



Université Mohamed Khider de Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique

MÉMOIRE DE MASTER

Sciences et Technologies
Electronique
Systèmes embarqués

Réf. : ...

Présenté et soutenu par :
BELHADJ Mahdi

Le : mardi 9 juillet 2019

Etude et simulation d'un système de reconnaissance des expressions faciale

Jury :

Mme. TOUMI Abida	MCA	Université de Biskra	Président
Mr. OUAFI Abdelkrim	PR	Université de Biskra	Encadreur
Mme. BARKAT Aicha	MAA	Université de Biskra	Examineur

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement Supérieur et de la recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Electronique
Option : Systèmes embarqués

Mémoire de Fin d'Etudes
En vue de l'obtention du diplôme :

MASTER

Thème

**Etude et simulation d'un système de
reconnaissance des expressions faciale**

Présenté par :

BELHADJ Mahdi

Avis favorable de l'encadreur :

OUAFI Abdelkrim

Signature

Avis favorable du Président du Jury

Nom et prénom

TOUMI Abida

Signature

Cachet et signature

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement Supérieur et de la Recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Electronique
Option : Système embarqué

Thème :

**Etude et simulation d'un système de
reconnaissance des expressions faciale**

Proposé par : Pr. OUAFI Abdelkrim
Dirigé par : Pr. OUAFI Abdelkrim

Dédicace

*A tout respect et amour je dédie ce travail
à mes très chers parents, à mes frères et
sœurs, et tous mes amis.*

Remerciement

En tout premier lieu, je remercie « Allah » tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés.

Je tiens tout d'abord à témoigner toute ma reconnaissance à Mr OUAFI Abdelkrim professeur de l'Université de Mohamed Khider Biskra, pour son appui scientifique, sa disponibilité et ses conseils qu'il m'a apporté lors de l'encadrement de mon mémoire.

Je remercie l'ensemble des membres du jury qui m'ont fait l'honneur d'accepter de lire et de juger ce mémoire : Dr. TOUMI Abida, et, Dr. BARKAT Aïcha.

J'aimerais exprimer ma gratitude à tous les enseignants, de l'université de Mohamed Khider Biskra, qui m'ont inculqué les fondements scientifiques tout au long de mon cursus universitaire

Je remercie également mes parents de m'avoir toujours soutenu pendant toutes mes années d'études. Mes sincères remerciements vont aussi à mes frères et à mes sœurs, et tous mes amis.

Liste des tableaux

Tableau1.1 : Les avantages et les inconvénients de différentes modalités comportementales	12
Tableau1.2 : Les avantages et les inconvénients de différentes modalités morphologiques.. ..	13
Tableau1.3 : Les avantages et les inconvénients de différentes modalités biologiques	13
Tableau 3.1 : La matrice de confusion utilisant descripteur HOG et le classifieur SVM (sans égalisation d'histogramme).....	44
Tableau 3.2 : La matrice de confusion utilisant descripteur HOG et le classifieur SVM (avec égalisation d'histogramme).....	44
Tableau 3.3 : La matrice de confusion utilisant descripteur HOG et le classifieur euclidienne (sans égalisation d'histogramme).....	44
Tableau 3.4 : La matrice de confusion utilisant descripteur HOG et le classifieur euclidienne (avec égalisation d'histogramme)	45
Tableau 3.5 : La matrice de confusion utilisant descripteur LBP et le classifieur SVM (sans égalisation d'histogramme).....	46
Tableau 3.6 : La matrice de confusion utilisant descripteur LBP et le classifieur SVM (avec égalisation d'histogramme).....	46
Tableau 3.7 : La matrice de confusion utilisant descripteur LBP et le classifieur euclidienne (sans égalisation d'histogramme).....	47
Tableau 3.8 : La matrice de confusion utilisant descripteur LBP et le classifieur euclidienne (avec égalisation d'histogramme)	47
Tableau 3.9 : Comparaison des performances de notre approche par rapport à différentes approches de l'état de l'art.....	49

Liste des figures

Chapitre 1 Principe de la reconnaissance des expressions faciales

Figure 1.1 : Un système biométrique [4].....	6
Figure 1.2 : Image d'empreinte digitale [6].....	8
Figure 1.3 : Image visage	9
Figure 1.4 : Image d'iris [7].....	9
Figure 1.5 : Image géométrie de la main.....	10
Figure 1.6 : Image signature	10
Figure 1.7 : Principe de reconnaissance vocale.....	11
Figure 1.8 : Image de veines de la main.....	12
Figure 1.9 : Architecture d'un système de reconnaissance des expressions faciales	15
Figure 1.10 : Les six émotions universelles [14]	16
Figure 1.11 : Roue des émotions de Plutchik [16].....	17
Figure 1.12 : L'opérateur LBP	18
Figure 1.13 : Exemple d'histogramme obtenu dans méthode de HOG [22].....	19
Figure 1.14 : Ensemble des points caractéristiques ASM model [31]	21

Chapitre 2 Méthodes de reconnaissance des expressions faciales

Figure 2.1 : Architecture d'un système de reconnaissance des expressions faciales	24
Figure 2.2 : L'opérateur LBP	25
Figure 2.3 : Un ensemble voisinage d'opérateur LBP [37]	26
Figure 2.4 : Schéma d'obtention d'histogramme LBP.....	26
Figure 2.5 : Un exemple d'opérateur ILBP	27
Figure 2.6 : Normalisation des blocs (rectangle bleu représenter un bloc)	30
Figure 2.7 : Différents types des séparateurs.....	32
Figure 2.8 : Hyperplan optimal	33

Chapitre 3 Simulation et résultats

Figure 3.1 : Architecture du système de reconnaissance des expressions faciales.....	37
Figure 3.2 : Prétraitement d'image d'entrée. (a) image en niveau de gris, (b) détection du visage, (c) visage coupé.....	38
Figure 3.3 : Exemple d'égalisation d'histogramme.....	39

Figure 3.4 : Visualisation de descripteur HOG d'une image avec différents tailles de la cellule	40
Figure 3.5 : Extraction des caractéristiques avec l'opérateur LBP selon la taille des cellules	41
Figure 3.6 : Exemples de 6 expressions de la base de données JAFFE	42
Figure 3.7 : Représentation graphique des résultats avec l'opérateur HOG (score)	45
Figure 3.8 : Représentation graphique des résultats avec l'opérateur LBP (score)	47
Figure 3.9 : Représentation graphique des meilleurs résultats de descripteurs HOG et LBP avec les six émotions.....	48
Figure 3.10 : Organigramme du système REF	50
Figure 3.11 : Interface de l'application de reconnaissance des expressions faciales.....	51

Liste des abréviations

AAM : Active appearance model.

ADN : Acide désoxyribonucléique.

ASM : Active shape model.

BDD : Base de données.

CLBP: Complete local binary patterns.

ELBP: Extended local binary patterns.

EUC: Distance euclidienne.

HOG: Histogram of oriented gradients.

ILBP: Improved local binary patterns.

JAFFE: Japanese Female Facial Expression.

KNN: K-nearest neighbors.

LBP : Local binary pattern.

LDB: Local difference binary.

LPQ: Local phase quantification.

MATLAP : Matrix laboratory.

MPD : Modèle de distribution de points.

PHOG: Pyramid of histogram of oriented gradients.

REF : Reconnaissance des expressions faciales.

RGB : Rouge vert bleu.

SVM : Support vector machine.

TFCT : Transformée de Fourier à court terme.

VLBP: Volume local binary patterns.

Résumé

Ce mémoire présente une étude du système de reconnaissance des expressions faciales. Nous présenterons un état de l'art des approches récentes dans ce domaine, et nous allons être guidés par une étude approfondie de ces techniques en termes d'efficacité et de performance. Afin de sélectionner les descripteurs appropriés et efficaces pour la mise en œuvre de notre système.

Notre système est basé sur trois phases principales. La première étape c'est la prétraitement, où après la détection du visage, l'image est préparée en éliminant les effets de la lumière et les parasites. L'étape suivante c'est l'extraction des caractéristiques qui identifie et extrait les informations des zones du visage qui nous intéressent. La troisième étape c'est la classification, où l'expression faciale de l'image est classifiée, à partir des informations extraites à l'étape précédente.

Enfin, nous avons essayé d'obtenir des meilleurs résultats en effectuant plusieurs tests avec différents descripteurs : HOG et LBP, et différents types de classificateurs : SVM et euclidienne, en utilisant de la base de données JAFFE.

Mots clés : Reconnaissance des émotions, Expressions Faciales, SVM, HOG, LBP, JAFFE, Extraction des caractéristiques, Prétraitement, Classification.

المخلص

تقدم هذه الرسالة دراسة لنظام التعرف على تعبيرات الوجه. حيث سنقدم مراجعة عامة لأحدث الأساليب في هذا المجال، كما ستقودنا دراسة شاملة لهذه التقنيات من حيث كفاءتها واداءها، الى اختيار الواصفات المناسبة والفعالة لتنفيذ وتحقيق نظامنا.

نظامنا مبني على ثلاث مراحل اساسية، المرحلة الأولى هي ما قبل المعالجة، حيث بعد كشف الوجه وتحديده، يتم إعداد الصورة المتحصل عليها من خلال القضاء على التشويش وأثار الضوء، المرحلة التالية هي استخراج الخصائص حيث تحدد وتستخرج المعلومات من مناطق الوجه التي تهتمنا، والمرحلة الثالثة هي التصنيف، حيث يتم تصنيف تعبير الوجه للصورة، اعتمادا على المعلومات المستخرجة في الخطوة السابقة.

وفي الأخير، حاولنا الحصول على أفضل النتائج من خلال إجراء العديد من الاختبارات باستخدام واصفات مختلفة: HOG و LBP، وأنواع مختلفة من المصنفات SVM: والمصنف الإقليدي، باستخدام قاعدة بيانات JAFFE.

Sommaire

Liste des tableaux	iii
Liste des figures	iv
Liste des abréviations	vi
Résumé	vii
Introduction générale	1
Chapitre 1 Principe de la reconnaissance des expressions faciales	
1.1 Introduction	4
1.2 Définition de la Biométrie	5
1.3 Définition d'un système biométrique	5
1.4 Modes de fonctionnement d'un système biométrique.....	5
1.5 Schéma de fonctionnement d'un système biométrique.....	6
1.6 Principaux modules d'un système biométrique	7
1.7 Les modalités biométriques	7
1.7.1. Modalités morphologiques (physiologiques).....	7
1.7.2 Modalités comportementales.....	10
1.7.3 Modalités biologiques	11
1.7.4 Avantages et inconvénients de différentes modalités biométriques.....	12
1.8 Reconnaissance Des Expressions Faciales.....	14
1.8.1 Historique sur la reconnaissance de l'expression faciale	14
1.8.2 Système de reconnaissance des expressions faciales	14
1.8.3 Les 6 expressions faciales universelles	15
1.8.4 Différentes méthodes de reconnaissance des expressions faciales	17
1.8.4.1 Approches d'apparence	17
1.8.4.2 Approches géométriques	20
1.9 Conclusion.....	21
Chapitre 2 Méthodes de reconnaissance des expressions faciales	
2.1 Introduction	23

2.2	Système de reconnaissance des expressions faciales	24
2.3	Prétraitement.....	24
2.4	Extraction des caractéristiques	25
2.4.1	Motif binaire local (LBP : Local Binary Pattern).....	25
2.4.1.1	Concept de LBP	25
2.4.1.2	différentes extensions de LBP	27
2.4.1	Histogramme de gradient orienté (HOG : Histogram of oriented gradients)	28
2.4.1.1	Concept de HOG	28
2.4.1.2	Différentes extensions de HOG	31
2.5	Classification	31
2.5.1	Machine à vecteurs de support (Support vector machine) SVM.....	32
2.5.2	Distance euclidienne.....	33
2.6	Conclusion	34

Chapitre 3 Simulation et résultats

3.1	Introduction	36
3.2	Système de reconnaissance des expressions faciales	37
3.2.1	Prétraitement.....	38
3.2.2	Extraction des caractéristiques	39
3.2.2.1	Extraction des caractéristiques en utilisant HOG.....	39
3.2.2.2	Extraction des caractéristiques en utilisant LBP	40
3.2.3	Classification	42
3.3	Base de données	42
3.4	Résultats expérimentaux et discussions.....	43
3.4.1	Résultats obtenus	43
3.4.1.1	Histogramme de gradient orienté (HOG)	43
3.4.1.2	Motif binaire local (LBP)	46
3.5	Implémentation en temps réel	49
3.6	Conclusion	52
	Conclusion général	53
	Bibliographie	55
	Annexe	60

Introduction générale

La communication homme-machine a beaucoup évolué ces dernières années, passant de la communication via des supports électroniques tels que le clavier, la souris et l'écran à la communication directe entre l'appareil et le corps humain.

La main, le visage, les doigts, les yeux, la voix, ce sont les modalités biométriques plus utilisées dans les systèmes de la communication moderne, qui assure l'harmonie entre le dispositif électronique et le corps humain, en fonction de la capture du comportement observable de l'utilisateur, en utilisant des techniques d'intelligence artificielle, y compris la vision par ordinateur.

Le visage étant la partie la plus expressive et communicative d'un être humain, parce qu'il peut lui montrer différentes expressions émotionnelles exprimant l'émotion intérieure de la personne. Les expressions faciales provoquent des changements physiologiques sur le visage, tels que le mouvement des sourcils, la position de la bouche ouverts ou fermés, ou encore la manière de regarder fixement les yeux. Selon **Paul Ekman**, il existe six expressions universelles : la colère, le dégoût, la joie, le peur, la surprise, la tristesse, convenues dans toutes les races et les cultures. Il a également été constaté que les personnes aveugles de naissance ont aussi les mêmes expressions bien qu'elles n'aient jamais vues le visage des autres.

La capacité de l'ordinateur à reconnaître les expressions faciales de la personne constitue un nouveau défi pour la recherche scientifique moderne. D'autant plus que la communication entre humaine et les appareils électroniques a augmenté considérablement, alors que les chercheurs cherchent à développer des programmes intelligents capables de comprendre les expressions faciales en peu de temps,

Les techniques d'identification des émotions de la personne ne se limitent seulement plus sur l'état psychologique, elles sont également utilisées dans le domaine médical comme assistant du traitement des patients atteints d'autisme qui ont des difficultés à comprendre les émotions des autres et à communiquer avec eux, en plus d'aider les enseignants dans l'apprentissage à distance via internet. Elle est également utilisée pour développer aussi la technologie des jeux informatiques en révélant les émotions des joueurs et utiliser leurs réactions pour modifier le jeu ou leurs personnages virtuels.

Introduction générale

Ce mémoire traite un sujet de la reconnaissance des expressions faciales. Le premier chapitre est consacré à la présentation générale de la biométrie. Il décrit le principe de fonctionnement des systèmes biométriques et ses différentes modalités (morphologiques, comportementales, et biologiques), ensuite, les différentes techniques biométriques tels que les empreintes digitales, le visage, la rétine sont analysées en termes de leurs avantages et inconvénients. Après cela, nous nous intéresserons plus particulièrement sur le visage en examinant le principe de fonctionnement de système de reconnaissance des expressions faciales. Puis nous avons décrit en particulier les différentes méthodes d'analyse des expressions faciales, basées sur l'analyse de l'apparence ou sur l'analyse géométrique du visage, ainsi que leur efficacité dans l'identification des émotions. Nous concluons le chapitre en sélectionnant deux des méthodes étudiées pour adoption dans la suite de l'étude.

Le deuxième chapitre est consacré à l'étude des différentes étapes du système de reconnaissance faciale et les caractéristiques de chaque étape : prétraitement, extraction des caractéristiques, et classification, où nous allons donner une explication théorique des méthodes utilisées pour l'extraction des caractéristiques faciales : HOG et LBP, et les deux méthodes utilisées pour classifier les expressions faciales.

Le dernier chapitre porte sur l'implémentation et conception de notre système, où nous présenterons la base de données utilisée, aussi un ensemble des tests est réalisées, en fonction des paramètres de l'étape de prétraitement, et selon le descripteur et le classifieur utilisés. À la fin de ce chapitre, nous présentons les résultats obtenus et discutés.

Nous terminerons enfin par une conclusion générale dans laquelle nous résumerons le travail réalisé, et nous donnons quelques suggestions pour améliorer et développer ce système.

Chapitre 1

Principe de la reconnaissance des expressions faciales

1.1 Introduction	4
1.2 Définition de la Biométrie	5
1.3 Définition d'un système biométrique	5
1.4 Modes de fonctionnement d'un système biométrique.....	5
1.5 Schéma de fonctionnement d'un système biométrique.....	6
1.6 Principaux modules d'un système biométrique	7
1.7 Les modalités biométriques	7
1.7.1. Modalités morphologiques (physiologiques).....	7
1.7.2 Modalités comportementales.....	10
1.7.3 Modalités biologiques	11
1.7.4 Avantages et inconvénients de différentes modalités biométriques.....	12
1.8 Reconnaissance Des Expressions Faciales.....	14
1.8.1 Historique sur la reconnaissance de l'expression faciale	14
1.8.2 Système de reconnaissance des expressions faciales	14
1.8.3 Les 6 expressions faciales universelles	15
1.8.4 Différentes méthodes de reconnaissance des expressions faciales	17
1.8.4.1 Approches d'apparence	17
1.8.4.2 Approches géométriques	20
1.9 Conclusion.....	21

1.1 Introduction

Avec le développement technologique rapide, toutes nos données personnelles ont été sauvegardées dans des formats numériques, que ce soit dans des ordinateurs ou des appareils intelligents. Il est donc nécessaire de sécuriser nos informations contre le vol, et pour cette raison est apparu ce que nous appelons les systèmes biométriques.

Dans ce chapitre, on donne une idée générale sur la biométrie et ses différentes modalités, et à quoi sert un système biométrique, puis nous allons nous concentrer sur le visage et plus précisément le domaine de la reconnaissance des expressions faciales, tels que le processus de système de reconnaissance des expressions faciales (REF), et les différentes techniques existantes et leur efficacité pour déterminer l'émotion de la personne.

1.2 Définition de la Biométrie

La biométrie est une technique qui mesure des caractéristiques physiques (empreintes digitales, visage, iris, contour de la main, etc.) ou comportementales (signature, démarche, etc.) d'une personne pour déterminer son identité [1].

Chaque personne a des caractéristiques physiques unique et personnalisé, qui distinguent des autres (même pour les jumeaux).

