87.0
LGBPWP [88]	96.0	N/A^2	49.0	84.0	87.0
CHG [22]	96.0	N/A^2	49.0	84.0	87.0
DLBP [82]	96.0	N/A^2	49.0	84.0	87.0
Tan et al [113]	96.0	N/A^2	49.0	84.0	87.0

faciales, l'illumination et les variations temporelles dans les frameworks de modèle et de WPCA. La prédominance de LPQ sur les méthodes LBP (LGBPHS [135], HMBP [130], DLBP [82], Tan et al. [113] et POEM PDO [118]) et ELBP est plus convaincante sur la

sonde Dup 1 et Dup 2 ensemble, dont les images sont affectées par la variation de laps de temps, l'un des facteurs les plus difficiles de FR. En outre, le taux élevé de reconnaissance des cadres fondés sur la LPQ nous amène à de nouvelles découvertes que LPQ est une méthode efficace d'extraction de caractéristiques sous l'expression du visage, l'illumination et les variations de temps et la combinaison de LPQ avec WPCA peuvent constituer un excellent système de FR.

4.5.2 Résultats sur la base de données FRGC 2.0

Cette section présente les résultats de LPQ-WPCA sur la base de données FRGC 2.0 [42]. Les images de la galerie sont également exploitées pour la phase d'entraînement comme dans les systèmes fondés sur l'ELBP. Comme tous les résultats obtenus par les cadres, WPCA, nous adaptons chaque système par le nom de la description faciale qu'il a utilisé.

Il ressort clairement au tableau 4.12 et du tableau 4.13 que la méthode LPQ-WPCA surpasse d'autres systèmes de l'état de la technique dans les deux expériences. Notre résultat moyen est d'environ 10 fois plus élevé que le niveau de référence PCA [42] dans les tests le taux de reconnaissance moyen de la ligne de base PCA [42] est 5 fois inférieur à la nôtre. LPQ WPCA obtient de meilleurs taux de reconnaissance globale de 3,9% (46,6% contre 42,7%) dans l'expérience. Selon le meilleur de notre connaissance, les résultats de LPQ-WPCA sur la base de données FRGC 2.0 sont les meilleurs résultats rapportés à ce jour dans la littérature de reconnaissance faciale. Ces résultats sont cohérents avec ceux des bases de données AR et FERET dans les sections précédentes.

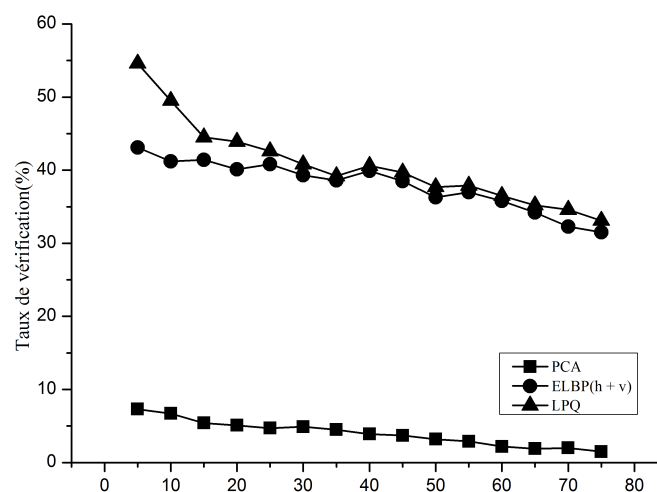


Figure 4.7: Exactitude des systèmes LPQ et ELBP(h + v) sur la base de données FRGC 2.0 .

Nous soutenons que les bons résultats du cadre LPQ-WPCA sur la base de données FRGC 2.0 ainsi que sa supériorité d'exactitude par rapport à la méthode fondée sur ELBP proviennent principalement de motifs fondés sur des phases significatives dans le vecteur caractéristique LPQ.

Paramètres de LPQ

Il existe plusieurs paramètres dans la construction d'un vecteur LPQ: la taille de fenêtre de la transformée STFT-M, l'argument de blanchiment- ρ et les sous-régions $R \times C$ pour capturer la séquence d'histogramme spatial final. Empiriquement sont fixées à 0,91 pour toutes les expériences. Les sous-régions divisées sont affectées respectivement de 8×9 et de 10×10 pour les bases de données AR et FERET. Sur la base de données de FRGC 2.0, puisque les images sont de tailles petites, elles sont divisées en 4×4 sous-régions rectangulaires non chevauchées pour faire l'extraction de caractéristiques avec LPQ. De plus, le paramètre M dépend de la taille de l'image lorsqu'il est réglé à 7 pour la base de données FRGC 2.0 et à 9 pour les fichiers AR et FERET.

Coût de calcul

Tableau 4.10: Temps de calcul de ELBP par rapport à d'autres méthodes d'extraction de caractéristiques

Method	Image size	Time (s)	Extraction Time (ms)	Image/s
ELBP(h)	128X128	4.40	3.6	260
LPQ	128X128	5.30	4.5	210
ELBP(h + v)	128X128	7.12	6	160
Gabor wavelets	80X88	96.23	80.4	12
MBC-A [128]	150X130	30.54	25.54	39

Pour étudier le rendement informatique de LPQ, les mêmes tests de référence décrits à la section 3.1.3.5 de ce chapitre sont effectués avec lui et les métriques obtenues sont comparées à celles de la méthode ELBP et à d'autres méthodes d'extraction de caractéristiques.

À partir des résultats de référence (tableau 4.15), il apparaît que la LPQ est très rapide à calculer, malgré le fait que sa vitesse est un peu plus lente que celle de l'ELBP (h). C'est à cause des huit opérations de convolution 1D qui sont nécessaires pour construire le vecteur de caractéristique LPQ alors que dans ELBP (h) (et aussi LBP [24]), seul le seuillage entre les pixels d'image et ses voisins est effectué. LPQ est plus rapide qu'Elbp (v + h) [89] puisque ELBP (v + h) a besoin d'une paire d'opérateurs ELBP pour former

une représentation faciale. La génération d'images d'ondelettes de Gabor est assez lente, bien que ce ne soit qu'une première étape et qu'elle fonctionne sur des images beaucoup plus petites (résolution 80 x 88). En se fondant sur la comparaison entre l'étape initiale des ondelettes de LPQ et de Gabor (dans le tableau 4.15), nous concluons que LPQ est d'environ $96.23 / 5.45 = 17.7$ fois plus rapide que toute méthode d'extraction de caractéristiques fondée sur des ondelettes de Gabor (LGBP [83], HGPP [84], FLPGMP [80], GFS [81], etc.).

4.5.3 Conclusion

Nous résumons ici les propriétés de la méthode d'extraction de la fonction LPQ. Pour chaque pixel, LPQ capture les modèles fondés sur la phase locale fondés sur la transformation STFT sur son voisinage carré dont le centre est le pixel lui-même. La représentation LPQ présente les propriétés suivantes:

1. Les modèles LPQ sont des caractéristiques de phase et ils sont flous invariants. Comme de code LPQ d'un pixel d'image est calculé sur la base de tous les pixels d'une fenêtre carrée centrée au pixel, la dépendance de ce pixel et de ses voisins est plus forte que dans les méthodes LBP et ELBP. Ainsi, LPQ est plus robuste que les méthodes LBP et ELBP lors de leur application sur l'image d'intensité. Cette supériorité de LPQ s'avère stable et cohérente dans un large éventail d'en ce qui concerne la FR comme l'expression faciale, l'illumination, la pose, les variations temporelles, les occlusions partielles et les images à basse résolution.
2. LPQ est une méthode efficace d'extraction de traits faciaux lorsque vous faites face à différents défis FR qui sont mentionnés dans la remarque ci-dessus. Les résultats expérimentaux étendus sur les bases AR, FERET et FRGC 2.0 sont des évidences évidentes pour cette conclusion.
3. LPQ est rapide à calculer et par conséquent, peut être encore attelé à développer des méthodes d'extraction de fonctionnalités plus avancées en utilisant une stratégie multirésolution / multiéchelle
4. Lorsqu'il rejoint LPQ avec WPCA, le cadre LPQ WPCA obtient d'excellents résultats et surpasse de nombreux systèmes à base d'ondelettes LBP et Gabor à la pointe de la technique dans la littérature FR

Outre ces précieuses caractéristiques, la méthode LPQ présente un inconvénient: son vecteur caractéristique est une description dense de la séquence d'histogrammes à 256 bits, et nécessite donc plus d'espace mémoire pour être stocké que les méthodes fondées sur ELBP mais n'affecte pas sa vitesse de traitement. Comment résoudre ce problème pour réduire la longueur du vecteur LPQ avant de l'alimenter dans les étapes de formation et de projection de WPCA est toujours une question ouverte.

4.6 Vérification de visage 3D

Dans cette partie, nous proposons un nouveau cadre pour la vérification du visage 3D en utilisant la réduction de dimension basée sur les descripteurs locaux dans un environnement non contrôlé. Les histogrammes des descripteurs locaux sont utilisés pour représenter distinctement les images faciales. À cet effet, différents descripteurs locaux sont évalués : LBP, TPLBP, FPLBP, 3DLBP et LPQ. Des expériences sur les combinaisons de ces descripteurs sont effectuées au niveau des caractéristiques à l'aide d'une concaténation simple des histogrammes. Tout système de reconnaissance biométrique comporte une phase très importante basée sur la réduction d'espace. Pour cela, nous consacrons un chapitre pour présenter les différentes techniques de projection d'espace. La classification n'est pas à négliger. Nous essayons de présenter l'essentiel des approches utilisées. La réduction de la dimension se pose comme une étape primordiale dans le processus de prétraitement des données (filtrage, nettoyage, élimination des points aberrants, etc.). En effet, pour des données appartenant à un espace de grande dimension, certains attributs n'apportent aucune information voire expriment du bruit, d'autres sont redondants ou corrélés. Ceci rend les algorithmes de décision complexes, inefficaces, moins généralisables et d'interprétation délicate. Les méthodes de réduction de la dimension de l'espace de représentation peuvent être divisées en méthodes d'extraction d'attributs et méthodes de sélection d'attributs. L'extraction d'attributs transforme l'espace d'attributs de départ en un nouvel espace formé de la combinaison linéaire ou non linéaire des attributs initiaux.