1.3 Définition d'un système biométrique

Un système biométrique est essentiellement un système de reconnaissance de formes qui utilise les données biométriques d'un individu, extrayant un ensemble de caractéristiques à partir des données acquises, ces caractéristiques sont comparées à la signature dans la base de données.

On peut dire qu'un système de contrôle biométrique, est un système automatique de mesure basé sur la reconnaissance de caractéristiques propres à l'individu [2].

1.4 Modes de fonctionnement d'un système biométrique

On peut définir trois modes de fonctionnement d'un système biométrique :

a) Enrôlement

L'enrôlement est la première phase de tout système biométrique. Il s'agit de l'étape pendant laquelle un utilisateur est enregistré dans le système pour la première fois.

Pendant l'enrôlement, la caractéristique biométrique est mesurée en utilisant un capteur biométrique afin d'extraire une représentation numérique. Cette représentation est ensuite réduite, en utilisant un algorithme d'extraction bien défini, afin de réduire la quantité de données à stocker et ainsi faciliter la vérification et l'identification.

De l'application et du niveau de sécurité souhaité, le modèle biométrique retenu, est stocké soit dans une base de données centrale, soit sur un élément personnel propre à chaque personne [3].

b) Identification

Dans le mode d'identification le système biométrique doit répondre à la question «Qui je suis?» , pour déterminer l'identité d'un individu inconnu. Dans ce cas le système comparer

l'échantillon biométrique d'un utilisateur avec tous les échantillons biométriques contenus dans la base de données biométrique (BDD) du système, (vérification 1 : N). Si l'échantillon correspond à celui d'une personne de la base ou échoue, si l'individu n'a pas une identification dans la base des données biométriques.

c) Vérification

Dans ce cas le système répond à la question « suis-je ce que je prétends être ? ».Par oui ou non c'est à dire le système compare entre l'échantillon qui présente l'utilisateur (un identifiant « Id »), et les donnée stockée dans sa propre base (les modèles d'apprentissage).Donc c'est un test 1 : 1.Le système doit vérifier que l'identité de la personne est bien celle proposée par l'utilisateur.

1.5 Schéma de fonctionnement d'un système biométrique

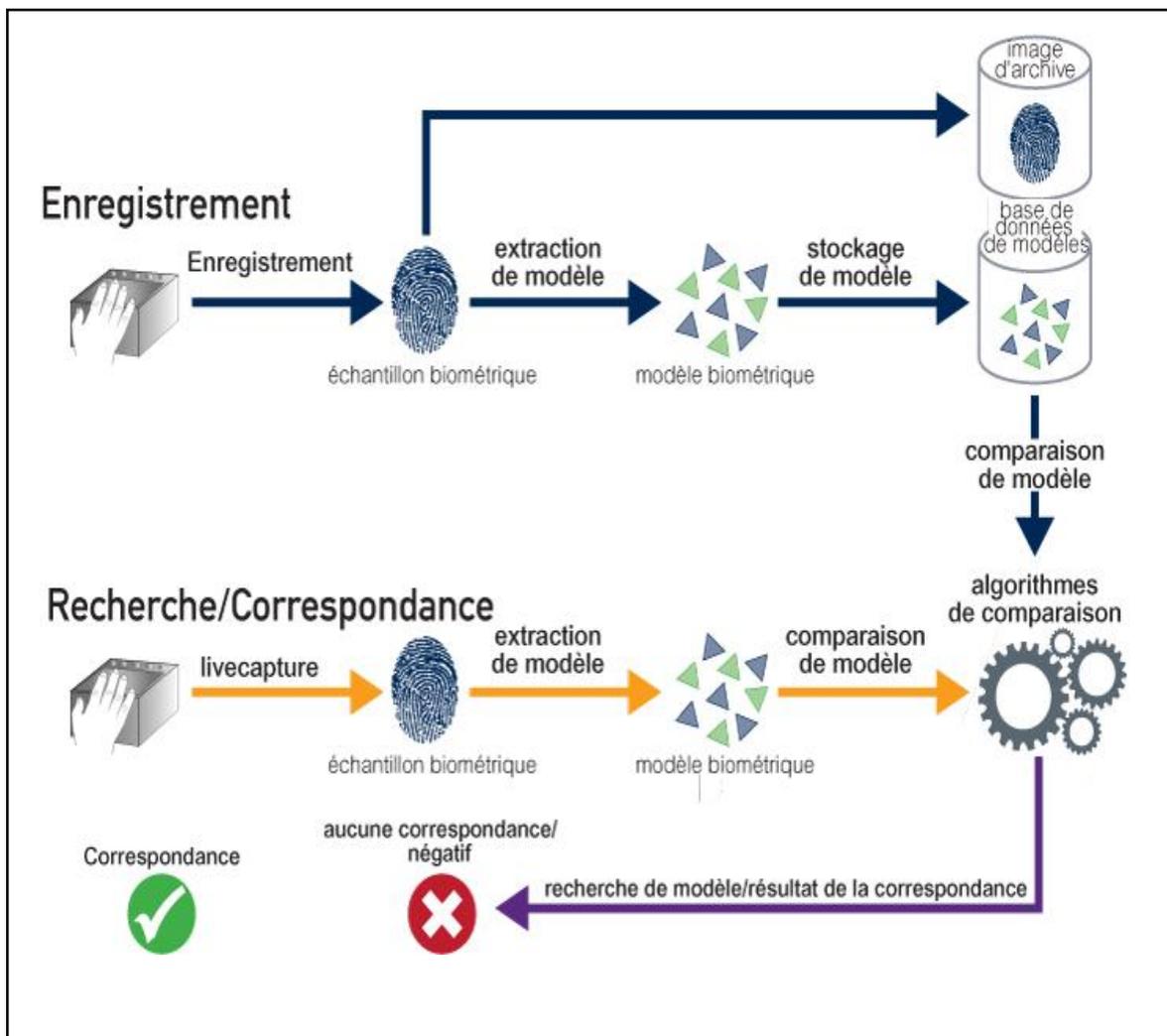


Figure1.1 : Un système biométrique [4].

1.6 Principaux modules d'un système biométrique

a) Module de capture

Correspond à la lecture des données biométriques (caractéristiques physiologiques, comportementales ou biologiques) d'un individu à l'aide de moyen d'un capteur biométrique (un appareil photo, un lecteur d'empreintes digitales, une caméra de sécurité,... etc.), afin d'extraire une représentation numérique. Cette représentation est ensuite utilisée pour l'enrôlement, la vérification ou l'identification.

b) Module extraction des données

Permet l'obtention des données biométriques acquises par le module de capture et d'extraction d'informations. Il extrait seulement l'information pertinente, afin de former une nouvelle représentation des données, ce module peut avoir un test de qualité pour contrôler les données biométriques acquises (extraction les informations pertinentes à partir des données biométriques brutes).

c) Module de correspondance

Il compare l'ensemble des caractéristiques extraites par le module d'extraction des données avec le modèle enregistré dans la base de données du système. Ce module détermine donc le degré de similarité (ou de divergence) entre les deux.

d) Module de décision

Détermine l'identité d'une personne basée sur le degré de similitude entre les caractéristiques extraites et les modèles stockés.

1.7 Les modalités biométriques

Il y a plusieurs techniques de la technologie biométrique qui sont appelés aussi les modalités biométriques, qui sont classifiées en trois catégories : biologique, comportementale ou morphologique. Dans cette partie nous allons expliquer quelques exemples des différentes modalités biométriques avec leurs avantages et inconvénients :

1.7.1. Modalités morphologiques (physiologiques)

Cette catégorie est basée sur la traite physique d'un individu (unique et permanent) tels que les suivants :

a) L'empreinte digitale

L'identification à l'aide des empreintes digitales, est l'une des premières techniques biométriques qui se base sur le fait que chaque personne a des empreintes uniques. Les lecteurs d'empreintes digitales scannent puis relèvent des éléments permettant de différencier les empreintes. Ces éléments sont appelés minuties (fig1.2).

Les minuties sont des changements de continuité de l'empreinte digitale. Il existe plusieurs types de minuties : lac, bifurcation, delta ou impasse...etc. Généralement une quarantaine est extraite de la zone scannée, statistiquement il est impossible de trouver douze points identiques chez deux individus. [5]



Figure 1.2 : Image d'empreinte digitale [6].

b) Le visage

Le visage c'est la caractéristique biométriques la plus utilisée et la plus acceptable parmi les humains pour s'identifier entre eux. Il s'agit de faire une photographie plus ou moins évoluée, pour en extraire un ensemble des caractéristiques du visage (joues, yeux, nez, bouche...etc.), (fig1.3).

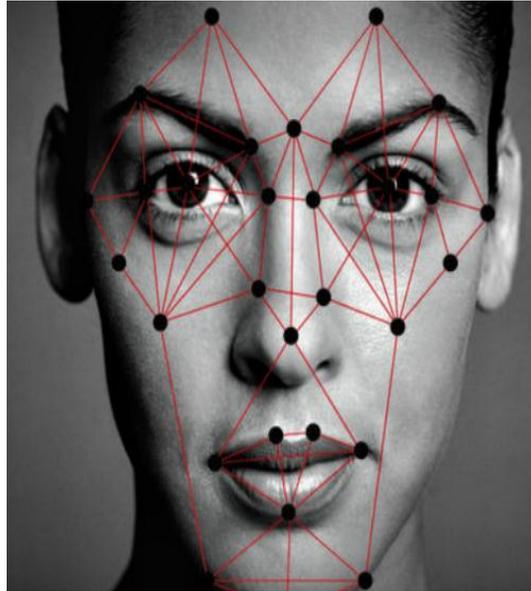


Figure1.3 : Image visage

c) L'iris

Est basé sur l'analyse du motif de l'iris sur la partie de l'œil visible. L'iris correspond à la partie colorée de l'œil (fig 1.4), et ils varient d'une personne à l'autre (même entre l'œil gauche et droit), et c'est cette partie qui est utilisée en biométrie. L'acquisition biométrique de l'iris est faite à l'aide d'une caméra.

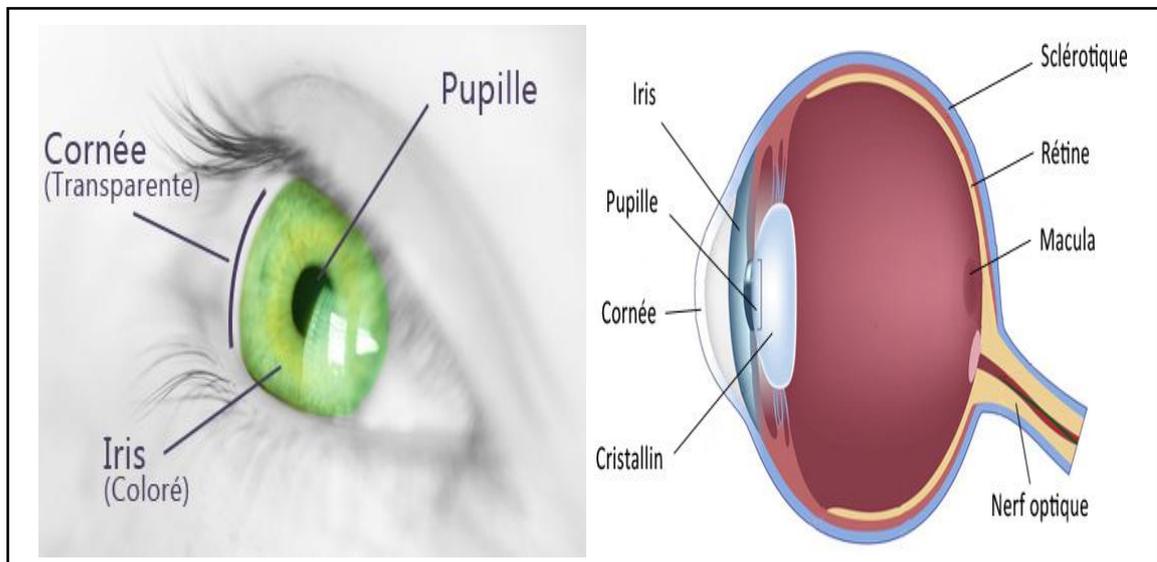


Figure1.4 : Image d'iris [7].

d) La rétine

La rétine possède des caractéristiques uniques pour chaque individu et pour chaque œil de l'être humain. Cette modalité est basée sur l'analyse de la cartographie des vaisseaux sanguins.

Elle fournit plus de 400 points caractéristiques. La mesure doit s'effectuer à très faible distance du capteur (quelques centimètres), qui effectue ensuite un balayage de la rétine.

e) Géométrie de la main

Cette modalité consiste à mesurer les caractéristiques de la main (forme longueur, largeur des doigts, l'articulation...etc.), (fig1.5).



Figure 1.5 : Image géométrie de la main.

1.7.2 Modalités comportementales

Ces modalités sont basées sur l'analyse comportementale d'un individu :

a) La signature

Dans ce cas l'utilisateur doit être signé avec un stylo électronique sur une palette graphique (fig1.6). Le système va mesurer plusieurs caractéristiques comme la vitesse, l'accélération, et la pression d'écriture...etc, pour identifier l'individu.



Figure1.6 : Image signature.

b) La voix

Est basée sur l'analyse fréquentiel de la voix sans contact physique avec la lecture du système. Les systèmes d'identification de la voix sont basés sur les caractéristiques de voix, uniques pour chaque individu (fig1.7). Ces caractéristiques de la parole sont constituées par

une combinaison des facteurs comportementaux (vitesse, rythme, etc....) et physiologiques (tonalité, âge, sexe, fréquence, accent, harmoniques, ...) [8].



Figure 1.7 : Principe de reconnaissance vocale.

c) La démarche

Ce modèle identifie les utilisateurs par l'analyse des mouvements lors de la marche (les bras, niveau des articulateurs...etc.).

d) Dynamique de la frappe au clavier

La dynamique de la frappe au clavier est une technique basée sur la façon de taper sur un clavier. Il s'agit d'un logiciel qui calcule le temps de frappe et la durée en l'air (entre les frappes) et la fréquence des erreurs, ces paramètres expriment l'état comportemental de l'individu, la fatigue, le stress...etc.

1.7.3 Modalités biologiques

Ces modalités sont basées sur l'étude des traces biologiques d'un individu :

a) L'ADN

L'ADN c'est une trace spécifique à chaque personne. Il peut être déterminé à partir d'une goutte de sang, fragment de peau ...etc. Ce modèle est basé sur l'analyse du patrimoine génétique pour identifier une personne inconnue.

b) Veines de la main

Avec la biométrie du réseau veineux, la reconnaissance de l'identité est réalisée avec la paume de la main, et plus précisément le motif et le flux veineux (fig1.8). Le motif et le flux

restent identiques de la petite enfance à la mort, et sont clairement identifiables, grâce à un scanner infrarouge et une caméra grande angle intégrée. Le système capte, en quelques millisecondes, la structure veineuse et donc l'identité univoque d'une personne enregistrée, [9].

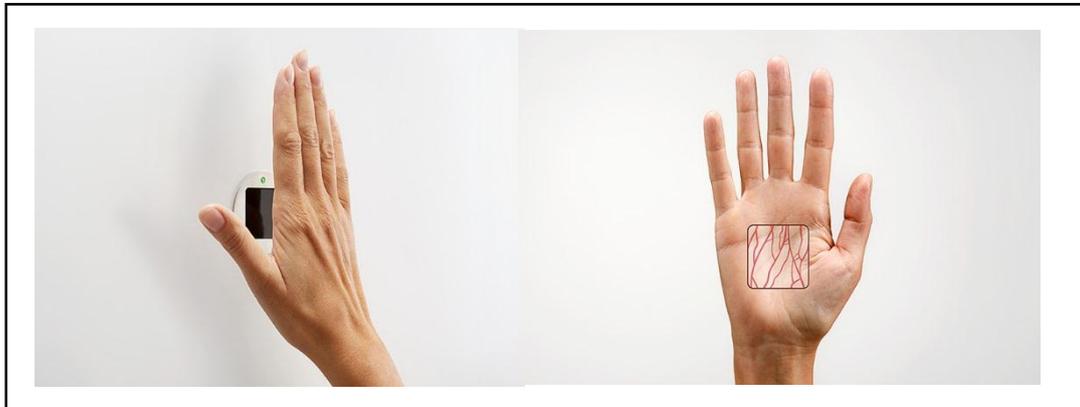


Figure1.8 : Image de veines de la main.

1.7.4 Avantages et inconvénients de différentes modalités biométriques

Les tableaux suivants résument les avantages et les inconvénients de différentes modalités [10] [5] :

Modalités comportementales		
Le modèle	Les avantages	Les inconvénients
La signature	- La signature écrite sur un document peut être conservée des certains documents.	- Besoin d'une tablette graphique. - Peu fiable
La voix	-sans contact physique. - Impossible d'imiter la voix. - Technique simple et peu Coûteuse.	- Sensible à l'état physique et émotionnel de l'individu. -Taux de faux rejet et fausse acceptation élevés. -Fraude possible par enregistrement. - Sensible aux bruits ambiants. -Taux de faux rejet et fausse acceptation élevés.
La démarche	- Transparente pour l'utilisateur	- Technique encore en phase expérimental
Dynamique de la frappe au clavier	- Pas de matériel supplémentaire, juste un programme simple. - Réduit sensiblement la nécessité de changement de mot de passe et la sollicitation des services informatique.	- Dépend de l'état (physique, émotion, fatigue...) - Problèmes avec différentes formes de clavier AZERTY/QUERTY.

Tableau1.1 : Les avantages et les inconvénients de différentes modalités comportementales.

Modalités morphologiques		
Le modèle	Les avantages	Les inconvénients
L'empreinte digitale	<ul style="list-style-type: none"> -Technique éprouvée et rapide. -Petite taille de lecteur est moins chère. -Traitement rapide. 	<ul style="list-style-type: none"> -L'utilisateur doit placer son doigt correctement sur le lecteur. -Certains systèmes peuvent accepter un doigt cassé
Le visage	<ul style="list-style-type: none"> - Usage aisé. - pas de contact physique - Technique peu coûteuse 	<ul style="list-style-type: none"> - Sensible aux changements (barbe, lunette...). - sensible à l'environnement (éclairage, position, expression du visage...).
L'iris	<ul style="list-style-type: none"> - Vrais jumeaux non confondus 	<ul style="list-style-type: none"> - L'iris est aisément visible et peut être photographié. - Le problème de sécurité est alors lié aux vérifications effectuées lors de la prise de vue. (Problème identique pour les empreintes, mais moins pour la rétine).
La rétine	<ul style="list-style-type: none"> - L'empreinte rétinienne est peu exposée aux blessures. - Les taux faussent acceptation et sont faibles. - La rétine est stable durant la vie d'un individu 	<ul style="list-style-type: none"> - L'utilisateur doit se placer son œil près du capteur. - Mauvaise acceptation du public (l'œil est un organe sensible).
Géométrie de la main	<ul style="list-style-type: none"> - très bonne ergonomie bonne acceptabilité. -Simple à utiliser. 	<ul style="list-style-type: none"> - système encombrant et coûteux. - Perturbation possible par des blessures.