L'approche proposée est évaluée avec différents algorithmes de réduction de la dimensionnalité : PCA, OLPP et la méthode combinée PCA+EFM. Enfin, le SVM multi-classes est utilisé comme classificateur pour effectuer la vérification entre les imposteurs et les clients. La méthode proposée est testée sur la base de données CASIA3DV4 et les résultats montrent que notre méthode permet d'obtenir de hautes performances de vérification.

4.6.1 Aperçu sur l'approche proposée

Dans cette étude, les caractéristiques locales du visage sont utilisées pour améliorer la performance d'un système de vérification de visage 3D à partir d'images de profondeur 2.5D. Tout d'abord, la région du visage est détectée puis prétraitée. Ensuite, l'image faciale est passée à travers l'un des descripteurs locaux LBP, TPLBP, FPLBP ensuite découpée en plusieurs blocs rectangulaires. À cet effet, l'histogramme de chaque bloc est extrait. Ensuite, les histogrammes de tous les blocs sont concaténés dans un vecteur caractéristique unique. Ces vecteurs sont caractérisés par une grande dimensionnalité. Afin de réduire la dimension de ces vecteurs, différentes méthodes, notamment PCA, OLPP et PCA+EFM sont utilisées. Enfin, le SVM multi-classes est adopté pour la classification et la vérification. L'approche proposée se divise en deux phases : la phase d'apprentissage (l'enrôlement) et la phase de test. Nous utilisons la base de données de visage CASIA3DV4, qui contient 123 personnes avec différentes variations d'illumination, d'expressions et illumination combinée avec des expressions. Les sujets (personnes) sont

formés en tant que clients et imposteurs selon un protocole spécifique au cours de la phase d'apprentissage. Le même traitement est effectué dans les deux phases de manière à obtenir les vecteurs caractéristiques.

Dans la phase de test, les vecteurs caractéristiques sont appariés avec chaque échantillon défini dans la phase de l'apprentissage à l'aide de SVM multi-classes afin d'accepter la personne de test en tant que client ou la rejeter comme imposteur. L'organigramme de l'approche proposée est montré dans la figure 4.8. Généralement, le prétraitement est un processus nécessaire pour le domaine de traitement d'image. L'objectif du processus de prétraitement est l'amélioration de l'effet visuel des images faciales et leur clarté afin d'être vraiment expressive de la forme de visage humain. Notre méthode de prétraitement et détection est décrite en détail dans le chapitre 3, section 3.4. Cependant, dans ce chapitre, nous utilisons les images de profondeur 2.5D comme entrée pour le système RV3D. La figure 4.2 montre un exemple de prétraitement de l'image de profondeur.

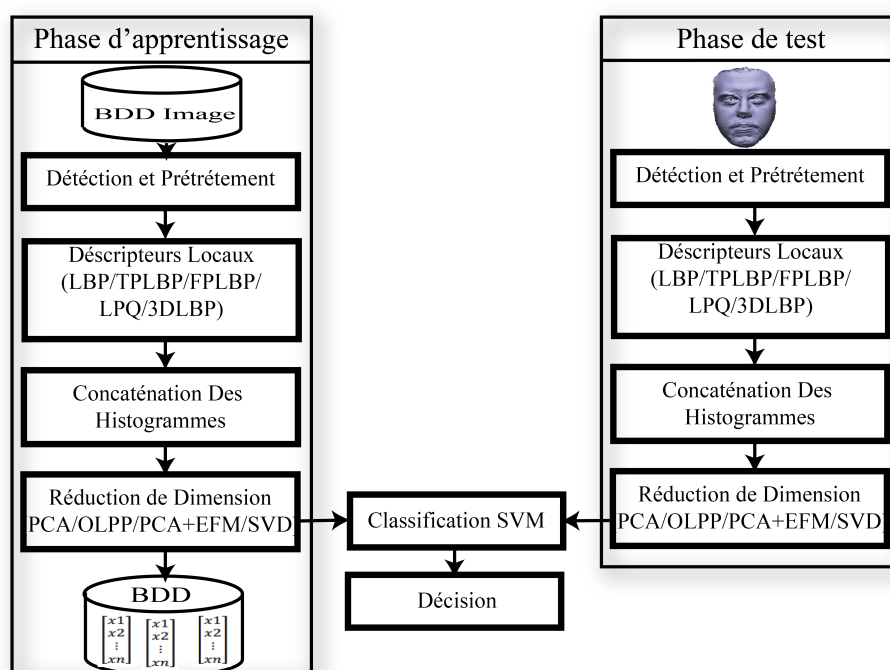


Figure 4.8: Organigramme de l'approche proposée.

4.6.2 Les descripteurs locaux

La représentation de l'image faciale est un processus important pour l'efficacité du système de vérification. Dans lequel, l'information appropriée est obtenue pour être invariante et

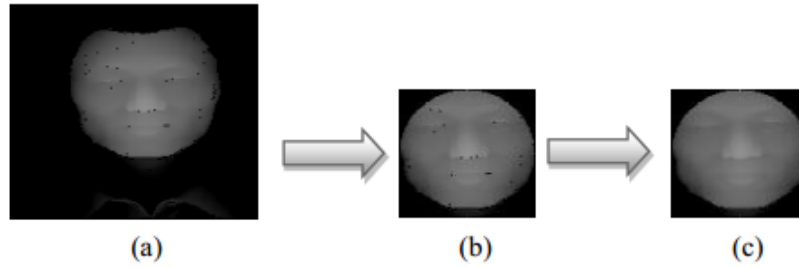


Figure 4.9: Prétraitement de l'image de profondeur 3D ;(a) Image d'entrée, (b) Image détectée, (c) Image prétraitée.

discriminante. Les descripteurs locaux ont été considérés comme l'une des méthodes les plus efficaces adoptées à cet effet.

Les descripteurs locaux peuvent être utilisés dans la reconnaissance faciale, l'analyse de texture, la reconnaissance d'objets et la détection des régions d'intérêt. Récemment, ces méthodes ont été largement utilisées pour les systèmes de reconnaissance de visage [122, 130- 137]. L'idée principale des descripteurs locaux est de représenter les images faciales de manière discriminative avec leurs micros caractéristiques d'apparences locales.

Motif binaire local (LBP)

L'opérateur LBP a été initialement proposé par T. Ojala et al [140] en 2002 afin d'exprimer la texture des patches de l'image. Il a été largement appliqué avec divers algorithmes de systèmes de reconnaissance de visage comme une méthode d'extraction de caractéristiques locales [98]. L'opérateur LBP de base attribue à chaque pixel un motif binaire. Le LBP d'une image de pixel est produit par le seuillage du voisinage $3 * 3$ avec le pixel central (si la valeur du pixel central était supérieure à la valeur du son pixel voisin ou non) et le transfert comme un code binaire qui est converti en un nombre décimal. La figure 4.10 donne une illustration de l'opérateur LBP de base. Après cela, cet opérateur a été étendu pour utiliser de différents rayons de voisinages R et différents points d'échantillonnage P [94] ce qui permet d'extraire les caractéristiques dans différentes échelles. Soit I_c un pixel dans l'image d'entrée, ses pixels voisins sont $(I_1, I_2, \dots, I_{P-1})$. La réponse LBP du pixel I_c est calculée comme suit :

$$LBP_{P,R} = \sum_{P=0}^{P-1} s(I_P - I_c) 2^P \quad (4.1)$$

$$s(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (4.2)$$

Le LBP étendu sélectionne les pixels voisins comme un ensemble de points d'échantillonnage réparti uniformément le long d'un cercle avec comme centre le point I_c (pixel central) et

un rayon R comme représenté dans la figure 4.11.

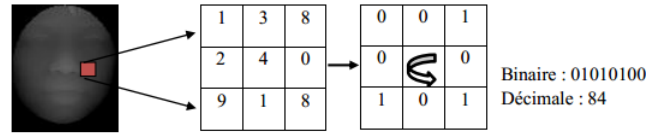


Figure 4.10: Une illustration de LBP basique.

Dans la littérature la notion LBP est généralement utilisée pour désigner l'opérateur LBP basique, tandis que la notion $LBP_{P,R}$ est utilisée pour représenter l'étendue LBP où, l'indice P représente le nombre des points d'échantillonnage et l'indice R représente le rayon du cercle.

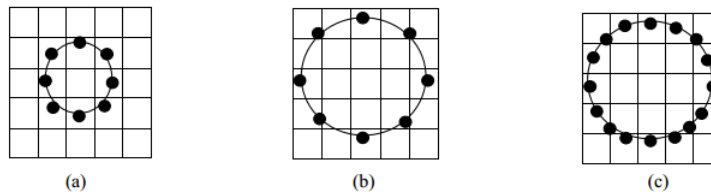


Figure 4.11: Exemples d'opérateur $LBP_{P,R}$ (a) $LBP_{8,1}$ (b) $LBP_{8,2}$ (c) $LBP_{16,2}$.

Le nombre total des valeurs des différentes sorties est 2^P sont générées par l'opérateur $LBP_{P,R}$ avec certaines valeurs correspondantes aux mêmes motifs suivant la rotation. Afin d'éliminer l'effet de la rotation et pour améliorer encore la puissance discriminante du LBP d'autres méthodes étendues sont proposées, telle que le Rotation-invariant LBP ($LBP_{P,R}^{ri}$) et l'Uniform-rotation-invariant LBP ($LBP_{P,R}^{uri}$) [22, 140, 142]. La figure 4.12 illustre une représentation d'image de profondeur avec le descripteur LBP et différentes valeurs de P et R .

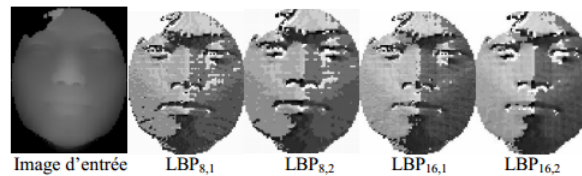


Figure 4.12: Représentation d'une image de profondeur avec le descripteur LBP.

4.6.3 LBP basée sur les patches (Patch base LBPs)

Les descripteurs LBP basées sur les patches sont placés dans une famille de nouveaux descripteurs d'images introduites en 2008 par Wolf et al [78]. Au lieu de faire la comparaison entre les pixels voisins et un pixel central pour donner à chaque pixel son motif, la

comparaison ici se fait entre les patchs voisins et un patch central. Deux types de ces descripteurs ont été proposés, Three-Patch Local Binary Patterns (TPLBP) et Four-Patch Local Binary Patterns. (FPLBP).

Motif binaire local à Trois-patchs (TPLBP)

Le descripteur TPLBP produit une valeur d'un seul bit dans le code attribué à chaque pixel à travers une comparaison de la similarité entre trois patchs différents. La similarité entre les patchs est calculée en utilisant la mesure de distance avec la norme L2. Pour chaque pixel de l'image, on considère un patch de taille $w * w$ centré dans ce pixel. Les patchs voisins sont distribués uniformément dans un anneau de rayon R autour du patch central (voir figure 3.13). Ce dernier est comparé avec les deux patchs qui se situent entre une valeur définie α . La valeur d'un seul bit est définie en fonction de celle des deux patchs la plus similaire. Le descripteur TPLBP est donné par l'équation suivante :

$$TPLBP = \sum_{i=0}^{S-1} f(d(P_0, P_i) - d(P_0, P_{i+\alpha \text{ mod}(S)}))2^i \quad (4.3)$$

L'exemple ci-dessous explique le codage du pixel selon le descripteur TPLBP avec $\alpha = 2, S = 8, et w = 3$:

$$\begin{aligned} TPLBP = & f(d(P_0, P_1) - d(P_0, P_3))2^0 + f(d(P_0, P_2) - d(P_0, P_4))2^1 \\ & + f(d(P_0, P_3) - d(P_0, P_5))2^2 + f(d(P_0, P_4) - d(P_0, P_6))2^3 \\ & + f(d(P_0, P_5) - d(P_0, P_7))2^4 + f(d(P_0, P_6) - d(P_0, P_8))2^5 \\ & + f(d(P_0, P_7) - d(P_0, P_1))2^6 + f(d(P_0, P_8) - d(P_0, P_2))2^7 \end{aligned} \quad (4.4)$$

Où d est la mesure de distance L_2 entre les patchs, P_0 patch central, P_i les patchs voisins du patch central ($i = 1, \dots, 8$) et f est la fonction de seuillage qui est calculée par :

$$f(x) = \begin{cases} 1 & \text{si } \tau \geq 0 \\ 0 & \text{si } \tau < 0 \end{cases} \quad (4.5)$$

τ : Le seuil de comparaison est choisi légèrement plus grand que 0 (dans notre cas=0.01) afin d'assurer une certaine stabilité dans les régions uniformes [56]. La figure 4.14 illustre une représentation de l'image de profondeur avec le descripteur TPLBP. La rangée du haut de gauche à droite : l'image de profondeur d'entrée, le code image TPLBP ($S=8$ et $r=3$) avec $w=3, 5, 7$ et 9 . La rangée du bas le code image TPLBP ($S=8$ et $r=5$) avec $w=3, 5, 7$ et 9 .

Motif binaire local à Quatre-Patch (FPLBP)

Le FPLBP est une autre variante de descripteur LBP basée sur les patchs presque avec la même idée de TPLBP. Cependant, il y a une différence dans le nombre d'anneaux et

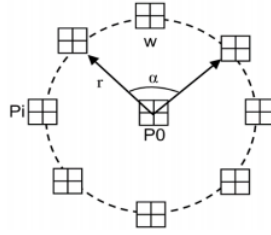


Figure 4.13: La procédure de descripteur TPLBP.

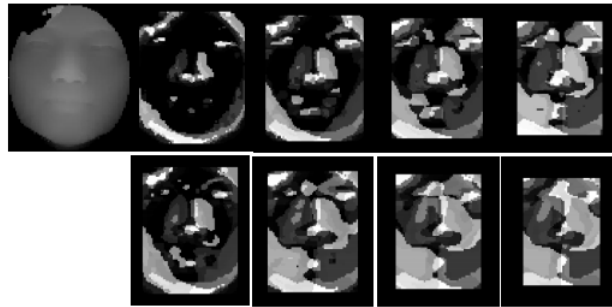


Figure 4.14: La représentation TPLBP d'une image de profondeur avec différentes valeurs de rayons et de tailles des patches w

des patches utilisés pour la comparaison. L'opérateur FPLBP utilise deux anneaux avec deux rayons r_1 et r_2 centrés dans un pixel de l'image. Les S patches sont uniformément répartis autour de ces deux anneaux (voir figure 3.15). Dans ce cas, la comparaison est produite entre deux centres de patches symétriques dans l'anneau intérieur et deux centres de patches symétriques dans l'anneau extérieur positionnés le long d'un cercle dans le sens des aiguilles d'une montre $\alpha(\alpha = 1)$ [121]. Après la comparaison, un bit dans chaque pixel de l'image est défini en fonction de celui des deux patches les plus similaires. Pour les S patches le long de chaque cercle, nous avons $\frac{S}{2}$ paires de centres symétriques qui sont la longueur des codes binaires produits. Le code FPLBP est donné par :

$$FPLBP = \sum_{i=0}^{\frac{S}{2}-1} f(d(P_{1,i}, P_{2,i+\alpha \text{ mod}(S)}) - d(P_{1,i+\frac{S}{2}}, P_{2,i+\alpha+\frac{S}{2} \text{ mod}(S)}))2^i \quad (4.6)$$

L'exemple ci-dessous explique le codage de pixel selon le descripteur FPLBP avec $\alpha = 2, S = 8, \text{ et } w = 3$:

$$FPLBP = (f((P_{1,0}, P_{2,1}) - d(P_{1,4}, P_{2,5}))2^0 + f((P_{1,1}, P_{2,2}) - d(P_{1,5}, P_{2,6}))2^1 + f((P_{1,2}, P_{2,3}) - d(P_{1,6}, P_{2,7}))2^2 + f((P_{1,3}, P_{2,4}) - d(P_{1,7}, P_{2,8}))2^3) \quad (4.7)$$

La figure 4.16 illustre une représentation de l'image de profondeur avec le descripteur FPLBP. La rangée du haut de gauche à droite : l'image de profondeur d'entrée, le code image FPLBP ($r_1 = 1, r_2 = 5, S = 8$) avec $w = 3, 5, 7$ et 9 . La rangée du bas de gauche à droite : le code image FPLBP ($r_1=1, r_2=5, S=12$) avec $w=3, 5, 7$ et 9 .

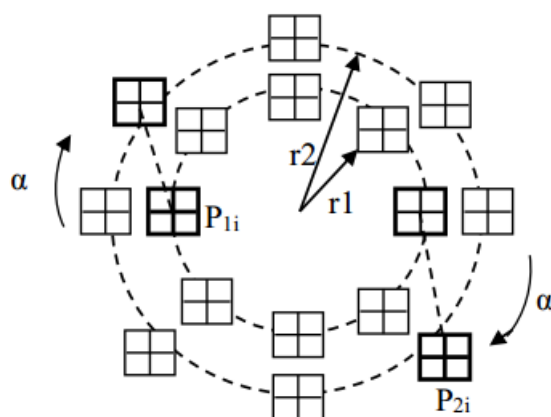


Figure 4.15: La procédure du descripteur FPLBP.

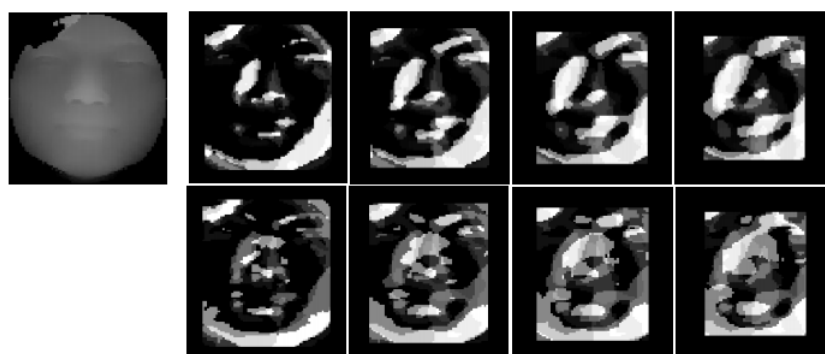


Figure 4.16: La représentation FPLBP d'une image de profondeur avec différentes valeurs de tailles de patches w et de nombre des patches voisins S .

4.6.4 Méthode de 3DLBP

Yonggan Huang et al. [69] ont proposé une version étendue des LBP, nommée 3DLBP. Outre les informations fournies par LBP, 3DLBP considère également la valeur de la profondeur notée DD(Depth Difference) entre le pixel central et son pixel voisin. Un ensemble de trois bits est utilisé pour coder les valeurs absolues des différences de profondeur. Il est ensuite combiné aux signes (0 ou 1) obtenus par les LBP. En se basant sur une étude statistique, les auteurs considèrent une valeur maximale de 7 pour ces valeurs de différence. En effet, dans leur étude, les auteurs ont noté que plus de 93% des DD obtenues entre les pixels avec un rayon $R=2$ sont inférieures à 7. Ceci est dû à l'aspect lisse de la profondeur du visage, ce qui n'est pas le cas pour les images en niveaux de gris où les valeurs de différence entre les pixels voisins peuvent être arbitrairement grandes selon la texture et les conditions de l'environnement. Une DD est donc représentée par quatre bits i_1, i_2, i_3, i_4 , le bit i_1 est le code représentant le signe, et i_2, i_3, i_4 est les trois bits représentant la valeur absolue de la différence de profondeur ($|DD|$). Toute différence supérieure au seuil est ramenée à 7. Les bits sont ensuite séparés en 4 couches. Pour

chaque couche, les bits correspondants de l'ensemble des DD des pixels voisins sont concaténés pour générer un code binaire. Au total, pour chaque pixel, on obtient 4 codes binaires P_1, P_2, P_3, P_4 , qui sont ensuite convertis en 4 valeurs décimales. L'histogramme de chaque couche est calculé, puis les histogrammes sont concaténés pour former un descripteur unique pour l'image. La Figure 4.17 illustre le mode fonctionnement du 3DLBP.