Tableau1.2 : Les avantages et les inconvénients de différentes modalités morphologiques.

Modalités biologiques		
Le modèle	Les avantages	Les inconvénients
L'ADN	<ul style="list-style-type: none"> - L'ADN est facile à obtenir (cheveux, salive...) 	<ul style="list-style-type: none"> - Coûteux - Obtenir le résultat nécessite beaucoup de temps.
Veines de la main	<ul style="list-style-type: none"> - Les lecteurs de la biométrie des réseaux veineux n'enregistrent pas de photographie respect de l'intimité. - fonctionnent même avec les mains tachées 	<ul style="list-style-type: none"> - Technique encore en phase expérimental.

Tableau1.3 : Les avantages et les inconvénients de différentes modalités biologiques.

1.8 Reconnaissance Des Expressions Faciales

1.8.1 Historique sur la reconnaissance de l'expression faciale

L'expression faciale est un aspect important du comportement et de la communication non verbale, où le changement dans le visage, est perceptible visuellement, dû à l'activation (volontaire ou non) de l'un ou de plusieurs des 44 muscles composant le visage (250000 expressions possibles). L'expression faciale, déjà étudiée par **Darwin** et **Duchenne** de Boulogne au dix-huitième siècle, a joué un rôle majeur dans la recherche sur les émotions depuis les travaux de **Sylvan Tomkins** dans les années 1960. Ses élèves **Paul Ekman** et **Caroll Izard** ont défendu l'idée d'un nombre limité d'émotions de base auxquelles sont associées des expressions faciales automatiques, universelles et innées [11] [12].

Durant la seconde moitié du dix-huitième siècle, le neurologue **Duchenne** de Boulogne réalise une série d'expériences sur l'expression faciale de l'émotion. Il utilise la photographie et la stimulation électrique des muscles de la face pour mettre en évidence les mouvements associés à l'expression des émotions. Il remarque notamment que les sourires exprimant une joie sincère se différencient des sourires volontaires par la contraction d'*orbicularis oculi*, un muscle situé autour des yeux.

Des recherches menées dans les années 1980 par **Paul Ekman** et son équipe ont permis de confirmer et de compléter ces résultats. **Ekman** a mis en évidence le fait que nous sommes pour la plupart incapables de contracter volontairement *orbicularis oculi*, et que ceux qui le peuvent n'arrivent généralement pas à contracter ce muscle de chaque côté au même moment.

En outre, les sourires de **Duchenne** sont généralement associés à une activité asymétrique dans le lobe frontal, considérée comme un signe d'affection positif [13].

1.8.2 Système de reconnaissance des expressions faciales

Les expressions faciales sont l'un des moyens de communication non verbaux résultant d'une réaction souvent involontaire, résultant principalement de contractions musculaires modifiant les traits du visage tels que la bouche, les sourcils. C'est un moyen essentiel de transmission d'informations sociales entre humains, d'une durée de quelques secondes.

Un système de reconnaissance des expressions faciales est constitué de trois étapes principales. La première étape c'est la détection du visage où déterminer la zone du visage. La deuxième étape c'est l'extraction de donnée à partir du visage d'individu pour extraire les

caractéristiques du visage tels que la position de la bouche, les yeux, et les sourcils et donner une meilleure description de l'expression. La dernière étape est la classification qui permet de classifiez spécifiez l'expression appropriée pour l'image à partir des caractéristique extraites dans l'étape précédente.

Le schéma suivant illustre un système complet de reconnaissance des expressions faciales :

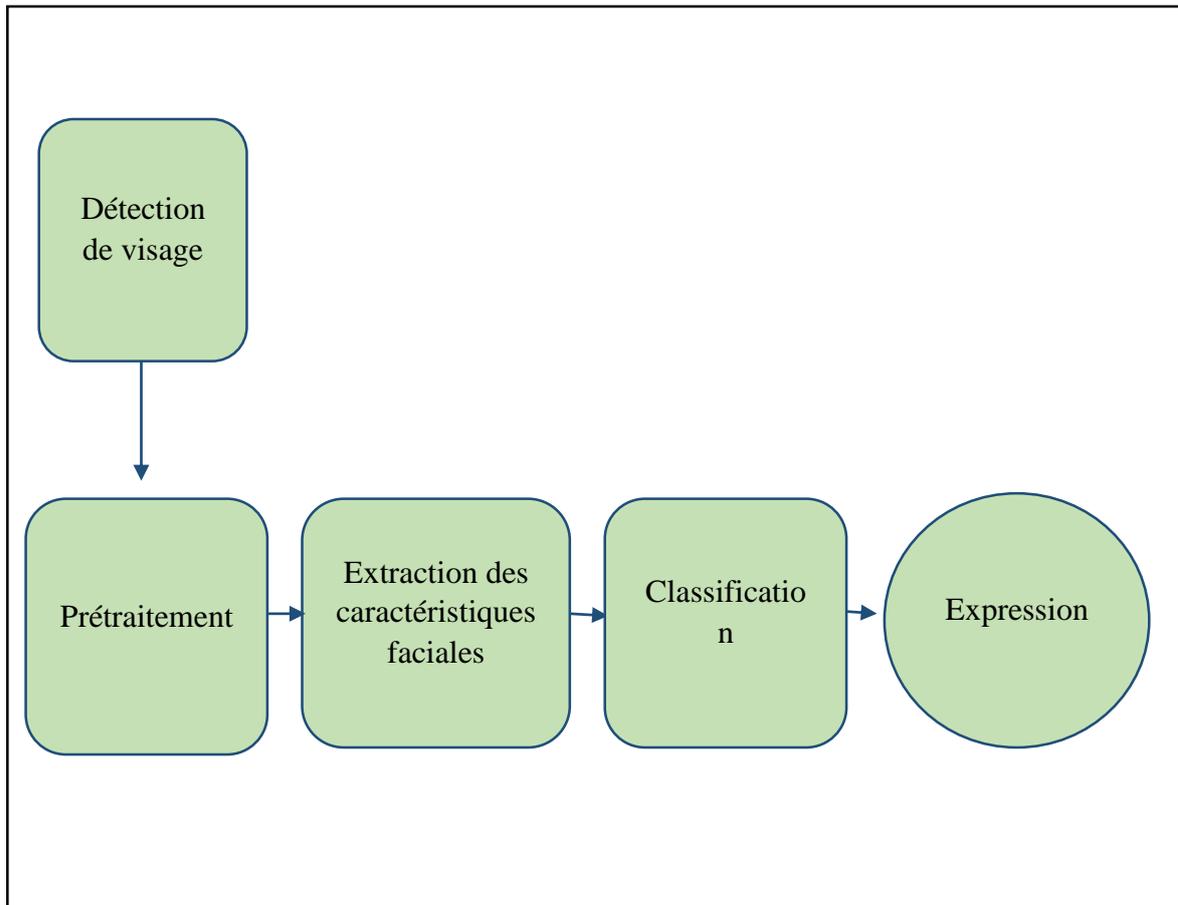


Figure1.9 : Architecture d'un système de reconnaissance des expressions faciales.

1.8.3 Les 6 expressions faciales universelles

Au commencement c'était **Charles Darwin**, qui affirmait dans '*L'expression des émotions chez l'homme et l'animal (1872)*', que celles-ci étaient innées et universelles. Il apportait également l'idée que les expressions faciales étaient intimement liées à ces émotions et pouvaient servir de moyen de mesure [14].

En 1972 **Paul Ekman** confirmé après ses expériences sur la reconnaissance d'émotions à partir d'expressions faciales dans différentes cultures tels qu'asiatique et latines. L'expression des émotions est universellement reconnue dans le monde.

Donc voilà les 6 émotions universelles, (fig1.10) :

-Colère

-Dégoût

-Joie

-Surprise

-Peur

-Tristesse

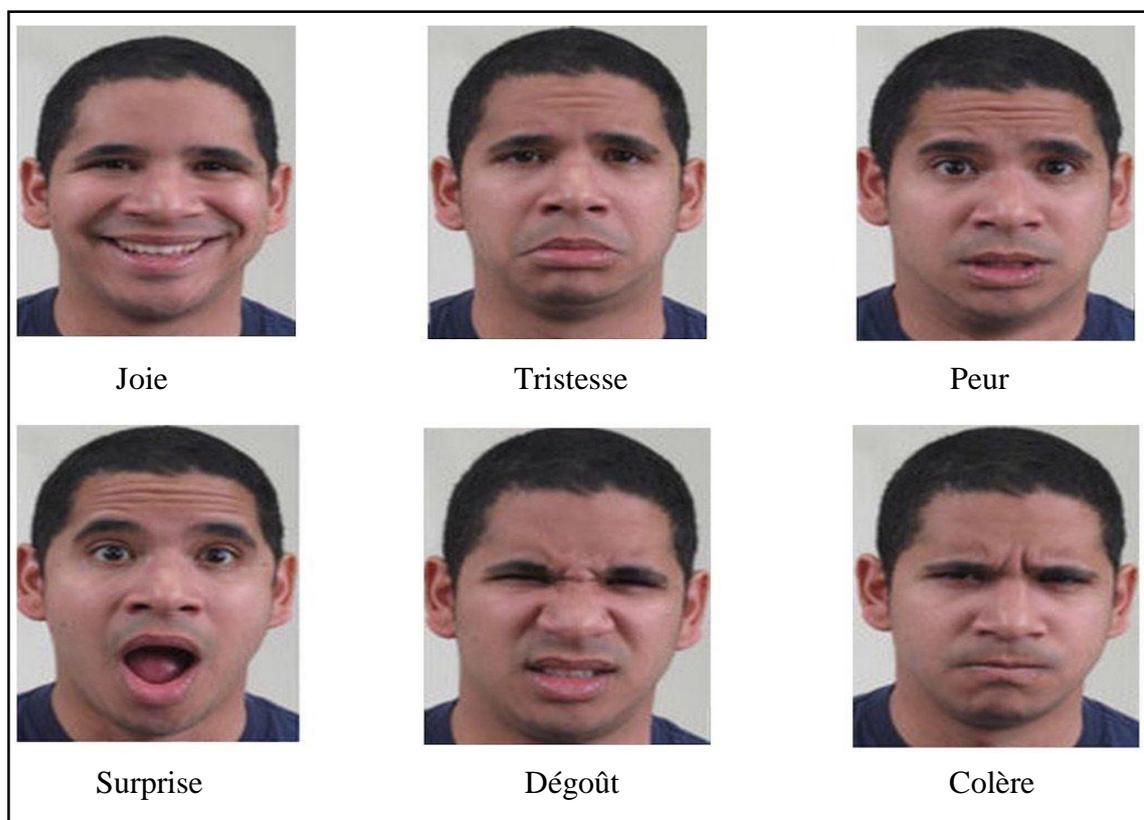


Figure 1.10 : Les six émotions universelles [14].

Biehl et Matsumoto réhabiliteront le mépris comme 7^{ème} émotion en 1997 [15]. Par ailleurs le professeur et psychologue américain **Robert Plutchik** considère qu'il y a huit émotions essentielles : la joie, la peur, le dégoût, la colère, la tristesse, la surprise, la confiance et l'anticipation. **Robert Plutchik** a aussi créé une roue des émotions pour illustrer diverses émotions motivatrices et nuancées. Il a proposé son modèle complexe en trois dimensions et son modèle de roue en deux dimensions en 1980 pour décrire les relations entre ces émotions [16] (fig1.11).

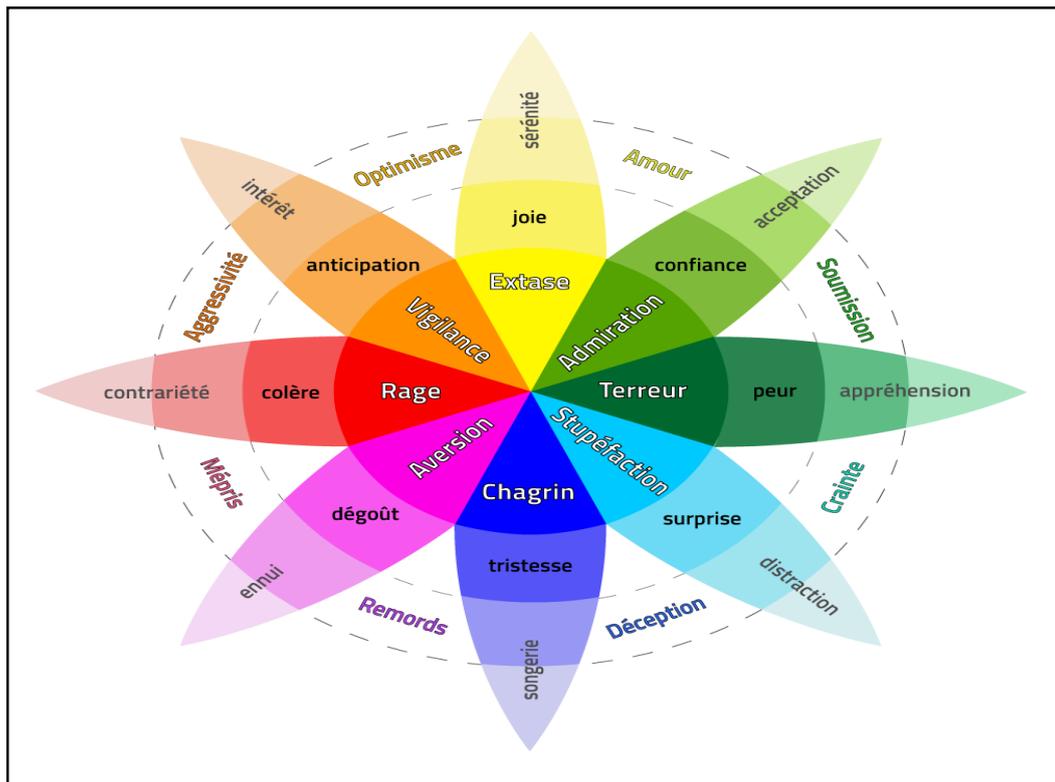


Figure1.11 : Roue des émotions de **Plutchik** [16].

Récemment et en 2014 l'**Université de Glasgow** a publié un article à propos d'une étude de **Rachael Jack** où il a trouvé que : colère et dégoût fusionnent, de même que peur et surprise, du fait de leur similarité dans les premiers instants [14]. Donc quatre émotions restent : joie, tristesse, colère/dégoût, peur/surprise.

1.8.4 Différentes méthodes de reconnaissance des expressions faciales

Les techniques de reconnaissance faciale intéressent de nombreux chercheurs, donc, il est apparu beaucoup d'algorithmes et des méthodes visant à obtenir le meilleur résultat possible en pour la reconnaissance des expressions du visage.

On peut diviser ces méthodes en deux approches :

- Les approches d'apparence.
- Les approches géométriques.

Chaque approche contient des différents modèles pour obtenir la reconnaissance des expressions faciales, nous allons mettre en évidence certaines de ces méthodes.

1.8.4.1 Approches d'apparence

a) Motif binaire local (LBP)

La méthode LBP (Local Binary Pattern) est l'un des techniques qui a récemment attiré l'attention dans la représentation des expressions faciales, a été proposée par **Ojala et al**, en 1996.

LBP basé sur la séparation de l'image en plusieurs petites régions pour extraire la caractéristique de chaque point ou bien pixel.

L'opérateur LBP travaille parmi huit voisins de pixel, en utilisant la valeur de pixel centrale comme seuil. Lorsque le pixel voisin a une valeur de gris supérieure au pixel central par rapport à la valeur attribuée à ce pixel, en obtient "1". Le même cas se produit lorsque les valeurs sont identiques. Ensuite, le code LBP est créé par concaténation de zéros et des uns en un code binaire comme illustré à la figure 1.12 ou en huit valeurs pour le pixel central [17].

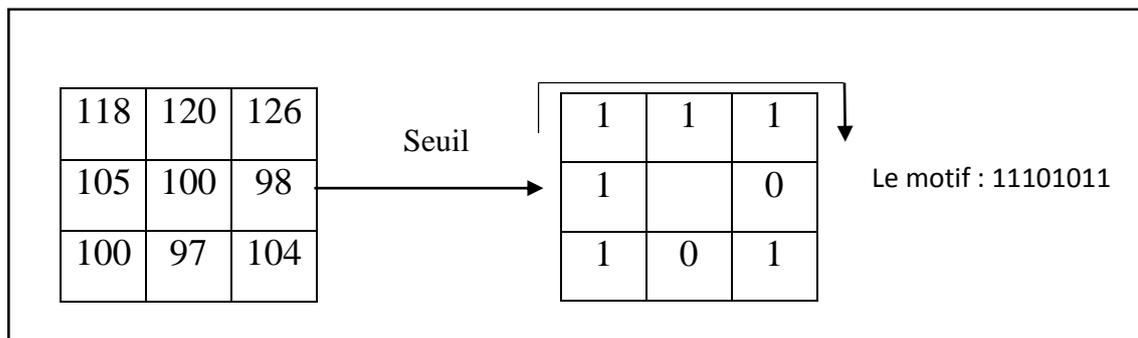


Figure 1.12 : L'opérateur LBP.

La majorité des recherches qui était basée sur la méthode de LBP pour améliorer la reconnaissance des expressions faciales, donnent une précision entre 91.7% et 95.8%. Une expérience basée sur la combinaison de (LBP+SVM) donne une efficacité de 90% [17], une autre expérience utilisant un modèle LBP de pièces du visage et réseau de neurones multicouches donne une efficacité de 95.8% [18].

b) Quantification de phase locale (LPQ)

La technique de LPQ (Local Phase Quantization), a été proposée par **Ojansivu**, [19]. Pour l'analyse des expressions faciales et reconnaissance des visages. Cette méthode est basé sur le calcul de la transformée de Fourier dans les voisinages locaux (Transformée de Fourier à Court Terme : TFCT).