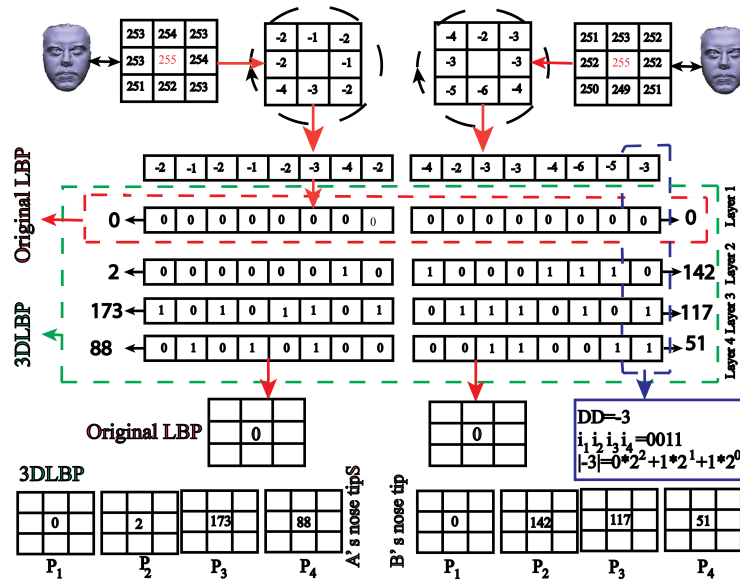


Figure 4.17: Exemple de calcul des 3DLBP.

Les auteurs ont évalué le descripteur proposé sur la collection FRGC. Les résultats montrent que la précision obtenue en utilisant les 3DLBP est meilleure que celle des LBP et des cartes de profondeur brutes. Bien que l'utilisation des 3DLBP augmente le pouvoir discriminant des LBP en utilisant, en plus du signe, la valeur de différence de profondeur dans le voisinage, ce descripteur souffre de certaines limites :

- La taille du vecteur caractéristique est très grande. Chaque image est représentée par 4 matrices de même taille. Ces matrices sont segmentées en régions et un histogramme par région pour chaque matrice est calculé. Ce qui donne un vecteur caractéristique de tailles $H*4*256$, ou H est le nombre de régions de découpage. Ceci implique un temps de traitement plus long ;
- La méthode de codage n'est pas consistante avec la variation de la profondeur. Une très faible variation dans le voisinage du pixel peut conduire à une grande différence dans les codes calculés. Par exemple, pour un pixel de valeur 10 et dont le voisinage est 13, 14, 15, 12, 11, 9, 11, 12, les différences de profondeur sont alors 3,4,5,2,-1,1,2, les codes 3DLBP obtenus sont 251, 96, 145, 174. Pour un changement mineur de la valeur du premier voisin de 13 vers 14, on obtient un code 3DLBP 251 , 224, 17, 46 . Ceci est dû à la manière de répartir les valeurs des 4 bits sur les différentes couches encodant les DD ;

- Les valeurs DD sont considérées inférieures ou égales à 7 et sont codées par conséquent sur 3 bits. Cependant, l'étude statistique sur laquelle les auteurs se basent est limitée à un petit rayon ($R=2$). Par conséquent une extension multi-échelles ne peut pas être envisagée en utilisant les 3DLBP, puisque le seuil de 7 pour les DD ne serait alors plus valide. Afin de généraliser cette approche pour qu'elle puisse être utilisée avec différents rayons, il est nécessaire de représenter la différence sur 8 bits, soit 256 différences de profondeur possibles. Cependant, cette généralisation a le défaut d'engendrer une représentation de taille importante : 9 matrices pour représenter un seul visage de profondeur.

4.6.5 Extraction des caractéristiques d'histogrammes

Après le traitement des images faciales avec l'un des descripteurs locaux, nous obtenons des images étiquetées (le code image) avec le modèle correspondant. Ensuite, l'image de visage est subdivisée en différents blocs rectangulaires. L'histogramme de chaque bloc est calculé. Ensuite, tous les histogrammes sont concaténés dans un vecteur caractéristique unique. La figure 3.18 illustre notre méthode d'extraction des caractéristiques. L'histogramme représente les informations concernant les micro structures telles que les bords, les tâches et les zones plates dans les régions locales [146]. Par conséquent, les histogrammes sont utilisés comme des caractéristiques discriminantes pour représenter l'information locale de l'image faciale. L'histogramme (H) de l'image d'entrée (I) est calculé par :

$$H(i) = \sum_{x,y} B(I_p(x,y) = i); \quad i \in [0, 2^p - 1] \quad (4.8)$$

$$B(v) = \begin{cases} 1 & \text{lorsque } v \text{ est vrai} \\ 0 & \text{autrement} \end{cases} \quad (4.9)$$

Où P est le nombre de bits. La taille du vecteur de caractéristiques de sortie est de 256 éléments multiplié par le nombre de blocs de chaque image et multiplié aussi par le nombre de bits de codage utilisé. Donc, un vecteur avec une grande dimension est produit pour la représentation de l'image faciale. La haute dimensionnalité des vecteurs caractéristiques conduit à un temps de calcul élevé ainsi qu'à une redondance de l'information entre les caractéristiques extraites, par conséquent, une faiblesse au niveau de la classification. Dans le but de limiter ces inconvénients, la réduction de dimensionnalité est effectuée. Cette étape est généralement.

4.7 Résultats de simulation et discussion

Cette section donne des détails sur les résultats expérimentaux du système proposé. Pour démontrer l'efficacité de l'approche proposée. Les expériences ont été réalisées sur la base

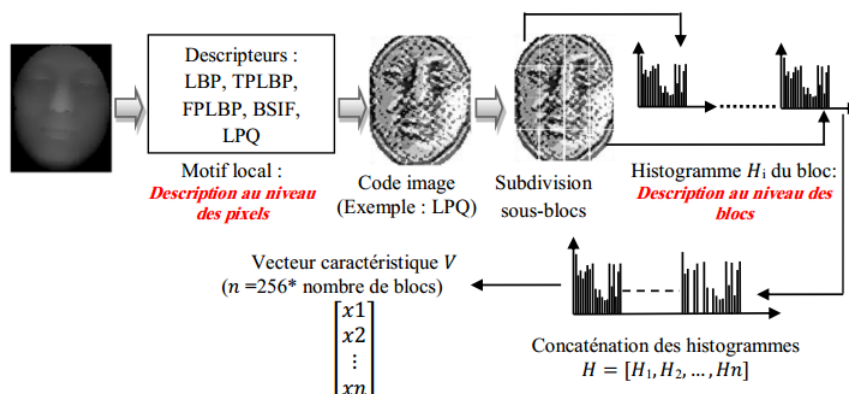


Figure 4.18: Les descripteurs et l'extraction de caractéristiques d'histogramme.

de données de visage FRGC 2.0, qui contient 123 personnes. Dans ce travail, nous utilisons 1.845 images où chaque personne est représentée par 15 modèles différents. La figure 4.18 montre les 15 images de profondeur de la même personne dans la BDD de visage FRGC 2.0. Comme le montre le tableau 4.11, notre protocole comprend trois partitions de données : Apprentissage, Évaluation et Test.

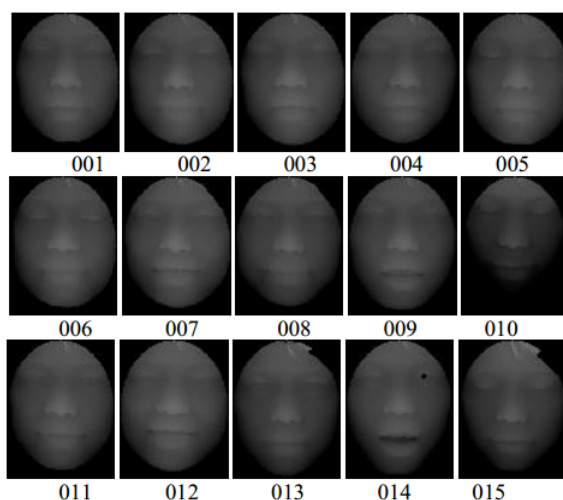


Figure 4.19: Image de profondeur de la même personne, première ligne variations d'illuminations, deuxième ligne variation d'expressions, troisième ligne variations d'expressions sous différentes illuminations.

Dans les premières expériences, le système de vérification 3D du visage proposé est évalué avec différents descripteurs locaux : LBP, TPLBP, FPLBP, 3DLBP et LPQ avec trois méthodes de réduction de dimension, PCA, DCT et SVD.

Dans la deuxième expérience, la combinaison des caractéristiques du descripteur LPQ

Tableau 4.11: Protocole utilisé pour la vérification

Dataset	Client	Imposteur
Apprentissage	500 images (1, 4, 8, 9, 10)	0 image
Evaluation	500 images (2, 6, 7, 14, 15)	195 images (1:15)
Teste	500 images (3, 5, 11, 12, 13)	150 images (1:15)

avec tous les autres descripteurs est utilisée avec la meilleure méthode de réduction de dimension.

Les résultats expérimentaux ont été obtenus avec les paramètres optimaux de chaque descripteur. Le descripteur LPQ est basé sur la quantification de la transformée de Fourier de la phase dans les voisins d'une fenêtre de taille 5×5 à chaque pixel. La taille optimale du filtre linéaire utilisé pour le descripteur 3DLBP est 17×17 . Tous les paramètres sont choisis suite aux résultats des expériences. La performance de notre système est testée avec le descripteur LBP. Le code d'image avec $p=16$ et $r=2$ nous donne le meilleur résultat. Les paramètres optimaux des descripteurs sont choisis expérimentalement. Ces paramètres sont montrés dans le tableau 4.12.

Tableau 4.12: Paramètres optimaux des descripteurs utilisés

Descripteur	Paramètres optimaux
LPB	$p=16, r=2$
3DLBP	$l=17$
LPQ	$R=5$
TPLBP	$r1 = 3, w = 8, s = 8, \tau = 0.01, r1 = 1$
FPLBP	$r2 = 5, w = 12, \tau = 0.01$

La figure 4.17 illustre une image de profondeur et le code de l'image LBP, TPLBP, FPLBP, 3DLBP et LPQ avec les paramètres optimaux.

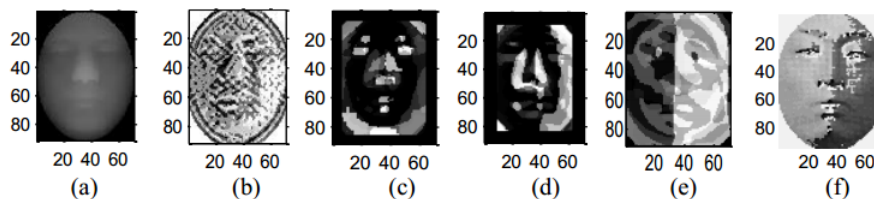


Figure 4.20: Représentation du visage basée sur les descripteurs locaux; (a) image d'entrée; (b) code image LPQ; (c) code image TPLBP, (d) code image FPLBP, (e) code image 3DLBP, (f) code image LBP

Les tableaux 4.13, 4.14 et 4.15 montrent les performances de la vérification avec les méthodes, PCA, DCT et SVD avec les quatre descripteurs locaux.

Tableau 4.13: Performance de la vérification avec PCA durant l'évaluation et le test

Descripteur Local	Evaluation	Test			
	TEE%	TFA%	TFR%	HTER%	TV%
TPLBP	9.50	7.30	8.20	8.10	84.2
FPLBP	8.25	6.40	6.80	6.30	87.2
3DLBP	7.2	7.15	5.30	6.60	87.5
LPQ	6.1	6.12	5.23	6.34	94.3

Tableau 4.14: Performance de la vérification avec DCT durant l'évaluation et le test

Descripteur Local	Evaluation	Test			
	TEE%	TFA%	TFR%	HTER%	TV%
TPLBP	6.1	3.4	3.8	3.2	91.58
FPLBP	5.6	2.6	4.3	3	92.30
3DLBP	4.2	2.3	2.6	1.9	95.3
LPQ	3.9	2.1	2.4	1.8	96.6

Tableau 4.15: Performance de la vérification avec SVD durant l'évaluation et le test

Descripteur Local	Evaluation	Test			
	TEE%	TFA%	TFR%	HTER%	TV%
TPLBP	5.3	3.2	3.7	3	92.3
FPLBP	4.3	2.3	4.2	2.9	93.2
3DLBP	4.1	2.1	1.9	1.6	96.3
LPQ	1.5	2	2.1	1.7	98.2

Nous pouvons constater que les deux algorithmes DCT et SVD sont parvenus à une amélioration considérable du taux de vérification par rapport à l'algorithme de base PCA avec plus de 9% à 11%. Par exemple avec l'opérateur TPLBP, le taux de vérification est égal à 84.2%, 91.58% et 92.3% avec PCA, DCT et SVD respectivement. Aussi, en utilisant le descripteur FPLBP nous obtenons 87.2%, 92.3% et 93.2% en utilisant PCA, DCT et SVD respectivement. Les résultats présentés dans les tableaux 4.13, 4.14 et 4.15 au cours d'évaluations et des tests indiquent que LPQ et 3DLBP conduisent aux hautes

performances de vérification ; nous obtenons un taux de vérification de 96.3% et 98.2% pour 3DLBP et LPQ et respectivement. D'autre part, nous pouvons confirmer ces résultats à travers les courbes ROCs des différentes expériences qui sont illustrées dans les figures 4.20, 4.21 et 4.22. La courbe ROC trace le taux de vérification en fonction de taux de fausse acceptation. La 3DLBP arrive à la deuxième meilleure performance de vérification après le LPQ avec toutes les méthodes de réduction de la dimension, sauf dans un cas lorsque nous utilisons l'algorithme de base PCA.

Nous obtenons un taux de vérification faible égal à 84.2% comme il est illustré dans le tableau 4.13. Les histogrammes LPQ explorent des caractéristiques très discriminantes pour la vérification, et ils surpassent toutes les autres caractéristiques des différents descripteurs avec un taux de vérification égal à 98.2% et un TEE égal à 1.5%.

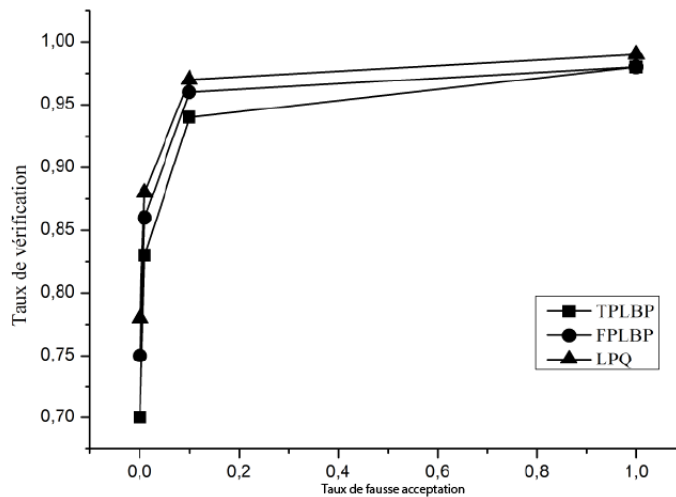


Figure 4.21: Courbes ROCs avec PCA

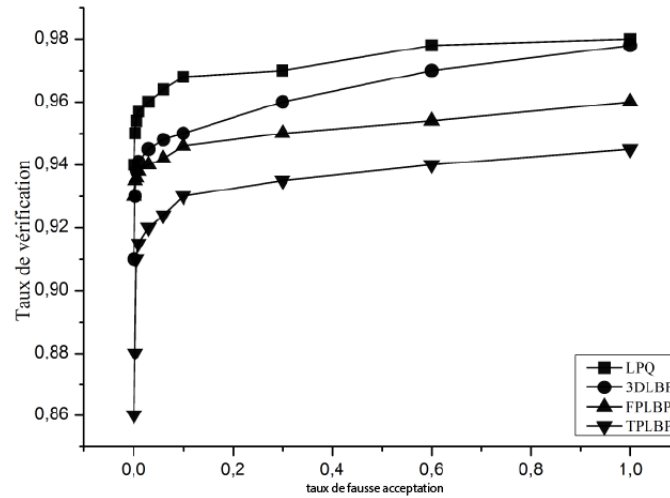


Figure 4.22: Courbes ROCs avec DCT

Nous pouvons dire que le descripteur LPQ basé sur Sud-est l'un des meilleurs cadre pour aborder l'effet de l'illumination et des d'expressions faciales dans le système de vérification du visage. Nous attribuons ce résultat au fait que les illuminations et les variations des expressions faciales n'ont aucun effet sur le domaine fréquentiel qui est représenté par le descripteur LPQ. Basée sur ces résultats, dans les deuxièmes expériences, la combinaison du descripteur LPQ avec tous les descripteurs est utilisée.

Le tableau 4.16 présente les performances de vérification de la combinaison LPQ avec les autres descripteurs basés sur PCA.

Tableau 4.16: Performance de la vérification avec PCA basé sur la combinaison de LPQ avec LBP, TPLBP, FPLBP durant l'évaluation et le test

Descripteur Local	Evaluation	Test				
	TEE%	TFA%	TFR%	HTER%	TV%	TC (s)
LPQ-LBP	1.45	0.5	2.1	1.3	97.2	2.30
LPQ-TPLBP	1.60	0.42	2.04	1.2	97.5	2.70
LPQ-FPLBP	1.20	0.31	1.50	0.80	98.20	2.70
LPQ-3DLBP	1.1	0.12	1.2	0.60	98.8	2.30

Dans ce tableau 4.16 les performances de vérification au cours de l'ensemble de test en utilisant $LPQ-3DLBP$ (TV = 98,20%, HTER = 0,80%) surpassent à la fois $LPQ-LBP$,

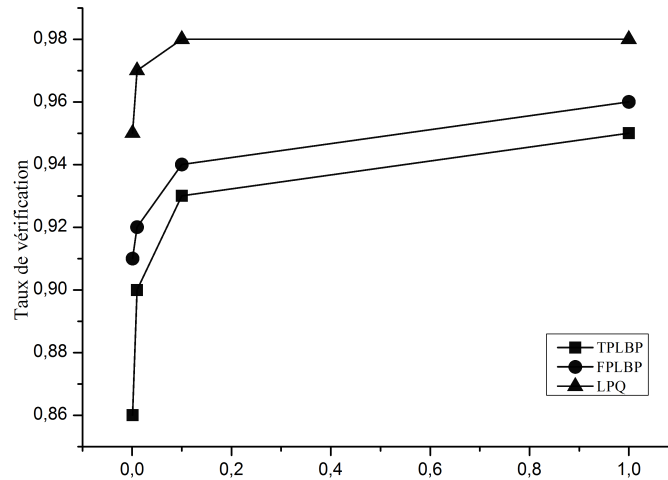


Figure 4.23: Courbes ROCs avec SVD

$LPQ - TPLBP$ et $LPQ - FPLBP$. Nous pouvons dire que la combinaison des histogrammes des caractéristiques locales $LPQ - 3DLBP$ basée sur la PCA lorsqu'on utilise des images de profondeur 3D donne des caractéristiques plus distinctes. Cette approche optimale produit une amélioration significative de la performance du système de vérification visage 3D.