À chaque pixel, les coefficients de Fourier locaux sont calculés pour quatre points de fréquence. Ensuite, les signes de la partie réelle et imaginaire de chaque coefficient sont quantifiés à l'aide d'un quantificateur scalaire binaire, pour calculer les informations de phase. Les coefficients binaires à huit bits résultants sont ensuite représentés sous forme d'entiers à l'aide d'un codage binaire. Cette étape est similaire à l'étape de construction de l'histogramme dans LBP. À la fin, nous obtenons un vecteur de caractéristiques à 256 dimensions [20]. Une expérience pour la reconnaissance des expressions faciales qui utilise la méthode de LPQ donne un résultat de 79.11% d'efficacité [21].

c) Histogramme de gradient orienté (HOG)

L'un des méthodes de reconnaissance des expressions faciales c'est HOG (Histograms of oriented gradients) qui a été proposée par **Dalal et Triggs** 2005, [22]. Les descripteurs HOG décrivent l'apparence locale et la forme du visage. L'implémentation de ces descripteurs peut être obtenue en divisant l'image en petites régions connectées (cellules) pour lesquelles un histogramme des directions de gradient (orientations des bords) est calculé. La combinaison de ces histogrammes représente alors le vecteur de caractéristiques.

L'algorithme implique différentes étapes, la première étape est le calcul des valeurs du gradient. La deuxième étape consiste à créer les histogrammes de cellules. Chaque pixel de la cellule génère un vote pondéré pour un canal d'histogramme basé sur l'orientation en fonction des valeurs trouvées dans le calcul du gradient (fig1.13). Les cellules sont regroupées en blocs plus grands, liés dans l'espace. Le descripteur HOG est alors le vecteur qui concatène les composants des histogrammes normalisés de tous les blocs les régions. Ici, plusieurs paramètres doivent être pris en compte : la taille de la cellule en pixels, le nombre de cellules par bloc et le nombre d'orientations [23]. Le taux de précision de la méthode HOG et son efficacité est environ 95.8%, à partir d'une expérience basé sur la méthode de HOG, [22].

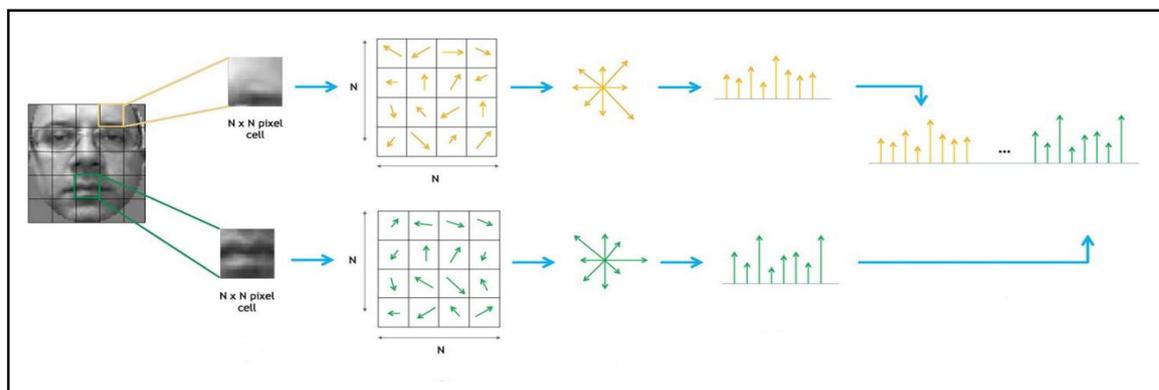


Figure1.13 : Exemple d'histogramme obtenu dans méthode de HOG [22].

d) Ondelette de Gabor

Gabor Wavelet en anglais ou ondelette du Gabor à son inventeur **Dennis Gabor** en 1946, [24].

Les ondelettes de Gabor sont un outil important pour l'analyse et le traitement des signaux dans les champs spatiaux et fréquentiels, comparaison à la transformation de Fourier traditionnelle [25]. La propriété importante de l'ondelette est qu'elle minimise le produit de ses écarts-types dans le domaine temporel et fréquentiel. En d'autres termes, l'incertitude dans les informations véhiculées par cette ondelette est minimisée. Cependant, ils ont l'inconvénient d'être non orthogonaux, ce qui rend difficile une décomposition efficace dans la base. Depuis leur création, diverses applications sont apparues, du traitement d'images à l'analyse de neurones dans le système visuel humain [26]. Une étude faite pour la reconnaissance d'expression du visage basée sur la transformation en ondelettes de Gabor et la correspondance de modèles élastiques donne une efficacité de 90.4% [27].

1.8.4.2 Approches géométriques

Les approches géométriques, ou basées sur des modèles, comparent une image d'entrée à un ensemble de modèles construits soit à l'aide d'outils statistiques, soit en analysant les caractéristiques faciales locales et leurs relations géométriques.

a) Modèle de forme active (ASM)

Les modèles de formes actives (Active Shape Model) a été développés par **Tim Cootes** et **Chris Taylor** en 1995, [28]. Le modèle de forme active (ASM) représente un modèle paramétrique déformable dans lequel un modèle statistique de la variation de forme globale par rapport à l'ensemble d'apprentissage (c'est-à-dire la base de données).

Ce modèle de forme est contraint par le MPD (modèle de distribution de points) à ne varier que de la même manière que dans l'ensemble de formation des exemples étiquetés. La forme d'un objet est représentée par un ensemble de points (fig1.14) [29]. Comme elle a donné les expériences basées sur cette méthode une précision est de 92.1% [30].

Le principe de fonctionnement d'ASM est basé sur l'alternance entre les deux étapes suivantes :

- Générez un modèle proposé en regardant autour de chaque point pour obtenir la meilleure position pour ce point.
- Compatibilité entre le modèle proposé et le modèle de distribution par points.

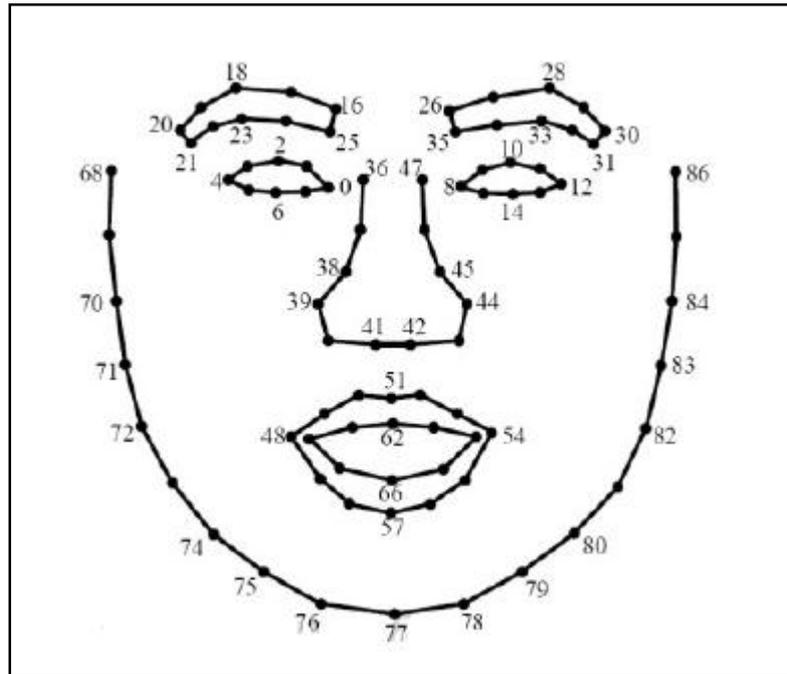


Figure1.14 : Ensemble des points caractéristiques de l'ASM model [31].

b) Modèles d'apparence active (AAM)

Un modèle d'apparence active (Active Appearance Models) est une amélioration du modèle précédent (ASM). (AAM) a été introduit par **Edwards, Cootes and Taylor** 1998, [32].

AAM est une méthode de segmentation à base statistique, où la variabilité de la forme et de la texture est capturée à partir du jeu de données [33].

L'algorithme utilise la différence entre l'estimation d'apparence actuelle et l'image cible pour piloter un processus d'optimisation. En tirant parti des techniques des moindres carrés, il peut être associé très rapidement aux nouvelles images [34]. La précision de cette méthode est environ 91.5% [35].

1.9 Conclusion

Dans ce chapitre, nous avons présenté les différentes modalités biométriques et nous avons spécifiée notre étude sur l'expression du visage. Après l'analyse théorique de différentes approches de REF et leur efficacité, nous avons choisis deux approches LBP et HOG pour les appliqués dans notre travail dans ce mémoire.

Chapitre 2

Méthodes de reconnaissance des expressions faciales

2.1 Introduction	23
2.2 Système de reconnaissance des expressions faciales	24
2.3 Prétraitement.....	24
2.4 Extraction des caractéristiques	25
2.4.1 Motif binaire local (LBP : Local Binary Pattern).....	25
2.4.1.1 Concept de LBP	25
2.4.1.2 différentes extensions de LBP	27
2.4.1 Histogramme de gradient orienté (HOG : Histogram of oriented gradients)	28
2.4.1.1 Concept de HOG	28
2.4.1.2 Différentes extensions de HOG	31
2.5 Classification	31
2.5.1 Machine à vecteurs de support (Support vector machine) SVM.....	32
2.5.2 Distance euclidienne.....	33
2.6 Conclusion.....	34

2.1 Introduction

L'expression faciale est une réaction émotionnelle visible de l'extérieur reflète l'état psychologique d'une personne. Cela se manifeste par des changements dans les traits du visage. Les déformations des traits faciaux représentent l'information sur l'expression du visage, que a regroupe précédemment dans six émotions universelles colère, dégoût, joie, peur, surprise, tristesse (voir chapitre 1).

Les recherches sur les expressions faciales a débuté depuis le 20ème siècle depuis les travaux de **Silvan Tomkins** 1960, et ses élèves **Paul Ekman** et **Carroll Izard** qui définissait les six émotions universelles. Depuis lors, plusieurs systèmes ont été proposez pour définir l'émotion d'humaine. Avec l'évolution de vision artificielle et la communication homme-machine, nous trouvons qu'il existe plusieurs techniques qui tentent de définir l'émotion d'une personne avec précision et haute performance.

Dans ce chapitre nous allons représenter précisément les différentes étapes d'un système de reconnaissance des expressions faciales, ce qui est résumé en trois étapes principales : prétraitement, extractions des caractéristiques et classification.

2.2 Système de reconnaissance des expressions faciales

Généralement, un système de reconnaissance des expressions faciales est constitué de trois phases principales : la première étape c'est prétraitement où détecter et localiser le visage et éliminer tous les parasites tels que contraste et illumination. La prochaine étape consiste à extraire les caractéristiques du visage qui définissent l'expression (position des yeux, les sourcils, la bouche). Finalement et à partir de ces informations extraire l'expression de l'individu sera classifié à une catégorie d'expressions (joie, tristesse, peur, colère, dégoût) (fig2.1).

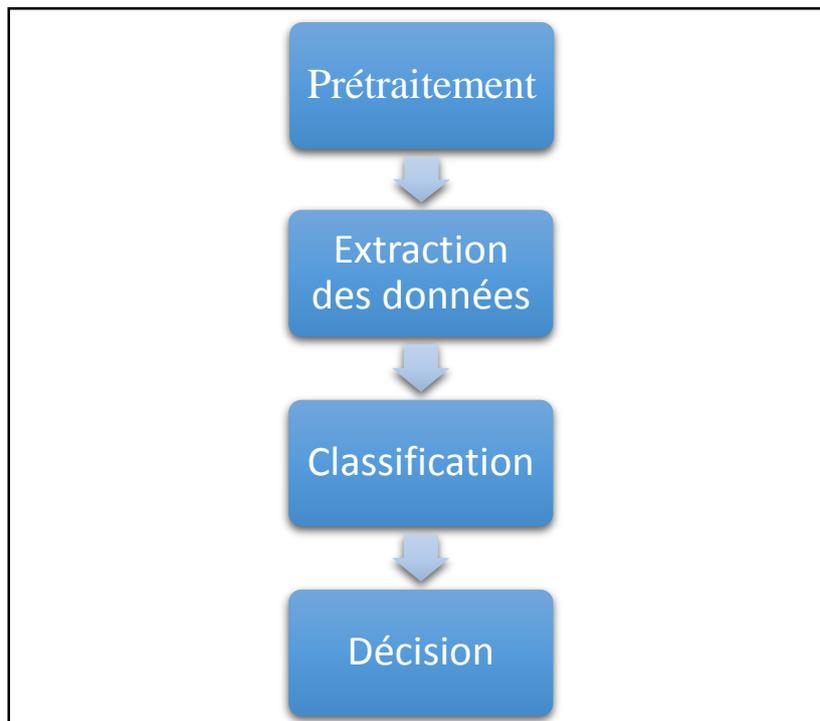


Figure 2.1 : Architecture d'un système de reconnaissance des expressions faciales.

2.3 Prétraitement

Tous les systèmes de reconnaissance faciale comportent une étape de prétraitement. Initialement, on a délimité la zone d'intérêt par un rectangle pour localiser la zone du visage dans l'image. Après la détection du visage nous obtenons une image n'est pas pure à cause des parasites comme la variation lumineuse et l'arrière-plan. Pour cela, il est nécessaire d'éliminer le bruit et les zones non informatives dans l'image, afin d'obtenir le bon fonctionnement du système de reconnaissance des expressions faciales.

Toutes les images doivent être de taille et de couleur uniformes (RGB, niveaux de gris...), puis appliquer une normalisation sur le visage pour réduire les effets de l'illumination.

2.4 Extraction des caractéristiques

Après avoir localisé un visage dans une image, la prochaine étape est l'extraction des caractéristiques qui est l'étape la plus importante du système de reconnaissance d'expression faciale. Extraire des informations à partir du visage détecté pourraient être obtenues en décrivant des points d'intérêt associés aux caractéristiques faciales, tels que les yeux, les sourcils, la bouche, afin, que on obtient un vecteur représente tous les données.

Il existe plusieurs techniques d'extraction des données du visage, comme mentionner dans le chapitre précédent, nous avons adopté sur deux méthodes : LBP et HOG, comme le montre ce qui suit :

2.4.1 Motif binaire local (LBP: Local Binary Pattern)

2.4.1.1 Concept de LBP

Comme nous l'avons dit dans la définition de LBP dans le chapitre précédent, le concept de LBP était basé sur le calcul du voisinage de chaque pixel dans un bloc de (3*3). Nous prenons la valeur centrale comme seuil, pour cela donc nous obtenons "1" lorsque la valeur du pixel voisin est supérieure ou égale au pixel central et "0" est obtenue lorsque la valeur du pixel voisin est inférieure au pixel central. Après l'obtention d'un code binaire nous le convertissons en code décimal comme le montre la figure 2.2 ci-dessous.

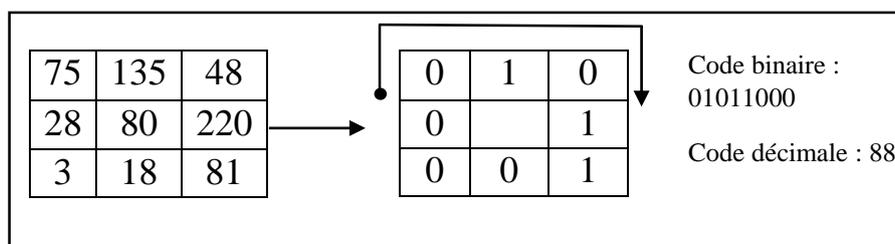


Figure 2.2 : L'opérateur LBP.

L'un des inconvénients de l'opérateur fondamental de LBP est qu'il est incapable de saisir certaines caractéristiques dominantes. Afin de pouvoir traiter la texture à des échelles complètement différentes, il a donc fallu travailler sur l'opérateur et ainsi émerger l'utilisation de voisinage de différentes tailles. Un voisinage local est défini comme un groupe de points d'échantillonnage de la même dimension répartis sur un cercle et centrés à l'intérieur des pixels. Les interpolations sont interpolées à l'aide d'une interpolation bilinéaire, permettant ainsi de définir tout rayon et toute variété de points d'échantillonnage dans le voisinage, [36].

Pour les échantillons de l'opérateur LBP étendu, où la notation (P, R) désigne un voisinage de P à l'intérieur duquel les points sont équidistants (la même dimension) sur un cercle de rayon R créant un ensemble voisin de symétrie circulaire (fig2.3) [36].

Nous disons que le modèle binaire local est uniforme lorsque nous avons au plus deux transitions binaires de 0 à 1 ou vice versa par exemple 00000000 (0 transitions), et 01110000 (2 transitions) sont uniformes, le modèle binaire 11001001 (4 transitions), et 01010011 (5 transitions) n'est pas uniforme.

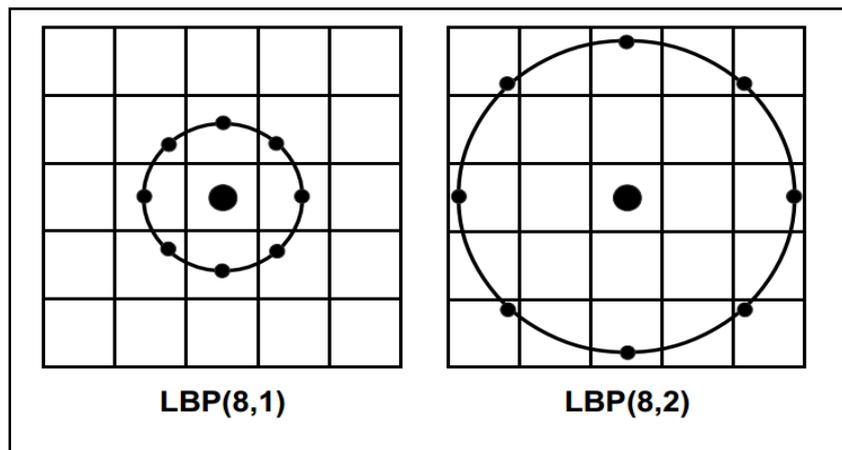


Figure 2.3 : Un ensemble voisinage d'opérateur LBP [37].