Les résultats représentés dans le tableau 4.16 montrent que les méthodes étendues TPLBP et FPLBP peuvent surpasser le descripteur LBP, mais il n'y a pas une grande différence entre eux. Nous pouvons voir qu'il n'y a pas de gain de performance significatif dans les méthodes combinées. Le descripteur de LPQ dépasse certaines méthodes combinées telles que : $LPQ - TPLBP$, $LPQ - FPLBP$ et $LPQ - LBP$. Cela est dû au fait que ces descripteurs partagent beaucoup de caractéristiques similaires (redondance) qui ont conduit à un affaiblissement du processus de classification.

Dans le tableau 4.16, nous présentons le temps de calcul (TC) de chaque méthode. Il n'y a pas de TC significatif entre toutes les expériences (1,90 (s) à 2.70 (s)), car dans toutes les méthodes proposées, nous utilisons presque la même taille de vecteurs caractéristiques. Comme nous pouvons l'observer dans le tableau 4.16, le TC des méthodes combinées a augmenté. Ceci s'explique dans ce cas par la concaténation de deux vecteurs caractéristiques où le premier est le descripteur LPQ et le second est l'un des autres descripteurs locaux utilisés.

Nous examinons aussi les effets de la variation de la taille des vecteurs caractéristiques. Nous pouvons déterminer les meilleures dimensions du vecteur caractéristique expérimentalement en variant la taille entre 10 et 100. La figure 4.23 illustre la variation des taux de vérification à travers la dimension du vecteur caractéristique.

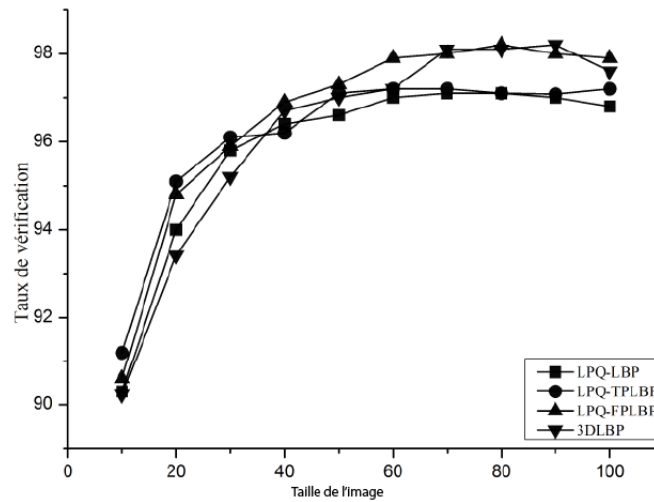


Figure 4.24: Variations de taux de vérification en fonction du nombre de caractéristiques

À partir de cette figure, nous affirmons que le vecteur de projection discriminante optimale des quatre méthodes combinées ($LPQ - LBP$, $LPQ - FPLBP$ et $LPQ - TPLBP$) en utilisant notre méthode PCA + EFM est situé entre 70 et 90.

Le tableau 4.17 présente une comparaison de nos meilleurs résultats en termes de taux de vérification avec plusieurs techniques existantes dans la littérature.

Tableau 4.17: Comparaison des taux de vérification avec l'état de l'art

Auteur	Méthodes	Base de données	TV(%)
X.Wang [92]	PCA, Gabor, LBP	FRGC 2.0	91.71
Y.A.Li [85]	Caractéristique géométrique, LDA, géodésique.	FRGC 2.0	91.10
Ouamane.A [89]	LBP, SIFT, SLF, SVM	FRGC 2.0	96.26
Notre méthode	$LPQ - FPLBP$, SVM	FRGC 2.0	98.20

4.8 Conclusion

Dans cette section, nous avons présenté une validation expérimentale de notre système : la vérification automatique du visage 3D en présence des variations d'illuminations et d'expressions faciales en utilisant des images de profondeur. Cinq des meilleurs descripteurs locaux LBP, LPQ, TPLBP, FPLBP et 3DLBP ont été utilisés. Les histogrammes de chaque descripteur sont extraits sur des blocs rectangulaires dans l'image faciale et concaténés dans un seul vecteur caractéristique. Afin de réduire la grande dimensionnalité de ces vecteurs, trois méthodes de réduction de dimension sont évaluées et comparées : PCA, PCA+EFM et OLPP. D'après nos résultats, nous avons démontré que la méthode combinée *LPQ - 3DLBP* basée sur PCA donne des performances de vérification très élevées avec un taux de vérification de 98,20% et HTER de 0,80% durant l'ensemble de tests. Nos résultats montrent que la méthode PCA a atteint une différence dans le TV pouvant aller jusqu'à 10%, et de 5,8% pour le TEE par rapport à la méthode de base PCA.

Conclusion générale

Dans cette thèse, nous nous sommes intéressés aux problèmes de la reconnaissance automatique de visages en présence de différentes variations, en mettant l'accent principal sur l'extraction des caractéristiques, la tâche la plus importante de chaque système de reconnaissance de visages. Nous avons proposé plusieurs méthodes visant des représentations faciales robustes contre une diversité de facteurs stimulants, nous avons présenté le procédé du système de reconnaissance faciale et les différents problèmes et défis scientifiques rencontrés dans le monde réel. Nous avons présenté les bases de données du visage 2D et 3D les plus utilisées dans la littérature pour l'évaluation des différents algorithmes de reconnaissance de visage ainsi que les critères d'évaluation et de mesure des performances pour les deux modes de reconnaissance vérification et identification. Un état de l'art des méthodes récentes et travaux antérieurs sur la détection et la reconnaissance de visage 3D est mis au point. Ceci nous a permis de choisir la méthodologie adéquate à la problématique et faire un choix judicieux pour la conception du système d'analyse d'images de visage en présence d'expressions et de la pose.

Avant d'énumérer ce qui a été fait dans cette thèse, nous souhaitons raconter en bref, et dans l'ordre chronologique, les concepts sous-jacents de toutes les propositions que nous avons faites. Notre première constatation est que l'information du visage est orientée dans la direction horizontale et les caractéristiques faciales importantes, comme les yeux et la bouche, sont en forme d'ellipse horizontale. C'est la motivation pour le descripteur de modèle binaire local elliptique horizontal (ELBP). Ensuite, la propriété complémentaire d'une paire symétrique d'un ELBP horizontal et vertical est découverte. Ensuite, les avantages de la phase locale de la quantification (LPQ) sont trouvés et exploités.

Nous avons analysé les avantages et les inconvénients des méthodes actuellement utilisées ce qui nous a permis d'explorer plusieurs pistes et concevoir de nouvelles stratégies de reconnaissance. Nous avons proposé plusieurs méthodes ciblées à des représentations robustes du visage contre la diversité des facteurs difficiles qui sont mentionnés ci-dessus. Pour cela, cinq parmi les descripteurs locaux les plus efficaces LBP, LPQ, 3DLBP, TPLBP et FPLBP ont été utilisés afin de fournir plus de puissance discriminative pour les images de profondeur. L'image d'entrée est subdivisée en plusieurs petits blocs rectangulaires et les histogrammes de chaque bloc ont été extraits. Ensuite, tous les histogrammes des différents blocs sont concaténés dans un seul vecteur de caractéristiques. Ce dernier est considéré comme une signature biométrique efficace pour la représentation des personnes enregistrées dans la base de données. De cette manière, l'information locale est représen-

tée (yeux, nez, bouche, front...) à partir de l'image faciale d'une façon discriminante. Malgré l'existence de nombreuses variantes depuis qu'il a été introduit dans la littérature FR, LBP est encore largement utilisé comme descripteur élémentaire pour la construction de nombreuses méthodes d'extraction de caractéristiques faciales avancées. Cela est probablement modifié avec l'avènement de ELBP, notre nouvelle méthode détaillée au chapitre 3, en raison de sa supériorité de performance considérable par rapport à LBP, en particulier sous des obstacles difficiles, tout en ayant la même complexité et le coût informatique (ELBP horizontal) ou n'affecte pas la vitesse trop (lorsque la combinaison d'ELBP horizontal et vertical est utilisée). Sur la base des caractéristiques orientées de l'information faciale et des caractéristiques critiques, comme les yeux et la bouche, nous avons montré que l'ELBP horizontal, en utilisant un échantillon d'ellipse horizontale pour calculer les modèles ELBP, est plus pertinent et efficace pour FR que LBP, qui a été conçue à l'origine pour la classification des textures problème. De plus, en réalisant qu'une paire de ELBP's symétrique (horizontale et verticale) est des descripteurs complémentaires, ils sont ainsi combinés pour capturer des fonctionnalités plus utiles à partir des images de visage. En offrant des résultats prometteurs lors de la connexion du cadre basé sur WPCA et en raison d'une vitesse de traitement rapide, ELBP est susceptible d'extraire des fonctionnalités locales dans des méthodes plus sophistiquées et d'utiliser dans d'autres tâches liées à l'analyse de visage, par exemple la détection de visage et la classification par genre.

Des expériences sur les combinaisons de ces descripteurs au niveau des caractéristiques à l'aide d'une concaténation simple des histogrammes sont effectuées. Afin de réduire la grande dimensionnalité des vecteurs caractéristiques, trois méthodes de réduction de la dimensionnalité sont comparées la PCA, DCT et SVD. Les descripteurs PCA et LPQ conduisent à de hautes performances de vérification ; nous obtenons un taux de vérification de 97,80% et 98,70% pour DCT et LPQ respectivement. Les histogrammes LPQ explorent plus de caractéristiques discriminantes pour la vérification et surpassent toutes les autres caractéristiques des différents descripteurs (TV=98.70% et un HTER=0.80%). Nous attribuons ce résultat au fait que les illuminations et les variations d'expressions faciales n'ont aucun effet sur le domaine fréquentiel qui est représenté par le LPQ. Basés sur ces résultats, nous avons combiné le descripteur LPQ avec tous les autres descripteurs. En plus, nous avons également démontré à travers une étude comparative que la méthode combinée LPQ-FPLBP basée sur DCT est le meilleur chemin pour résoudre le problème des variations d'illuminations et d'expressions faciales grâce à ces performances élevées avec une TV égale à 98.80% et un HTER égal à 0.80%. Nous avons souligné aussi le rôle important du processus de la réduction de dimensions. Nos résultats expérimentaux montrent que la méthode DCT a obtenu un gain plus de 10% dans le TV et un gain de 5,8 % dans le TEE par rapport à la méthode de base PCA. D'autre part, nous pouvons voir qu'il n'y a pas un gain de performance significatif dans les méthodes combinées. Le descripteur LPQ dépasse certaines méthodes combinées telles que : LPQ-TPLBP, 3DLBP et LPQ-LBP. Cela est dû au fait que ces descripteurs partagent beaucoup de caractéristiques similaires qui ont conduit à un affaiblissement du processus de classification.

Pour la vérification 2D les résultats expérimentaux ont montré que notre méthode surpassait de manière significative presque tous les autres systèmes. Même sans l'utilisation d'une technique de normalisation de l'illumination, les performances de précision sont comparables à celles des principaux concurrents. En ce qui concerne l'aspect informatique, la méthode proposée est plus rapide que tous les autres algorithmes avancés basés sur les ondelettes de Gabor.

Grâce à des expériences approfondies dans diverses conditions, LPQ, un descripteur local basé sur la phase dont les détails se trouvent dans le chapitre 3, a été trouvé beaucoup plus robuste que beaucoup d'ondelettes de Gabor, LBP et ses méthodes basées sur les variantes.

- La fusion des descripteurs locaux avec une méthode statistique

La fusion des descripteurs locaux parfois conduit à un affaiblissement du processus de classification, parce qu'ils partagent beaucoup de caractéristiques similaires, alors dans ce cas nous avons une redondance d'informations faciales locales. Ce que nous avons observé lors de la fusion de LPQ avec BSIF. Nous arrivons à un bon résultat lors de l'utilisation du descripteur LPQ séparément par rapport à l'utilisation de la méthode de fusion LPQ-BSIF. Pour cela, nous trouvons intéressante l'idée de se concentrer sur le développement d'une méthode statistique de fusion des descripteurs en tenant compte de la relation de dépendance entre les descripteurs avec laquelle ils sont fusionnés. Une bonne compréhension de ce point devrait permettre de mieux combiner les caractéristiques des deux descripteurs de façon à bénéficier pleinement de leur fusion.

- Comme la partie supérieure du visage (au-dessus du nez) contient une plus grande quantité de caractéristiques faciales importantes que la partie inférieure, il est attendu que cela peut être une suggestion à appliquer pour améliorer le discriminant des descripteurs locaux comme ELBP et LPQ dans lequel le vecteur final est une combinaison de séquences d'histogrammes calculées à partir de sous-régions uniformément divisées. Nous pensons qu'une stratégie de division asymétrique devrait être meilleure.
- En perspective est fin de réduire la taille du vecteur LPQ, d'autre algorithme statistique ou méthodes peuvent être exploités.

Références bibliographiques

- [1] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1):71–86, January 1991. 28, 29, 35, 36, 53, 68, 69.
- [2] M. Grgic, K. Delac, and S. Grgic. Seface – surveillance cameras face database. *Multimedia Tools and Applications*, 51(3):863–879, October 2009. 15, 16, 17, 35, 56, 59, 60, 64, 85, 98, 99, 119, 120, 133, 134, 136, 142, 152, 153
- [3] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):711–720, 1997. 28, 29, 36
- [4] Y. Adini, Y. Moses, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):721–732, 1997. 36, 124
- [5] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(7):179–188, 1936. 36, 39
- [6] R. Gross, I. Matthews, and S. Baker. Fisher light-fields for face recognition across pose and illumination. In L. V. Gool, editor, *Pattern Recognition*, number 2449 in LNCS, pages 481–489. Springer Berlin Heidelberg, January 2002. 36
- [7] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):711–720, 1997. 28, 29, 36
- [8] M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski. Face recognition by independent component analysis. *IEEE Trans. Neural Netw.*, 13(6):1450–1464, November 2002. 37
- [9] J. Yang, D. Zhang, A.F. Frangi, and J. Yang. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(1):131–137, 2004. 37
- [10] J. Yang, D. Zhang, X. Yong, and J.-Y. Yang. Two-dimensional discriminant transform for face recognition. *Pattern Recognition*, 38(7):1125–1129, July 2005. 37

- [11] T. Ojala, M. Pietikäinen, and T. Maenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, 2002. 40
- [12] A. Hadid, M. Pietikäinen, and T. Ahonen. A discriminative feature space for detecting and recognizing faces. In *CVPR*, volume 2, pages 797–804, 2004. 41
- [13] [33] X. Feng, M. Pietikäinen, and A. Hadid. Facial expression recognition based on local binary patterns. *Pattern Recognition and Image Analysis*, 17(4):592–598, December 2007. 41
- [14] P. Karthigayani and S. Sridhar. A novel approach for face recognition and age estimation using local binary pattern, discriminative approach using two layered back propagation network. In *Proc. Trendz in Information Sciences and Computing (TISC)*, pages 11–16, 2011. 41
- [15] H. C. Lian and B. L. Lu. Multi-view gender classification using local binary patterns and support vector machines. *Advances in Neural Networks*, pages 202– 209, 2006. 41
- [16] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. *ECCV*, pages 469–481, 2004. 29, 40, 41, 42, 44, 45, 47, 53, 66, 76, 87, 89, 94, 100, 139
- [17] G. Zhang, X. Huang, S. Li, Y. Wang, and X. Wu. Boosting local binary pattern (lbp)-based face recognition. In *Advances in Biometric Person Authentication*, pages 179–186. 2005. 42, 43
- [18] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational learning theory*, pages 23–37. Springer, 1995. 42
- [19] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Li. Learning multi-scale block local binary patterns for face recognition. *Advances in Biometrics*, pages 828–837, 2007. 42
- [20] H. Jin, Q. Liu, H. Lu, and X. Tong. Face detection using improved LBP under bayesian framework. In *Proc. Image and Graphics*, pages 306–309, 2004. 42
- [21] M. Heikkilä, M. Pietikäinen, and C. Schmid. Description of interest regions with center-symmetric local binary patterns. *Computer Vision, Graphics and Image Processing*, pages 58–69, 2006. 42
- [22] J. Choi, W.R. Schwartz, H. Guo, and L.S. Davis. A complementary local feature descriptor for face identification. In *WACV*, pages 121–128, January 2012. 42, 83, 96, 117, 131, 148

- [23] W. Schwartz, H. Guo, and L. Davis. A robust and scalable approach to face identification. *ECCV*, pages 476–489, 2010. 42
- [24] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In *AMFG*, volume 4778, pages 168–182. Springer Berlin Heidelberg, January 2007. 42, 65
- [25] J. Ren, X. Jiang, and J. Yuan. Relaxed local ternary pattern for face recognition. In *ICIP*, pages 3680–3684, Melbourne, VIC, 2013. IEEE. 42
- [26] S. Liao and A.C.S. Chung. Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude. In *ACCV*, pages 672–679, Berlin, Heidelberg, 2007. Springer-Verlag. 42, 75, 83
- [27] D. Maturana, D. Mery, and A. Soto. Learning discriminative local binary patterns for face recognition. In *FG*, pages 470–475, 2011. 43, 83, 96, 131.
- [28] N. S. Vu and A. Caplier. Face recognition with patterns of oriented edge magnitudes. *ECCV*, pages 313–326, 2010. 43, 125, 126, 141
- [29] N. S. Vu and A. Caplier. Mining patterns of orientations and magnitudes for face recognition. In *IJCB*, pages 1–8, October 2011. 43, 96, 117, 131, 141, 148
- [30] G. Zhang, X. Huang, S. Li, Y. Wang, and X. Wu. Boosting local binary pattern (lbp)-based face recognition. In *Advances in Biometric Person Authentication*, pages 179–186. 2005. 42, 43
- [31] D. Maturana, D. Mery, and A. Soto. Learning discriminative local binary patterns for face recognition. In *FG*, pages 470–475, 2011. 43, 83, 96, 131
- [32] J. W. Shepherd, G. M. Davies, and A. W. Ellis. Studies of cue saliency. In G. Davies, H. Ellis, and J. Shepherd, editors, *Perceiving and Remembering Faces*, pages 105–131. New York: Academic Press, 1981.
- [33] V. Goffaux and S.C. Dakin. Horizontal information drives the behavioral signatures of face processing. *Frontiers in Psychology*, 1, 2010.
- [34] V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. *Image and Signal Processing*, pages 236–243, 2008.
- [35] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä. Recognition of blurred faces using local phase quantization. In *ICPR*, pages 1–4, 2008.
- [36] Z. Lei, T. Ahonen, M. Pietikainen, and S. Z. Li. Local frequency descriptor for low-resolution face recognition. In *FG*, pages 161–166. IEEE, 2011.