L'opérateur LBP produit 2^p ($p = 8$) valeurs de sortie, nous avons donc un vecteur de 256 valeurs. Dans [38] Ojala et al, ont observé que bien que seulement 58 des 256 motifs à 8 bits soient uniformes, près de 90% de tous les voisinages d'image observés sont uniformes et bon nombre des motifs restants contiennent essentiellement du bruit. Ainsi, l'histogramme de chaque motif uniforme est réduit de 256 à 59 valeurs. Le schéma ci-dessous montre les étapes pour obtenir l'histogramme de LBP final de l'image :

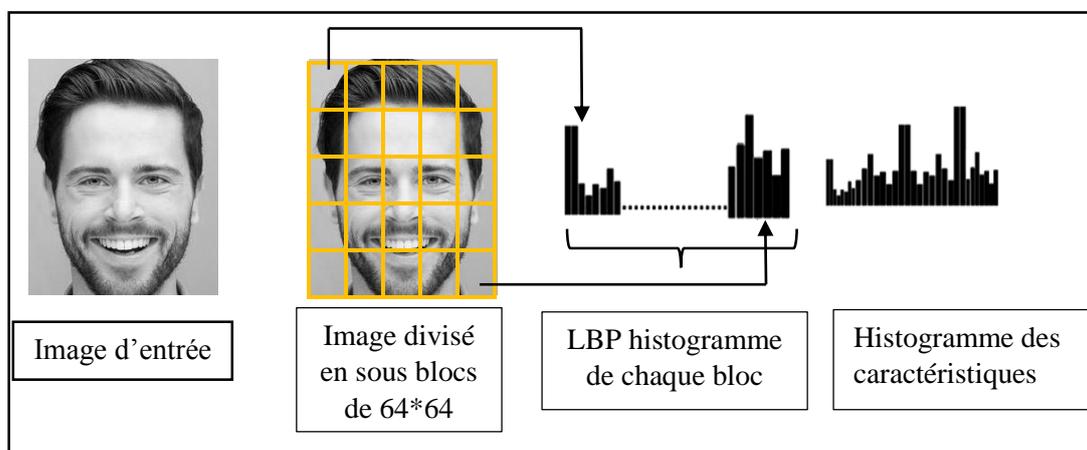


Figure2.4 : Schéma d'obtention d'histogramme LBP.

2.4.1.2 différentes extensions de LBP

Récemment l'opérateur LBP a été développé avec de nombreuses variantes pour améliorer les performances dans différentes applications. Ces variations se concentrent sur différents aspects de l'opérateur LBP d'origine [38] :

- amélioration de sa capacité de discrimination.
- amélioration de sa robustesse.
- sélection de son voisinage.
- extension aux données 3D.
- combinaison avec d'autres approches.

Dans cette section, nous examinons quelques extensions récentes de l'opérateur LBP :

a) Modèles binaires locaux améliorés (The Improved Local Binary Patterns ILBP)

Jin et al ont modifié l'opérateur LBP pour décrire des informations de structure plus locales et pour améliorer la capacité de discrimination [38]. La méthode ILBP compare tous les pixels (même le pixel central) à l'intensité moyenne de tous les pixels du bloc. Dans un bloc de (3*3) de voisinage, l'opérateur LBP (8, 1) produit 256 (2^8) valeurs. En revanche ILBP produit 511 motifs (2^9-1 , car tous les zéros et toutes les uns sont identiques) [38] (fig2.5).

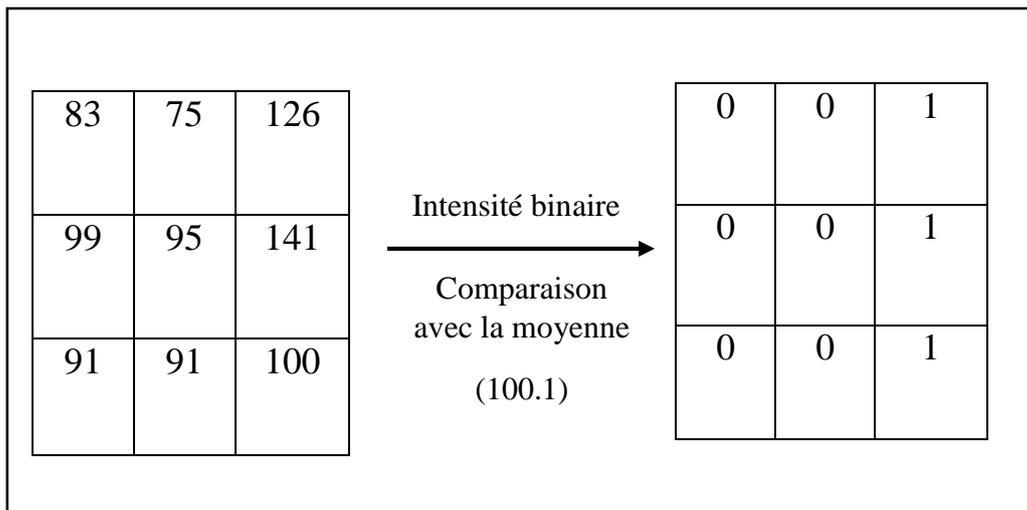


Figure2.5 : Un exemple d'opérateur ILBP.

b) Modèle binaire local complet (Complete Local Binary Patterns CLBP)

Plus récemment, **Guo et al**, a proposé une LBP complète (CLBP), [43].CLBP inclut les différences de signe et de valeur de gris entre un pixel central donné et ses voisins afin d'améliorer le pouvoir discriminant de l'opérateur LBP d'origine.

CLBP compare à nouveau la valeur absolue de GD (différences de valeur grise, gray differences en anglais) avec le pixel central donné pour générer un code similaire à LBP.

c) Modèle binaire local en volume (Volume Local Binary Pattern VLBP)

L'opérateur de modèle binaire local en volume VLBP (Local Binary Pattern), est une combinaison de mouvement et l'apparence pour décrire une texture dynamique. Il été produite à partir d'opérateur LBP original pour réaliser une représentation spatio-temporelle pour l'analyse de texture dynamique car l'opérateur original de LBP ne traite que les informations spatiales. L'idée de base de VLBP consiste à examiner la texture dynamique en tant qu'ensemble de volumes dans l'espace (X, Y, T), où X et Y désignent les coordonnées spatiales et T, l'indice de temps. Donc le voisinage de chaque pixel est ainsi défini dans un espace tridimensionnel [39].les informations de VLBP peut être définis et extraits dans des histogrammes

Il y a d'autres extensions de LBP tels que : ELBP, LBP-TOP, Soft LBP...etc.

2.4.1 Histogramme de gradient orienté (HOG : Histogram of oriented gradients)

2.4.1.1 Concept de HOG

HOG est un descripteur de forme et d'apparence qui compte les occurrences des orientations du gradient dans des parties spécifique d'une image. Comme nous l'avons dit précédemment dans le chapitre 01, initialement, l'image est divisée en régions adjacentes de petite taille ($N \times N$ pixels), appelées cellules, et en calculant l'histogramme des directions du gradient ou des orientations des contours pour les pixels de chaque cellule. Ces histogrammes sont triés et combinés dans un seul histogramme. Soit $Thog_{fs}$ représente le nombre total des caractéristiques calculées à l'aide du descripteur HOG, $Bimg$ indiqué des blocs par image, la taille de bloc est représentée par Bs et Nb indique le nombre de cases utilisés, ensuite, le nombre total des caractéristiques peut être calculé à l'aide de la formule suivante [40] :

$$Thog_{fs} = Bimg * Bs * Nb \quad (2.1)$$

Pour de meilleurs résultats, les histogrammes locaux sont normalisés en contraste, en calculant une mesure de l'intensité sur des zones plus larges que les cellules, appelées des blocs, et en utilisant cette valeur pour normaliser toutes les cellules du bloc. Cette normalisation permet une meilleure résistance aux changements d'illuminations et aux ombres [41].

C'était brièvement le concept général de descripteur de HOG, pour l'obtention d'histogramme finale il faut passer par cinq étapes nécessaires comme suite :

a) Calcul du gradient

Le calcul rapide du gradient peut être fait, par exemple, par des masques de dérivation simple 1D centré. On applique un masque $[-1 \ 0 \ 1]$ puis un masque $[-1 \ 0 \ 1]^T$ ou non centré $[-1,1]$. Il y a d'autre type de masquage tels que filtre de Sobel (3*3), ou Derich ou, filtrage gaussien pour dimension 2D.

b) Construction de l'histogramme

Dans cette étape l'image est divisée en 8*8 cellules et on calcule l'histogramme d'orientation de gradient pour chaque cellule, qui vote alors pour une classe de l'histogramme, en fonction de l'orientation du gradient à ce point. Le vote du pixel est pondéré par l'intensité du gradient en ce point. Les histogrammes sont uniformes de 0 à 180° (cas non signé) ou de 0 à 360° (cas signé) [41].

c) Formation et normalisation des blocs

Dans l'étape précédente, nous avons créé un histogramme basé sur le calcul de l'orientation de gradient de l'image. Les valeurs des HOG peuvent varier de manière significative à cause de la sensibilité à l'éclairage (variation de l'éclairage ou contraste entre différentes parties d'une image). Cela change les valeurs de l'histogramme, et pour cela nous avons fait une normalisation sur l'histogramme.

La normalisation se fait en formant des blocs, chacun est constitué de quatre cellules de 6x6 pixels, ou effectue la normalisation sur ces blocs (fig2.6).

Il existe une corrélation entre le nombre de cellules de bloc et la précision de descripteur, si le nombre de cellules dans un bloc est élevé, le descripteur est plus discriminant. Mais la taille des cellules est alors plus petite. La taille des cellules doit cependant rester assez importante car, si les cellules sont trop petites, les informations qu'elles contiennent sont

trop peu nombreuses et donc peu pertinentes, de plus, un grand nombre de blocs donne un résultat final plus précis [42].

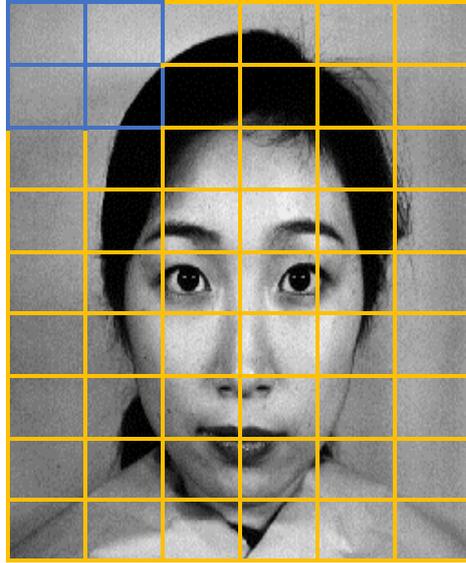


Figure 2.6 : Normalisation des blocs (rectangle bleu représente un bloc).

Soit V_{bloc} est le vecteur de caractéristique et N la taille du vecteur, l'histogramme est normalisé selon quatre types de normalisation [42] :

- Norme L1 : moyenne du vecteur :

$$V(i)_{bloc} = \frac{V(i)_{bloc}}{\sum_{k=1}^N V(k)_{bloc}} \quad (2.2)$$

- Norme L1 racine : la racine carrée de la norme L1 est prise afin de traiter le vecteur de caractéristiques comme une densité de probabilité :

$$V(i)_{bloc} = \sqrt{\frac{V(i)_{bloc}}{\sum_{k=1}^N V(k)_{bloc}}} \quad (2.3)$$

- Norme L2 : la moyenne par la norme 2 du vecteur :

$$V(i)_{bloc} = \frac{V(i)_{bloc}}{\sqrt{\sum_{k=1}^N V(k)_{bloc}^2}} \quad (2.4)$$

- Norme 2 à hystérésis : L'influence des gradients de hautes magnitudes est ici réduite en seuillant à 0,2 après une normalisation L2. Une normalisation unitaire est finalement réalisée pour obtenir la norme L2 à hystérésis [42].

2.4.1.2 Différentes extensions de HOG

a) Histogramme pyramidal de gradient orienté (Pyramid Histogram of Orientated Gradient PHOG)

Le descripteur HOG a été développé aux d'autres extensions telles que PHOG (**Bosch, Zisserman, & Munoz, 2007**) qui est utilisé généralement pour l'analyse d'expression faciale statique. Au début L'opérateur **Canny Edge** est appliqué pour extraire les contours des stimuli donnés [43]. Ensuite l'image du visage est divisée en grilles spatiales puis calculer l'histogramme de l'orientation (HOG) en utilisant le masque de Sobel (3*3) aux contours.

Pour obtenir le descripteur PHOG final, les histogrammes de gradients (HOG) aux mêmes niveaux sont concaténés. Le descripteur PHOG final est une concaténation de HOG à différents niveaux de la pyramide [44].

Aussi la méthode de HOG a été combinée avec autres descripteur comme Gabor ou LDB, [45].

2.5 Classification

La classification représente la dernière étape de système de reconnaissance des expressions faciales. Après l'extraction des caractéristiques du visage et son expression, où, à la fin de cette étape, nous pouvons connaître l'expression du visage approprié.

C'est l'étape qui fait la différence entre un système d'identification d'individus et un autre de vérification [46]. La classification consiste à une mesure de similarité afin d'estimer la différence entre le visage d'entrée et les visages qui sont stockés dans la base de données afin de déterminer l'émotion de l'individu parmi les six émotions universelle soit joie ,tristesse, dégoût, surprise, peur, ou colère.

Plusieurs méthodes sont utilisées pour classifier l'expression faciales d'un individu, on trouve des technique basé sur le calcul de distance tels que distance euclidienne qui calculer la racine de l'ensemble de soustraction de deux vecteurs de caractéristique le tous carré.

Il y a d'autres méthodes se basent sur la classification des caractéristiques par un seul classificateur comme **SVM, KNN, K-means** ,toutes ces méthodes consister à déterminer le modèle le plus proche de visage d'entre parmi les images de la base de donné.

2.5.1 Machine à vecteurs de support (Support vector machine SVM)

Le SVM est l'un des techniques d'apprentissage supervisé qui considérée comme des classificateurs linéaires, dans le but de résoudre les problèmes de discrimination et de régression.

La marge maximale c'est la distance entre la frontière de séparation et les échantillons les plus proches. Par exemple nous avons deux classes différentes : les cercles rouges et les étoiles jaunes comme citer dans figure2.7. L'objectif est de séparer les données et maximiser la distance entre ces deux classes et c'est exactement ce que fait le SVM et c'est qu'on l'appelle hyperplan séparateur ce qui représente la ligne droite qui séparant les deux catégories.

Le problème ici est qu'il existe plusieurs d'hyperplan séparateur plus valide mais sa performance en généralisation est différente. Ainsi le but d'un SVM est de trouver cette frontière optimale en maximisant la distance entre les points des échantillons et la frontière (distance entre cercle ou étoile et la frontière). les points des échantillons les plus proches de la frontière sont nommées vecteurs support. Plus que la marge est large, plus il sera efficace pour classer un nouveau exemple (fig2.8).

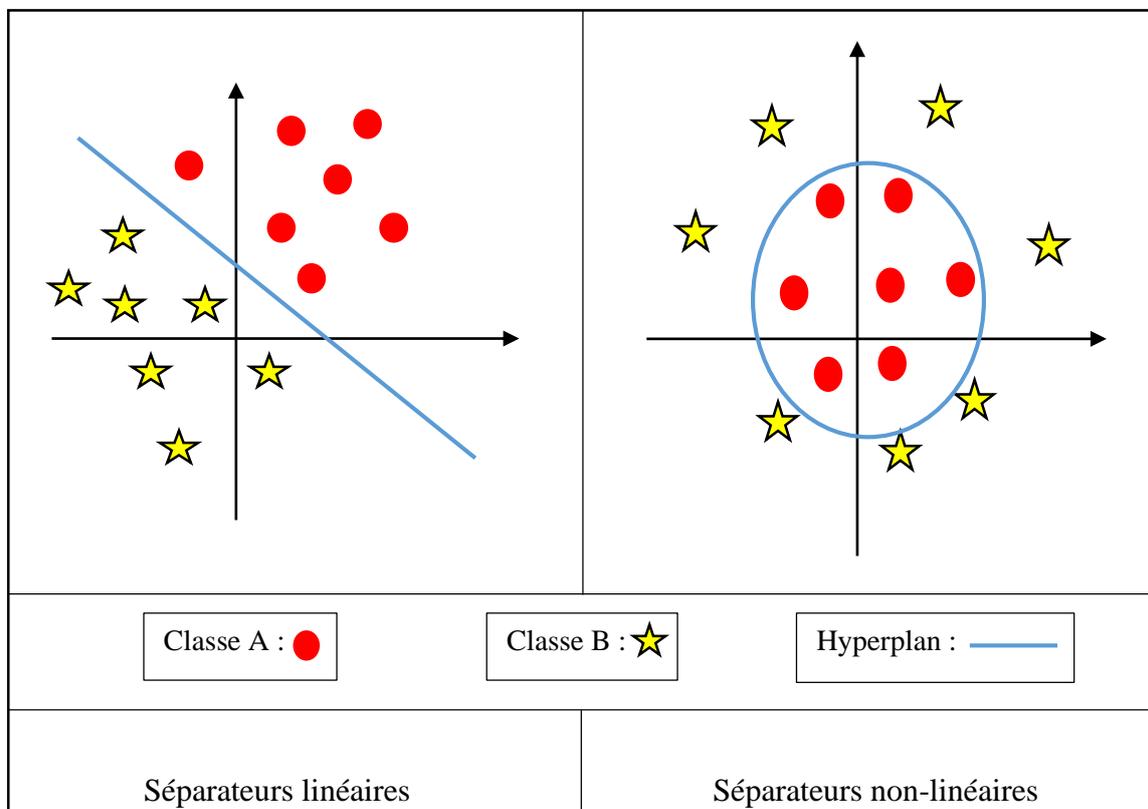


Figure 2.7 : Différents types des séparateurs.