- [37] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang. Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. In ICCV, volume 1, pages 786–791, 2005.
- [38] S. Shan, W. Zhang, Y. Su, X. Chen, and W. Gao. Ensemble of piecewise FDA based on spatial histograms of local (Gabor) binary patterns for face recognition. In ICPR, volume 4, pages 606–609, 2006.
- [39] B. Zhang, S. Shan, X. Chen, and W. Gao. Histogram of gabor phase patterns (HGPP): a novel object representation approach for face recognition. IEEE Trans. Image Process., 16(1):57–68, January 2007.
- [40] X. Tan and B. Triggs. Fusing gabor and LBP feature sets for Kernel-Based face recognition. In AMFG, volume 4778, pages 235–249. Springer Berlin Heidelberg, 2007.
- [41] X. Shufu, S. Shiguang, C. Xilin, and C. Jie. Fusing local patterns of gabor magnitude and phase for face recognition. IEEE Trans. Image Process., 19(5):1349–1361, May 2010.
- [42] K. Yan, Y. Chen, and D. Zhang. Gabor surface feature for face recognition. In 1st ACPR, pages 288–292, November 2011.
- [43] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss. The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. Pattern Anal. Mach. Intell., 22(10):1090–1104, 2000.
- [44] A. M. Martinez and R. Benavente. The AR face database. CVC Technical Report, 24, 1998.
- [45] S. Shan, W. Zhang, Y. Su, X. Chen, and W. Gao. Ensemble of piecewise FDA based on spatial histograms of local (Gabor) binary patterns for face recognition. In ICPR, volume 4, pages 606–609, 2006.
- [46] B. Zhang, S. Shan, X. Chen, and W. Gao. Histogram of gabor phase patterns (HGPP): a novel object representation approach for face recognition. IEEE Trans. Image Process., 16(1):57–68, January 2007.
- [47] X. Tan and B. Triggs. Fusing gabor and LBP feature sets for Kernel-Based face recognition. In AMFG, volume 4778, pages 235–249. Springer Berlin Heidelberg, 2007.
- [48] H. Cevikalp, M. Neamtu, and M. Wilkes. Discriminative common vector method with kernels. IEEE Trans. Neural Netw., 17(6):1550–1565, November 2006.

- [49] X. Shufu, S. Shiguang, C. Xilin, and C. Jie. Fusing local patterns of gabor magnitude and phase for face recognition. *IEEE Trans. Image Process.*, 19(5):1349–1361, May 2010.
- [50] K. Yan, Y. Chen, and D. Zhang. Gabor surface feature for face recognition. In 1st ACPR, pages 288–292, November 2011.
- [51] S. U. Hussain, T. Napoleon, and F. Jurie. Face recognition using local quantized patterns. In BMVC, September 2012.
- [52] M. Yang, L. Zhang, S. C.-K. Shiu, and D. Zhang. Robust kernel representation with statistical local features for face recognition. *IEEE Trans. Neural Netw.*, pages 1–1, 2013.
- [53] Z. Chai, Z. Sun, H. Mendez-Vazquez, R. He, and T. Tan. Gabor ordinal measures for face recognition. *IEEE Trans. Inf. Forensics Security*, 9(1):14–26, January 2014.
- [54] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, November 2004.
- [55] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, volume 1, pages 886–893, 2005.
- [56] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli. On the use of sift features for face authentication. In CVPR Workshop, 2006.
- [57] C. Rosenberger and L. Brun. Similarity-based matching for face authentication. In ICPR, pages 1–4, 2008.
- [58] E. Bailly-Baillié, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, and F. Porée. The BANCA database and evaluation protocol. In *Audio-and Video-Based Biometric Person Authentication*, pages 625–638. Springer, 2003.
- [59] A. M. Martinez and R. Benavente. The AR face database. CVC Technical Report, 24, 1998.
- [60] E. Meyers and L. Wolf. Using biologically inspired features for face processing. *Int Journal of Computer Vision*, 76:93–104, 2008.
- [61] J. Yang, J. Wright, T. S. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE Trans. Image Process.*, 19(11):2861–2873, 2010.
- [62] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1):71–86, January 1991.

- [63] W.W.W. Zou and P.C. Yuen. Very low resolution face recognition problem. In BTAS, pages 1–6, 2010.
- [64] P. Jonathon Phillips, Hyeonjoon Moon, Syed A. Rizvi, and Patrick J. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(10):1090–1104, 2000.
- [65] Marcio Teixeira Ross Beveridge, David Bolme and Bruce Draper. The CSU Face Identification Evaluation System User's Guide: Version 5.0., 2003.
- [66] P. Jonathon Phillips, Patrick J. Flynn, W. Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William J. Worek. Overview of the face recognition grand challenge. In *CVPR (1)*, pages 947–954, 2005.
- [67] K. Messer, J. Matas, J. Kittler, J. Luttin, and G. Maitre. XM2VTSDB: The extended M2VTS database. In *Audio- and Video-based Biometric Person Authentication, AVBPA'99*, pages 72–77, 1999. Washington, D.C., March 1999. 16 IDIAP-RR 99-02
- [68] P. Jonathon Phillips, Hyeonjoon Moon, Patrick Rauss, and Syed A. Rizvi. The feret evaluation methodology for face-recognition algorithms. In *cvpr*, page 137, Los Alamitos, CA, USA, 1997. IEEE Computer Society.
- [69] Fatima Boukhris, Abdallah Boumalk, El Houssain El Moujahid, Hamid Souifi, "La Nouvelle Grammaire de l'Amazighe", Centre de l'Aménagement Linguistique, Publications de l'IRCAM, Rabat, 2008.
- [70] Proposition d'ajout de l'écriture Tifnaghe au répertoire de l'ISO/CEI 10646 (format Unicode), 21/06/2004, CEISIC, IRCAM, Rabat, Maroc.
- [71] Abdessamad Malaoui, et al. "A new approach of a precise electric modeling of the semiconductors and dielectric samples." *The IEEE International Conference on Microelectronics*, pp 226-229, Marrakech, (2009).
- [72] E. Barsoukov, J. Ross Macdonald, *Impedance Spectroscopy Theory*, John Wiley and Sons, Inc., New Jersey, 2005.
- [73] N. Rizoug. Thèse de Doctorat, l'Ecole Centrale de Lille et l'Université des Sciences et Technologies de Lille, 2006
- [74] S. Skales, V. Dolecek, M. Slemnik. *Corrosion Science Journal*, 49, No3, pp. 1045-1055, 2007.
- [75] A. Malaoui, M. Mabrouki, M. Ankrim. A new precise method of electric modeling of thin film. *M. J. CONDENSED MATTER*, Vol 12, N 1, 2010

- [76] V. Ojansivu and J. Heikkilä, "Blur Insensitive Texture Classification Using Local Phase Quantization," in *Image and Signal Processing*. vol. 5099, A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 236-243.
- [77] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 971 -987, 2002.
- [78] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor based methods in the wild," in *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, 2008.
- [79] M. Yue, R. Qiuqi, W. Xueqiao, and M. Meiru, "Robust 3D face recognition using learn correlative features," in *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*, 2010, pp. 1382-1385.
- [80] C. Xu, S. Li, T. Tan, and L. Quan, "Automatic 3D face recognition from depth and intensity Gabor features," *Pattern Recognition*, vol. 42, pp. 1895-1905, 2009
- [81] W. Chen, M. J. Er, and S. Wu, "PCA and LDA in DCT domain," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2474-2482, 2005.
- [82] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1) :71-86, January 1991.
- [83] Vapnik, V. N. (2000). *The nature of statistical learning theory*.
- [84] J.R. Beveridge, K. She, B.A. Draper, and G.H. Givens. A nonparametric statistical comparison of principal component and linear discriminant subspaces for face recognition. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I-535-I-542 vol.1, 2001.
- [85] M. Turk, A. Pentland, Eigenfaces for recognition. *J. of Cognitive Neuroscience* 3, 72-86, 1991.
- [86] Freeman H. : On the encoding of arbitrary geometric configurations, *IRE Trans. Electron. Comput.* EC-10, 260-268, 1961.
- [87] Freeman H. : *Computer Processing of Line-Drawing Images*, *Computing surveys*, Vol 6, No 1, March 1974.
- [88] M. Turk, A. Pentland, Eigenfaces for recognition. *J. of Cognitive Neuroscience* 3, 72-86, 1991.

- [89] Belahcene, M., Ouamane, A., Boumehrez, M., Benakcha, A. (2014). Comparaison des méthodes de réduction d'espace et l'application des SVMs pour la classification dans l'authentification de visages.
- [90] OUAMANE, H. (2014). Identification de reconnaissance faciale avec des expressions (Master's thesis).
- [91] Miller, W. H., Seefeld, M. A., Newlander, K. A., Uzinskas, I. N., Burgess, W. J., Heerding, D. A., ... Moore, T. D. (2002). Discovery of aminopyridine-based inhibitors of bacterial enoyl-ACP reductase (FabI). *Journal of medicinal chemistry*, 45(15), 3246-3256.
- [92] Mian, A., Bennamoun, M., Owens, R. (2007). An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 29(11).
- [93] Bertrand, P. (1998). Comparaison de l'erreur quadratique moyenne intégrée pour différents estimateurs du coefficient de diffusion d'un processus. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 327(4), 399-404.
- [94] Harkat, M. F. (2003). Détection et localisation de défauts par analyse en composantes principales (Doctoral dissertation, Institut National Polytechnique de Lorraine-INPL).
- [95] Prudent, Y., Ennaji, A. (2004). Clustering incrémental pour un apprentissage distribué: vers un système évolutif et robuste. In *Conférence CAP*.
- [96] Girard, S., Chalmond, B., Dinten, J. M. (1998). Une ACP non linéaire basée sur l'approximation par variété. *Revue de statistique appliquée*, 46(3), 5-19.
- [97] Casin, P., Stachowiak, C., Marque, F. (2011). L'analyse en composantes principales de variables non stationnaires. *Math. Sci. Hum/Mathematics and Social Sciences*, 196(4), 27-40.
- [98] El Mahdi Barraha, S. S., Malaoui, A. (2015). Extended Set of *DCT – TPLBP and DCT – FPLBP* for Face Recognition. *World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 9(8).
- [99] Barraha, E. M., Ahdid, R., Safi, S., Malaoui, A. (2015). New Technique to Determination of Electronic Circuits for Semiconductor Components by Recognizing Nyquist Curve. *Journal of Computer Sciences and Applications*, 3(4), 100-104.
- [100] Barraha, E. M., Ahdid, R., Safi, S., Malaoui, A. (2017). 3DLBP and SVD Fusion for 3D Face Recognition Using Range Image. *Computer Science and Information Technology*, 5(2), 61-65.

- [101] Barrah, E. M., Safi, S., Malaoui, A. (2016, May). New technique for face recognition based on Singular Value Decomposition (SVD). In *Cloud Computing Technologies and Applications (CloudTech)*, 2016 2nd International Conference on (pp. 96-99). IEEE.