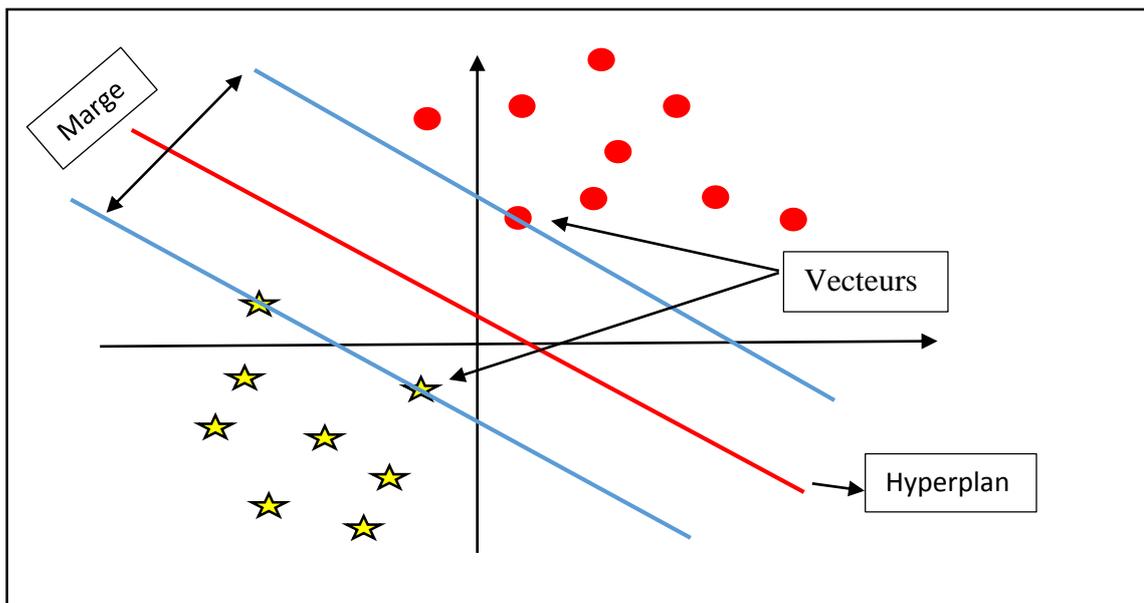


Figure 2.8 : Hyperplan optimal

Dans le cas linéaire on peut séparer les différents classes facilement mais dans le cas non-linéaire nous avons besoin d'espace plus grande dimension pour représenter les données d'entrée, et ceci à travers transformation des dimensions de \mathbb{R} (ensemble des nombres réels) à F (peut-être que la dimension est infinie) cette méthode est appelée '**Kernel trick**', ou astuce du noyau. Ceci est deuxième idée, dans cette nouvelle dimension probablement existe une séparation linéaire.

Les fonctions noyau permettent de transformer un produit scalaire dans un espace de grande dimension, ce qui est coûteux, en une simple évaluation ponctuelle d'une fonction [47].

$$K(x, y) = \lambda(x) \cdot \lambda(y) \quad (2.5)$$

Tels que x et y sont deux vecteurs d'entrée et λ c'est la transformation.

2.5.2 Distance euclidienne

Il y a plusieurs techniques de comparaison entre deux vecteurs caractéristiques du module d'extraction de caractéristiques d'un système de reconnaissance des expressions faciales. Parmi ces méthodes, celles qui dépendent du calcul de la distance (divergence).

Nous nous intéressons au calcul de la distance euclidienne entre des vecteurs représentant l'image d'entrée et avec chaque image stockés dans la base de données afin de déterminer leurs similitudes.

Soient $m \in \mathbb{N}^*$, $x = (x_1, \dots, x_m) \in \mathbb{R}$, et $y = (y_1, \dots, y_m) \in \mathbb{R}$, la loi de calcul de la distance euclidienne comme ce suit :

$$D(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (2.6)$$

2.6 Conclusion

Dans ce chapitre, nous avons présenté le processus d'un système de reconnaissance d'expressions faciales. Nous avons consacré notre étude à deux méthodes différentes et nous avons expliqué le principe de chacun et ses extensions récents. Nous avons également parlé sur des techniques de classification de l'expression faciale telle que SVM, distance euclidienne. Toutes ces méthodes seront implémentées puis comparées dans le chapitre 3.

Chapitre 03

Simulation et résultats

3.1 Introduction	36
3.2 Système de reconnaissance des expressions faciales	37
3.2.1 Prétraitement.....	38
3.2.2 Extraction des caractéristiques	39
3.2.2.1 Extraction des caractéristiques en utilisant HOG	39
3.2.2.2 Extraction des caractéristiques en utilisant LBP	40
3.2.3 Classification	42
3.3 Base de données	42
3.4 Résultats expérimentaux et discussions.....	43
3.4.1 Résultats obtenus	43
3.4.1.1 Histogramme de gradient orienté (HOG)	43
3.4.1.2 Motif binaire local (LBP)	46
3.5 Implémentation en temps réel	49
3.6 Conclusion.....	52

3.1 Introduction

Parmi toutes les fonctionnalités biométriques, la reconnaissance des expressions faciales est l'un des systèmes les plus importants puisqu'il est attaché aux changements physiologiques de l'être humain.

Pour cela, nous avons expliqué dans le chapitre précédent deux différents descripteurs d'extraction des caractéristiques (LBP, et HOG). Ces descripteurs sont utilisés pour l'extraction des changements dans le visage c.à.d. la position des sourcils, des yeux et la bouche. Ensuite, on a classifié l'information extraite selon six expressions à l'aide du classifieur SVM et la distance euclidienne.

Dans ce chapitre nous allons tester et comparer ces descripteurs (LBP, HOG) sur les images de la base de données JAFFE, pour mettre en évidence ses performances et ses précisions dans la reconnaissance des expressions faciales des individus. Ces descripteurs sont comparés selon des conditions de traitement prédéfinies, et également aussi en fonction du type de classifieur utilisé.

Afin de réaliser et appliquer tous les étapes de notre système de REF. Ces expériences sont implémentées sur le logiciel MATLAB version (2018a) qui est un environnement de calcul scientifique et de visualisation de données.

À la fin de ce chapitre, nous discuterons tous les résultats obtenus, et essayer de tirer quelques conclusions et faire des suggestions concernant le développement et l'amélioration des performances de ce système.

3.2 Système de reconnaissance des expressions faciales

Système de reconnaissance des expressions faciales est illustre dans le schéma suivant :

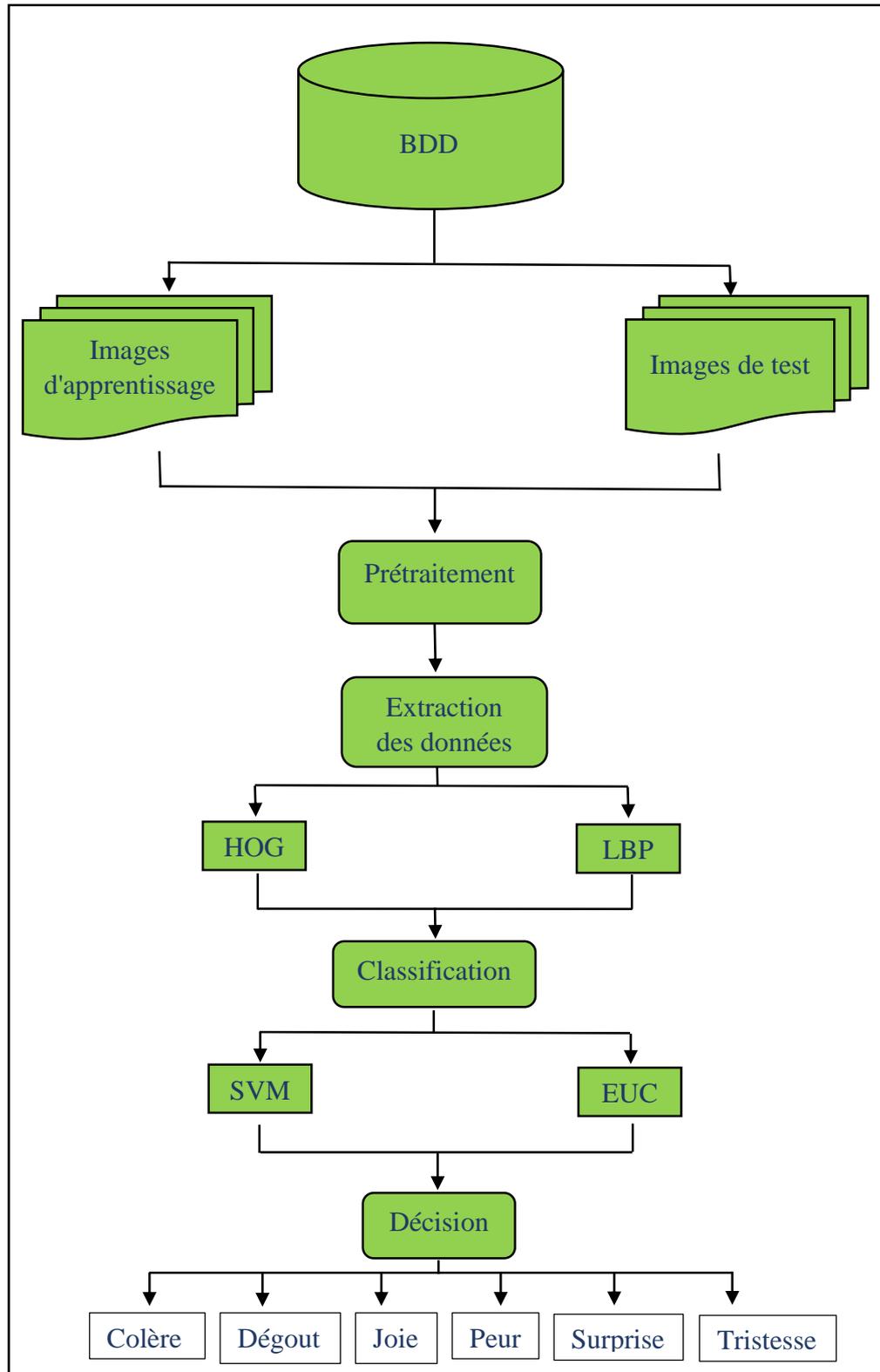


Figure 3.1 : Architecture du système de reconnaissance des expressions faciales

3.2.1 Prétraitement

Les images d'entrées de notre système sont d'abord converties en niveaux de gris. La détection des visages est effectuée par le détecteur de **Viola et Jones** [48], qui localise la zone de visage dans l'image à l'aide d'un rectangle qui spécifie les bords du visage. Ensuite, on coupe le visage pour éliminer les parties qui contiennent les informations inutile et concentrez sur la zone de visage, en particulier en ce qui concerne les sourcils, les yeux et la bouche (fig3.2).

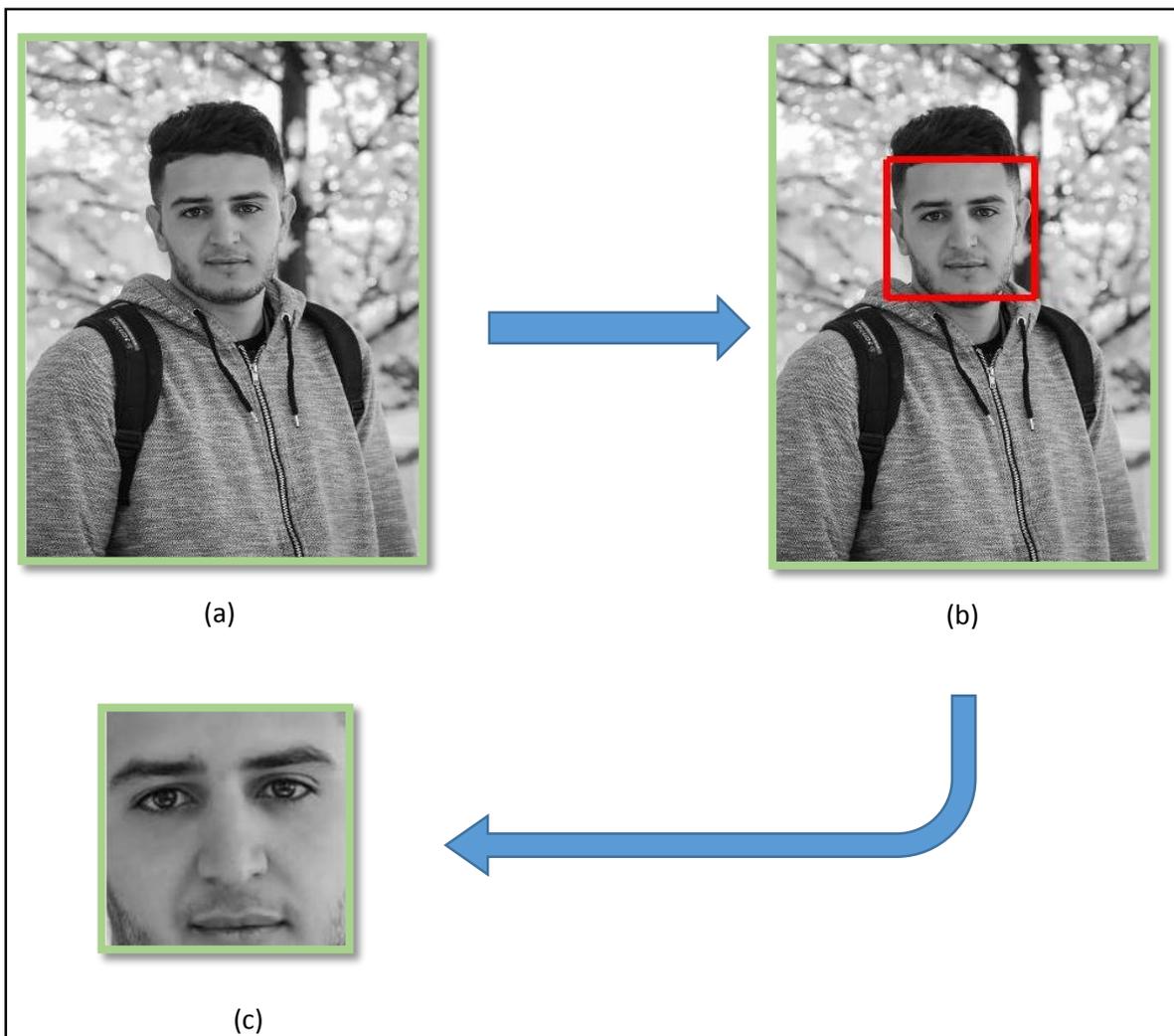


Figure 3.2 : Prétraitement d'image d'entrée. (a) image en niveau de gris, (b) détection du visage, (c) visage coupé.

À chaque fois qu'un visage est détecté, ensuite il est amélioré à l'aide d'une méthode d'égalisation histogramme qui est utilisée pour compenser les conditions d'éclairage et améliorer le contraste d'image. Elle consiste à appliquer une transformation sur chaque pixel de l'image, et donc d'obtenir une nouvelle image à partir d'une opération indépendante sur chacun des pixels afin d'obtenir un nouveau histogramme de l'image (fig3.3).

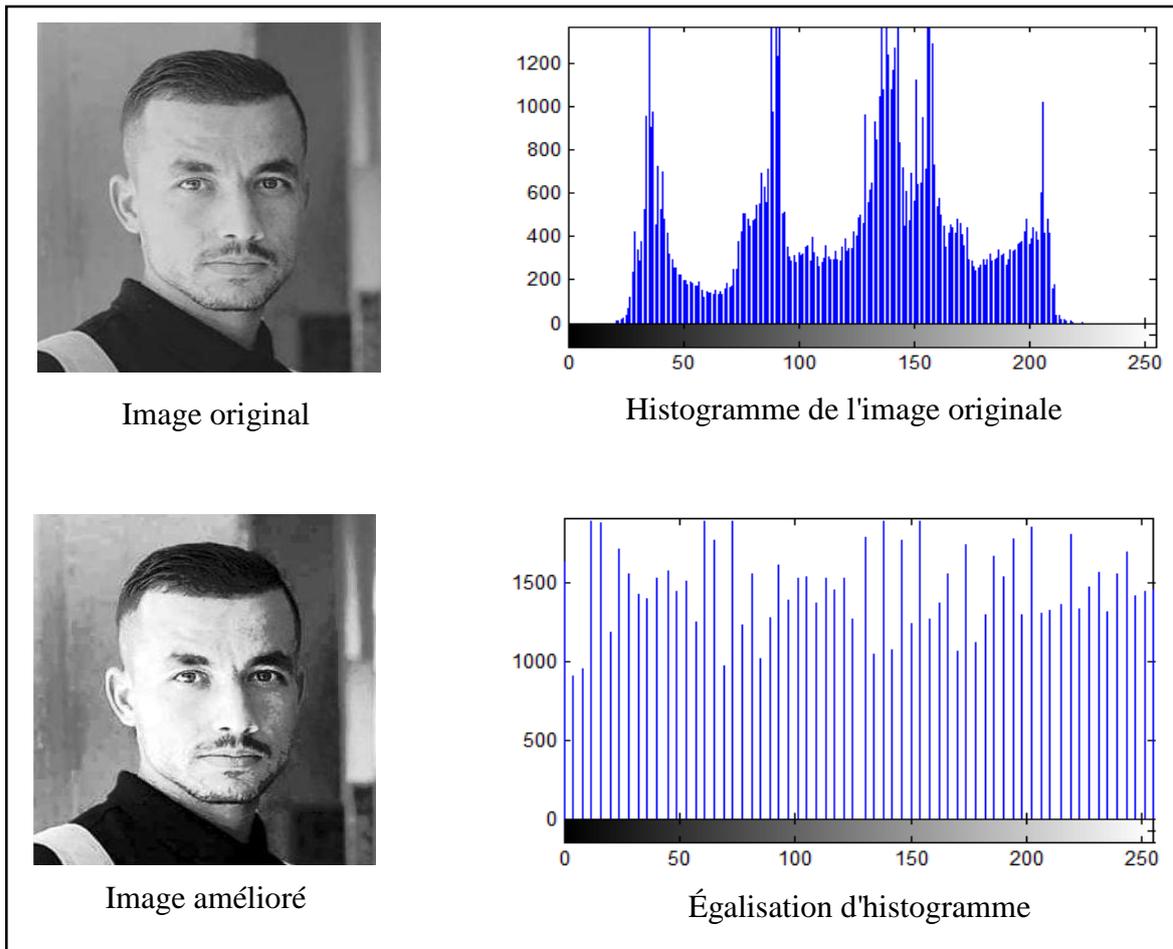


Figure 3.3 : Exemple d'égalisation d'histogramme

3.2.2 Extraction des caractéristiques

3.2.2.1 Extraction des caractéristiques en utilisant HOG

Toutes les images de visage extraites sont transmises au descripteur HOG pour l'extraction de caractéristiques. Le descripteur HOG divise l'image en blocs de taille de 16×16 . Nous obtenons un total de $16 \times 16 = 256$ blocs. Chaque bloc est constitué de 2×2 cellules de taille 8×8 . La plage des orientations du gradient est définie entre $[0^\circ-360^\circ]$ ensuite, il est quantifié à l'aide d'histogrammes de 9 cases (bins).

De cette manière, le nombre total de caractéristiques HOG pour une image d'entrée de taille 256×256 devient 8100.

Le descripteur HOG est caractérisé par deux principaux paramètres, la taille de la cellule et le nombre de cases d'orientation.

La taille de cellule représente la dimension du patch impliqué dans le calcul d'un seul histogramme. Quand on utilise des cellules de grande taille, les informations d'apparence

d'une région significative de l'image faciale sont comprimées dans un histogramme de cellule unique, puis certains détails utiles pour la classification ultérieure peuvent être perdus. Par contre, avec des cellules de petite taille, une analyse à haute résolution peut être effectuée mais, de cette manière, la distinction entre détails extraits utiles et inutiles est demandée au classificateur qui pourrait ne pas être en mesure d'effectuer cette tâche supplémentaire de la meilleure façon possible (fig3.4).

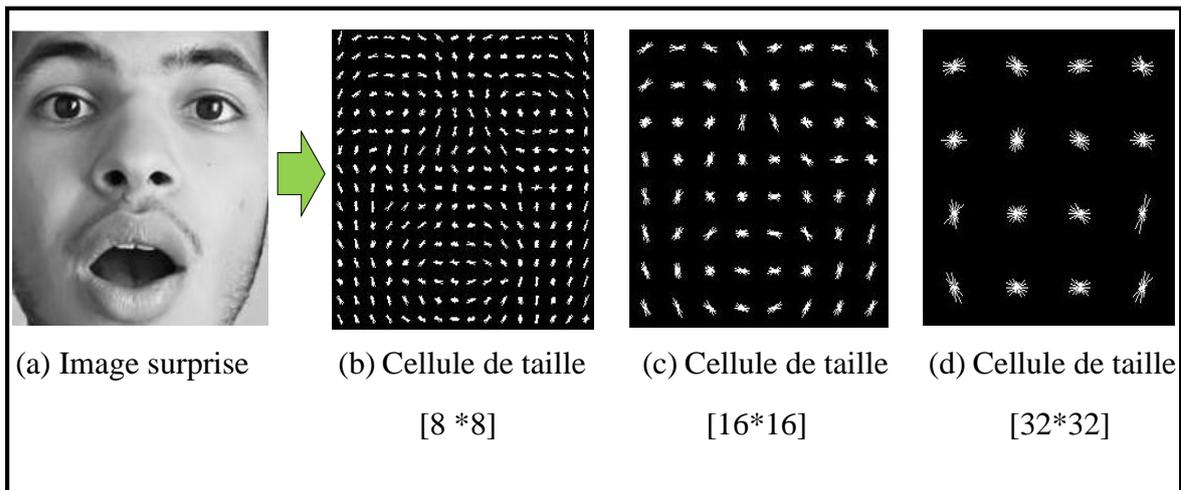


Figure 3.4 : Visualisation de descripteur HOG d'une image avec différents tailles de la cellule.

Le nombre de cases d'orientation fait référence aux niveaux de quantification des informations de gradient. Un faible nombre d'orientations pourrait entraîner une perte d'informations et donc une diminution des performances du système de REF. Au contraire, un grand nombre d'orientations pourraient écartier les informations le long des bacs, ce qui diminuerait également les performances de système de REF. Pour ces raisons, ces paramètres doivent être choisis avec précision.

3.2.2.2 Extraction des caractéristiques en utilisant LBP

Après la détection du visage dans une image et l'avoir coupé, il est transmis à l'étape d'extraction des caractéristiques faciales on utilisant l'opérateur LBP, où l'image résultante est divisée en blocs de taille de 3*3. Les pixels de chaque bloc sont soumis à un seuil déterminé par la valeur du pixel central.

Comme le voisinage est composé de 8 pixels, il est possible d'obtenir un total de $2^8=256$ différentes valeurs en fonction des valeurs de gris relatives du centre et des pixels du voisinage. Les histogrammes de tous les blocs sont calculés puis normalisés et concaténés dans un seul histogramme final.

Pour collecter les informations sur grandes régions, nous utilisons des cellules plus grandes, mais lorsque la taille de la cellule augmente, les détails locaux sont perdus. Pour cela nous avons spécifié la taille des cellules de 16×16 pixels pour obtenir une meilleure représentation des détails du visage (fig3.5).

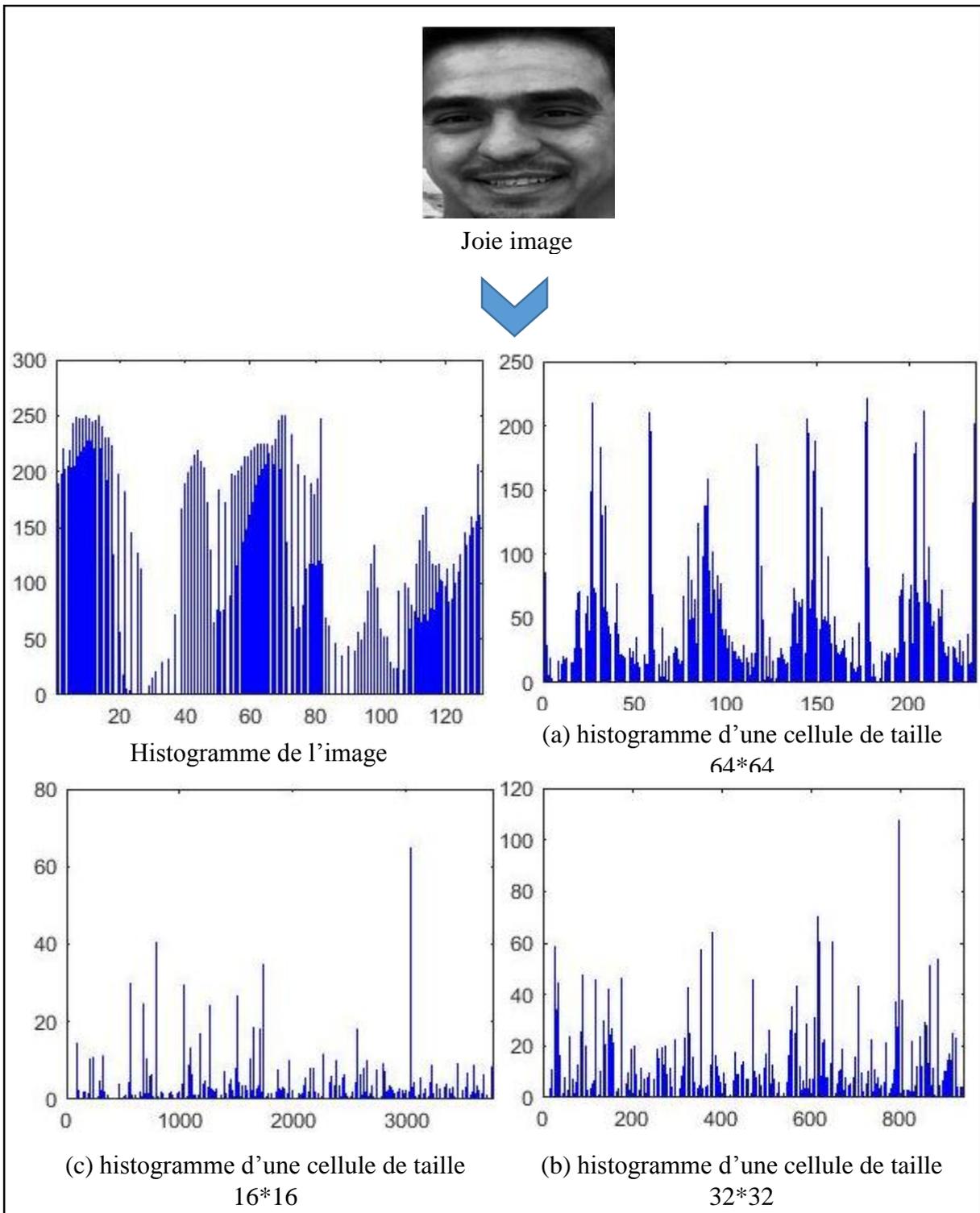


Figure 3.5 : Extraction des caractéristiques avec l'opérateur LBP selon la taille des cellules.

3.2.3 Classification

Avant de classifier les images et d'identifier les expressions faciales de chaque personne, les images originales des différentes personnes de la base de données utilisée (JAFPE) subissent la phase de prétraitement, afin d'obtenir une image plus nette du visage ainsi que l'élimination du bruit, ensuite extraire les propriétés de chaque visage en utilisant l'une des deux méthodes que a mentionnées précédemment.

Ces caractéristiques sont placées dans un vecteur. A la fin de cette étape, où obtient une matrice de taille (m*n), où chaque ligne représente un vecteur d'une image, où l'indice 'n' indique les valeurs de chaque vecteur, et l'indice 'm' représente le nombre des images de test ou d'entraînement.

Nous avons appliqué la tâche de classification en deux manières différentes : SVM, et le calcul de la distance euclidienne. Le vecteur de l'image test est comparé à la matrice des images d'entraînement du modèle de chaque classe, où la classe correspondant à l'expression du visage de test est détectée.

3.3 Base de données

Dans notre travail, on a utilisé la base de données JAFPE (Japanese Female Facial Expression), est une base de données d'image de visage gratuite. Elle contient sept expressions faciales (6 expressions faciales de base + 1 neutre) posées par 10 modèles féminins japonais y compris les individus : KA, KL, KM, KR, MK, NA, NM, TM, UY et YM, au total 213 images. Tous les personnes dans cette base ont les sept émotions universelles (colère, dégoût, peur, joie, tristesse, surprise et neutre) (fig3.6), chaque expression contient 2 ou 3 échantillons de l'individu avec un total de 30 images par émotion, et totalement on a 213 images de taille 256×256 pixel.

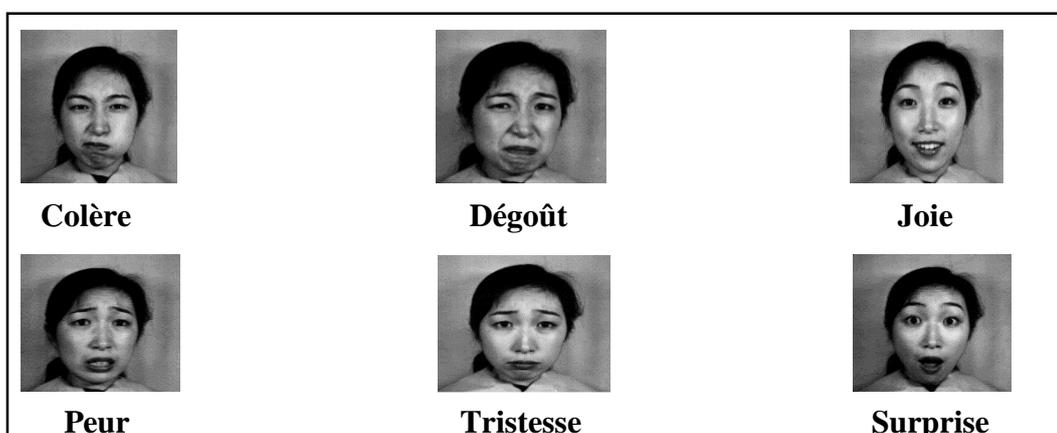


Figure 3.6 : Exemples des 6 expressions de la base de données JAFPE

3.4 Résultats expérimentaux et discussions

Dans nos expériences sur la base de données JAFFE, nous avons utilisé 183 images (sans les images neutres) divisées comme ce suit :

•Base d'apprentissage

Dans notre cas nous avons 120 échantillons ou image de 10 individus, trois ou deux images par individus pour chaque émotion : colère (20), dégoût (20), joie (20) peur (20), tristesse (20) et surprise (20). Ensuite les images de différents sujets appartenant à une même classe d'émotions soient regroupées pour former la base de données d'apprentissage. Par exemple les images qui représentent l'émotion 'joie' sont situées dans une classe unique, afin que nous obtenions une base de données de six classes représentant les six expressions faciales.

•Base de test

Nous avons choisi une ou deux images de chaque personne pour chaque expression, donc notre base de test est composée de 63 images avec différentes expressions faciales de 10 individus. À chaque fois on compare un visage de test avec la base d'apprentissage.

3.4.1 Résultats obtenus

Nous avons divisé nos résultats en deux phases en fonction du type de descripteur utilisé pour l'extraction des caractéristiques soit LBP ou HOG.

Chaque phase est constituée de 4 expériences différentes :

1^{er} : Classification à base de SVM.

2^{ème} : Classification à base de SVM et égalisation d'histogramme.

3^{ème} : Classification à base de calcul de distance euclidienne.

4^{ème} : Classification à base de calcul de distance euclidienne et égalisation d'histogramme.

Les résultats obtenus sont les suivants :

3.4.1.1 Histogramme de gradient orienté (HOG)

a) Le classifieur SVM

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	80	20	0	0	0	0
Dégout	0	100	0	0	0	0
Joie	0	0	72.72	27.27	0	0
Peur	0	0	0	81.82	0	18.18
Surprise	0	0	0	0	100	0
Tristesse	0	0	0	0	0	100
Total	89.09					

Tableau 3.1 : Matrice de confusion utilisant le descripteur HOG et le classifieur SVM (sans égalisation d'histogramme).

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	80	20	0	0	0	0
Dégout	0	100	0	0	0	0
Joie	0	0	81.82	18.18	0	0
Peur	0	0	0	90.91	0	9.09
Surprise	0	0	0	0	100	0
Tristesse	0	0	0	0	0	100
Total	92.12					

Tableau 3.2 : Matrice de confusion utilisant le descripteur HOG et le classifieur SVM (avec égalisation d'histogramme).

b) Le classifieur euclidienne

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	90	10	0	0	0	0
Dégout	10	70	0	0	0	20
Joie	0	0	72.72	9.09	9.09	9.09
Peur	9.09	0	0	63.64	0	27.27
Surprise	0	0	0	0	100	0
Tristesse	0	0	0	0	0	100
Total	82.72					

Tableau 3.3 : Matrice de confusion utilisant le descripteur HOG et le classifieur euclidienne (sans égalisation d'histogramme).

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	90	10	0	0	0	20
Dégout	10	80	0	0	0	10
Joie	0	0	81.82	0	9.09	9.09
Peur	0	0	0	90.91	0	9.09
Surprise	0	0	0	0	100	0
Tristesse	0	0	0	0	0	100
Total	90.45					

Tableau 3.4 : Matrice de confusion utilisant le descripteur HOG et le classifieur euclidienne (avec égalisation d'histogramme).

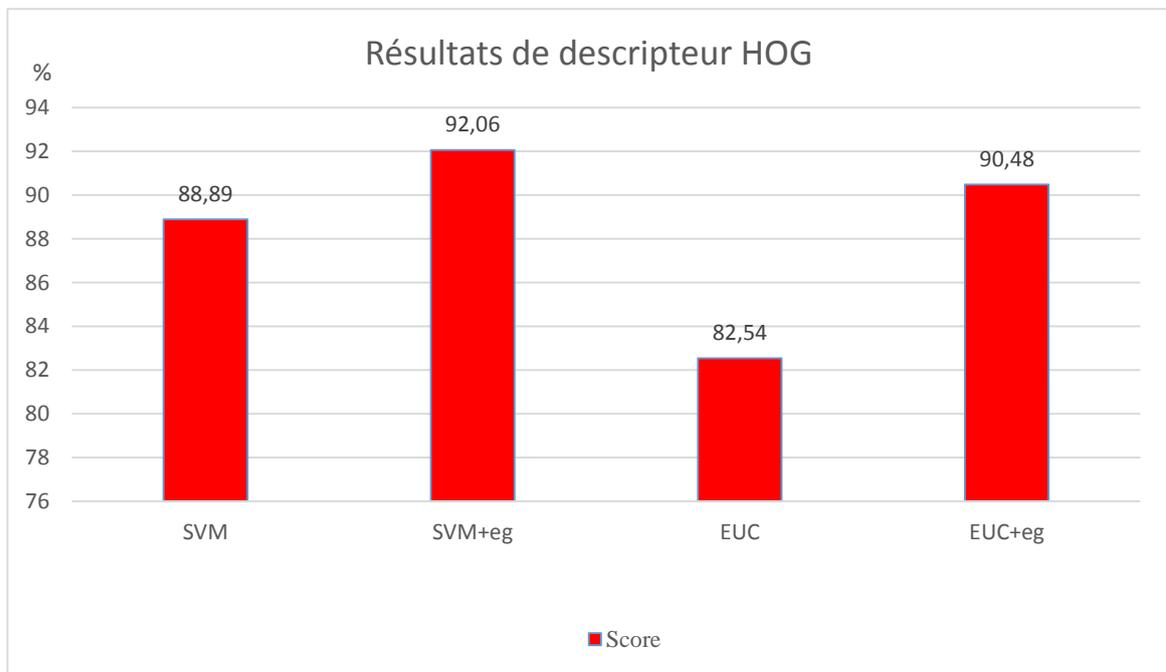


Figure 3.7 : Représentation graphique des résultats avec l'opérateur HOG (score).

Discussion1 :

La première phase de notre expérience est basée sur le descripteur HOG pour extraire les données faciales.

Le (tableau 3.1) montre la performance du système à l'aide du classifieur SVM. Nous avons obtenu un score de 88.89%. Après avoir appliqué l'égalisation d'histogramme aux images d'apprentissage et de test, on observe que le pourcentage de reconnaissance de toutes les expressions par le système était passé à 92,06%.

Les résultats obtenus en utilisant le classifieur euclidienne sont illustrés dans le (tableau 3.3), où on a un score de 82.54%. Ces résultats augmentent quand on applique l'égalisation d'histogramme sur les images de la base de données JAFFE, jusqu'à 90.48%.

L'effet d'égalisation d'histogramme, est évident dans l'expression de la joie, lorsque on utilise le classifieur SVM, où la reconnaissance de l'émotion a augmenté de 88.89% à 92.06%, ainsi que dans les expressions de colère, dégoût et de joie lorsque nous utilisons le classifieur euclidienne de 82.54% à 90.48%.

3.4.1.2 Motif binaire local (LBP)

a) Le classifieur SVM

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	90	10	0	0	0	0
Dégout	0	80	0	10	0	10
Joie	0	0	81.82	9.09	9.09	0
Peur	0	0	0	81.82	0	18.18
Surprise	0	0	0	10	90	0
Tristesse	10	0	0	0	0	100
Total	87.27					

Tableau 3.5 : Matrice de confusion utilisant le descripteur LBP et le classifieur SVM (sans égalisation d'histogramme).

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	80	10	0	0	0	10
Dégout	0	90	0	10	0	0
Joie	0	0	90.91	9.09	0	0
Peur	0	0	0	90.91	0	9.09
Surprise	0	0	0	10	90	0
Tristesse	0	0	0	0	0	100
Total	90.30					

Tableau 3.6 : Matrice de confusion utilisant le descripteur LBP et le classifieur SVM (avec égalisation d'histogramme).

b) Le classifieur euclidienne

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	70	10	10	0	0	10
Dégout	10	60	0	0	0	30
Joie	0	0	81.82	0	9.09	9.09
Peur	0	0	0	54.55	9.09	36.36
Surprise	0	0	0	0	100	0
Tristesse	10	0	0	0	0	100
Total	77.72					

Tableau 3.7 : Matrice de confusion utilisant le descripteur LBP et le classifieur euclidienne (sans égalisation d'histogramme).

Émotions	Colère %	Dégout %	Joie %	Peur %	Surprise %	Tristesse %
Colère	90	0	0	10	0	0
Dégout	10	50	10	20	0	10
Joie	0	0	90.91	0	0	9.09
Peur	0	0	0	81.82	0	18.18
Surprise	0	0	0	10	90	0
Tristesse	9.09	0	0	18.18	9.09	63.64
Total	77.72					

Tableau 3.8 : Matrice de confusion utilisant le descripteur LBP et le classifieur euclidienne (avec égalisation d'histogramme).

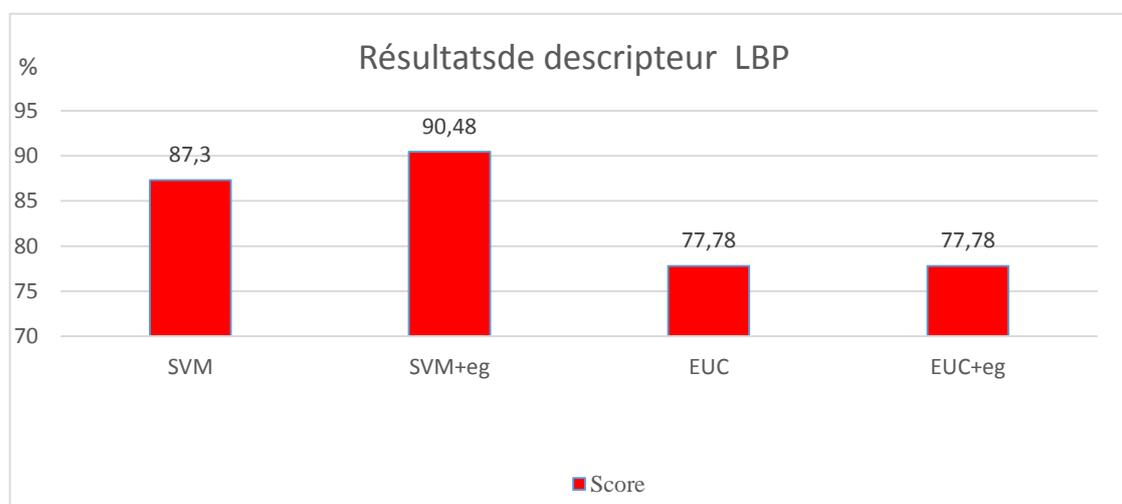


Figure 3.8 : Représentation graphique des résultats avec l'opérateur LBP (score)

Discussion 2 :

La deuxième partie de notre expérience est basée sur l'utilisation du descripteur LBP pour extraire les données faciales.

Le degré de reconnaissance de notre système qui est basé sur le descripteur LBP sur différentes expressions du visage est 87.30% avec le classifieur SVM. Après l'utilisation de l'égalisation d'histogramme sur les images de la base de données JAFFE, la performance du système a augmenté jusqu'à 90,48%.

Quand on adopte, le classifieur euclidienne, nous avons obtenus des résultats presque identique, avant et après l'utilisation de l'égalisation d'histogramme, de l'ordre de 77.78%.

L'égalisation d'histogramme a augmenté le degré de réussite de la reconnaissance des expressions de dégoût, joie, et de peur, avec une augmentation d'environ 10% pour chaque expression quand on utilise le classifieur SVM. Au contraire, lorsque on utilise le classifieur euclidienne, nous obtenons le même résultat même après égalisation d'histogramme qui est 77.78%, bien que les expressions de colère, dégoût, joie et de peur, où a amélioré leurs résultats, tels que la peur qui a augmenté de 54.55% à 81.82%.

- Après avoir analysé les résultats précédents, nous avons observé que l'égalisation d'histogramme avait un rôle important dans l'amélioration des performances du système, car il élimine l'effet d'illumination et ajuste le contraste d'image. L'image originale était très claire, cependant, après l'application de ce traitement, les zones sensibles d'expression faciales (les sourcils, la bouche, les yeux) deviennent plus éminent, ce qui facilite la classification de ces six expressions.

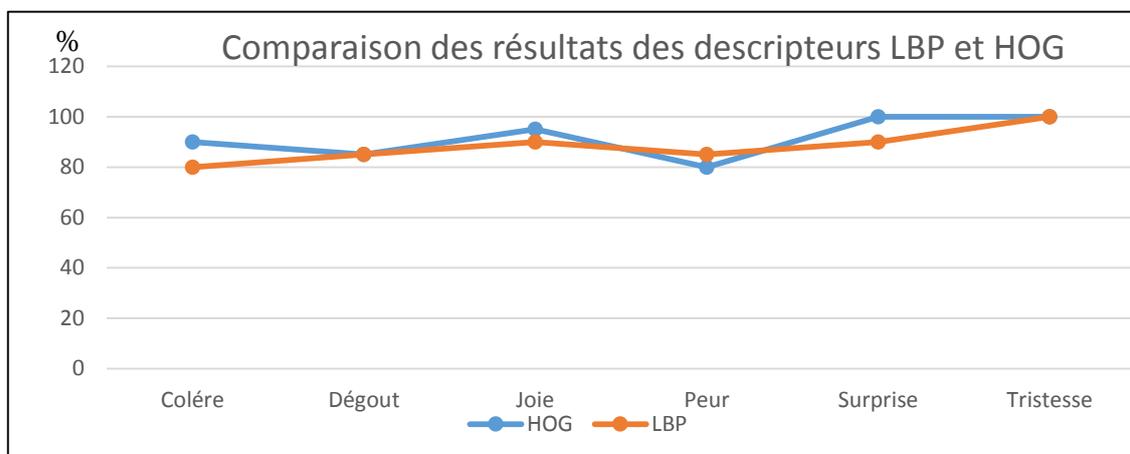


Figure 3.9 : Représentation graphique des meilleurs résultats de descripteurs HOG et LBP avec les six émotions.

•Comparaison entre les descripteurs HOG et LBP :

D'après les résultats obtenus précédemment, et des diagrammes en figure 3.7 et figure 3.8 et figure 3.9. Nous concluons que le descripteur HOG est plus efficace que le descripteur LBP dans l'identification des expressions faciales.

Autres études		Notre étude	
LBP+ Réseau de neurones	HOG+SVM	LBP+SVM	HOG+SVM
88.5%	89.5%	90,48%	92,06%

Tableau 3.9 : Comparaison des performances de notre approche par rapport à différentes approches de l'état de l'art [36] [49].

Discussion 3 :

Dans le tableau 3.5 les résultats obtenus sont comparés avec différents approches de l'état de l'art. Nous remarquons que les résultats obtenus sont très encourageants.

Dans une étude comparative de la reconnaissance des expressions faciale, la reconnaissance a basée sur les caractéristiques de LBP, donné un résultat de 88.5%, où la classification a été faite par le classifieur de réseau de neurones, sur la base de données JAFFE.

Une autre étude basée sur le descripteur HOG, où le degré de reconnaissance des expressions faciales a atteint 89.5%. Ils sont également basés sur le classifieur SVM pour la classification des expressions faciales, sur la base de données JAFFE.

Lorsqu'on compare les résultats des deux études précédentes et les résultats obtenus dans notre étude sur la reconnaissance des expressions faciales, en utilisant les descripteurs : LBP, et HOG. Nous avons observé que la performance de notre système était très bonne, ceci est illustré par la comparaison des résultats dans le tableau3.5.Où les résultats de notre travail, sont les meilleurs résultats de la reconnaissance des expressions faciales avec différents descripteurs (LBP, HOG), sur la base de données JAFFE.

3.5 Implémentation en temps réel

Dans la mise en œuvre d'un système de reconnaissance faciale, nous avons conçu une application qui fonctionne en temps réel à l'aide du logiciel MATLAB version (2018a). Où a utilisé une capture de l'image du visage à travers une caméra vidéo (webcam) après avoir

appuyé sur le bouton de capture. La phase de prétraitement est ensuite appliquée à l'image capturée, où affiche l'image traitée à l'écran. Le résultat est affiché directement dans des boîtes spécifiques selon le type de descripteur utilisé pour extraire les traits du visage : HOG, ou LBP. Le processus de classification est effectué par le classifieur SVM en raison de son efficacité par rapport le classifieur euclidien (fig3.10).

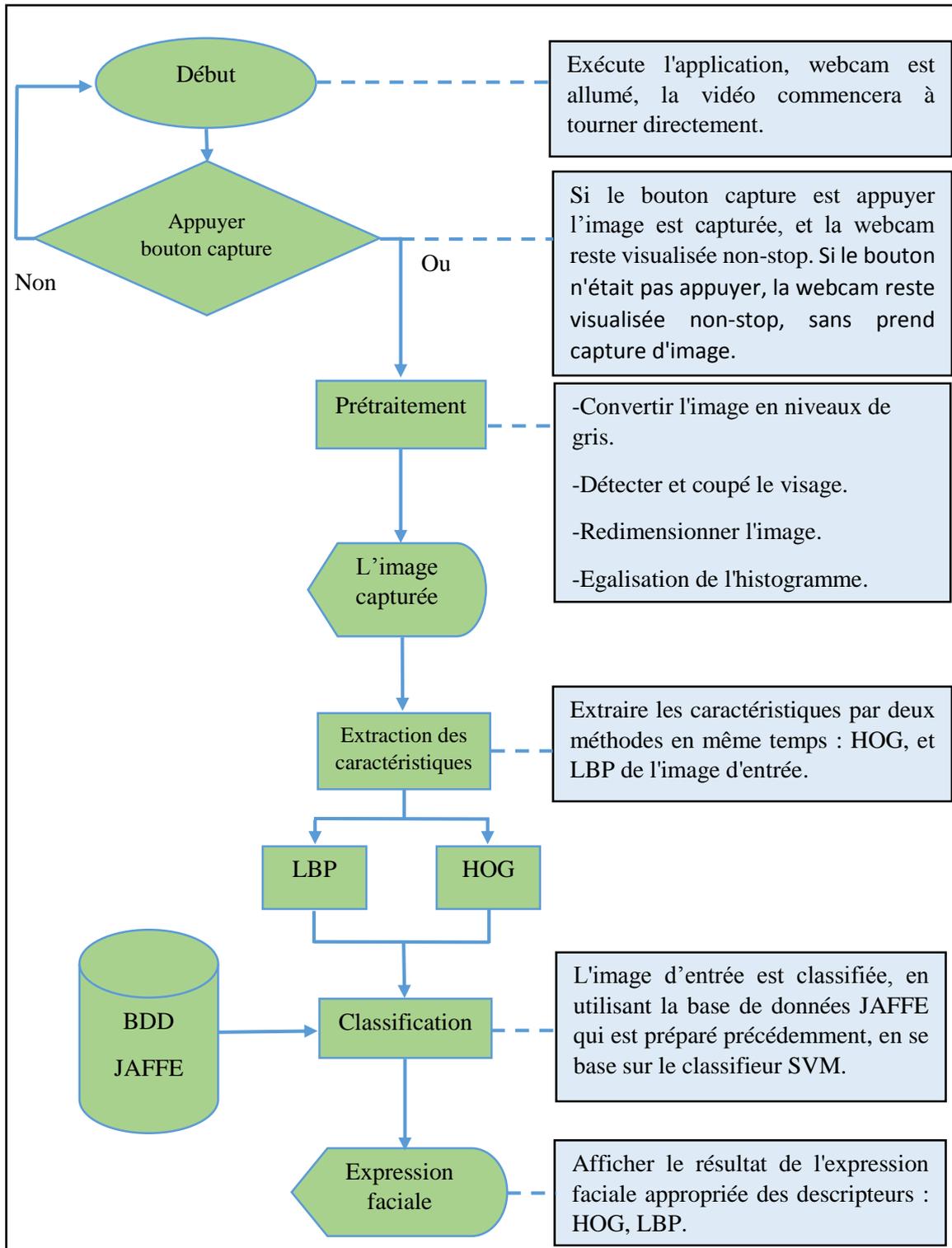


Figure 3.10 : Organigramme du système REF.

L'application de reconnaissance des expressions faciales, que nous avons réalisée, a donné des bons résultats dans la reconnaissance de l'émotion des différentes personnes testées.

La figure suivante illustre la forme finale de l'application réalisée à travers notre étude avec quelque exemple de différentes expressions faciales :

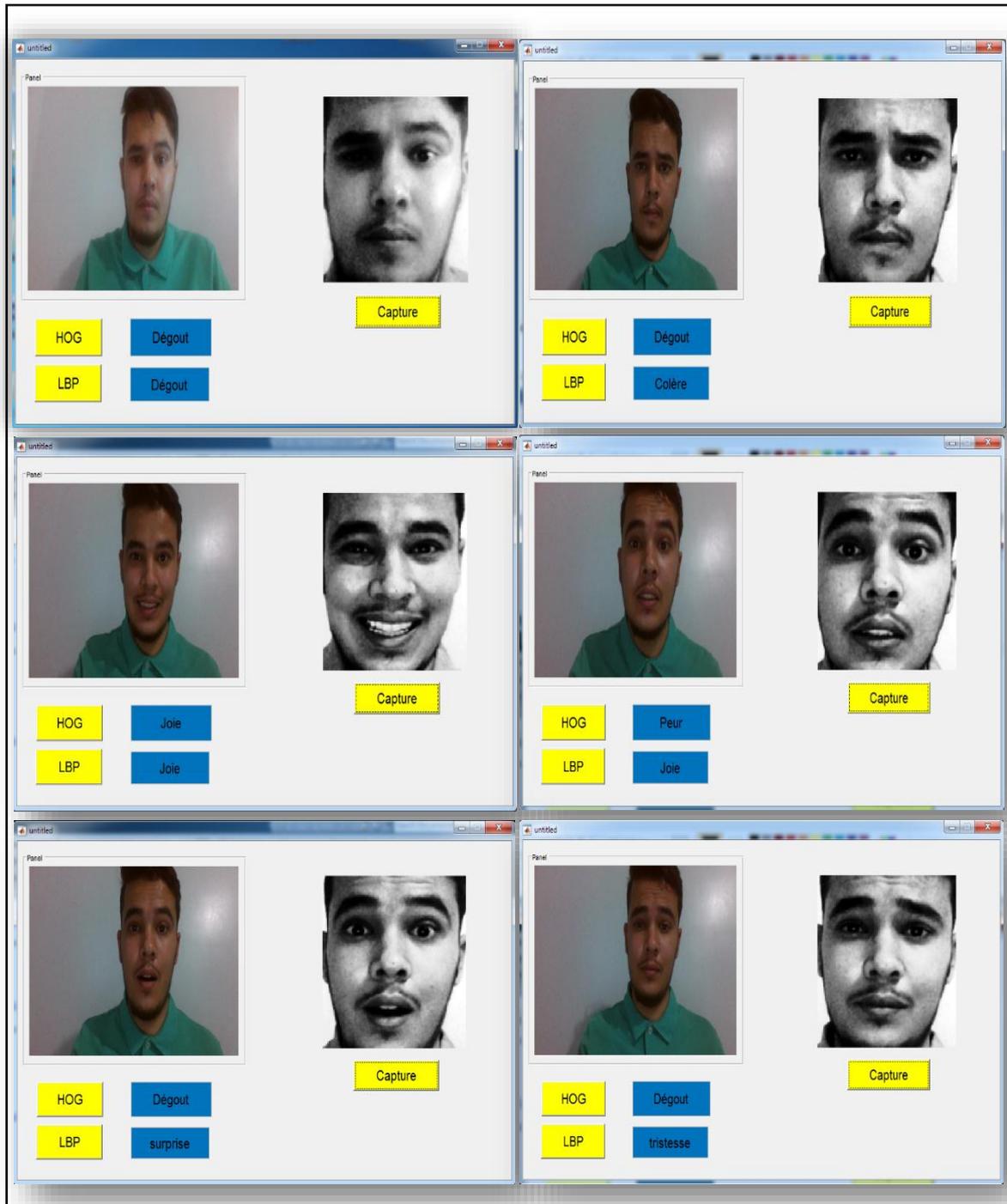


Figure 3.11 : Interface de l'application de reconnaissance des expressions faciales

3.6 Conclusion

Dans ce chapitre, nous avons présenté les différentes étapes de l'implémentation de notre système de reconnaissance des expressions faciale.

Nous avons effectué plusieurs expériences en se basant sur différentes aspect. Nous avons varié les paramètres de l'étape de prétraitement avec et sans l'égalisation d'histogramme, et les méthodes d'extraction des caractéristiques faciale : HOG et LBP, et aussi la méthode de classification multi-class qui est basée sur l'un des classifieur : SVM ou distance euclidien.

Les résultats obtenus pour la classification des émotions à l'aide du classifieur SVM étaient meilleurs et plus précis que le classifieur euclidien, pour les deux méthodes que nous avons utilisé précédemment.

Conclusion générale

L'objectif de ce mémoire est d'étudier, concevoir et implémenter un système de reconnaissance des émotions. Où de nombreux domaines sont devenus intéressés aux techniques de reconnaissance des expressions faciales (psychologie, psychologie sociale, neurophysiologie, informatique, domaine médical, marketing ...).

Dans notre travail nous avons présenté un état de l'art des différentes méthodes récentes dans ce domaine, qui sont divisés en deux catégories : les approches d'apparence (LBP, LBQ, HOG, Ondelette de Gabor...), et Les approches géométriques (ASM, AAM...). Avant d'aborder ces méthodes, nous avons démontré le principe de fonctionnement d'un système biométrique en général et y compris ses principaux modules, dans divers modalités biométriques (l'empreinte digitale, le visage, la signature, veines de la main...).

Après une étude approfondie des techniques utilisées pour identifier les expressions faciales, nous avons présenté les trois principaux étapes du système de reconnaissance de l'expression faciale, en commençant par l'étape de préparation de l'image ,puis l'extraction des caractéristiques, et la classification, en utilisant les techniques que nous avons choisi et expliquées dans le 2ème chapitre.

L'étape de prétraitement a été le premier défi que nous avons rencontré lors de la mise en œuvre du programme. Après la détection du visage par l'algorithme de Viola et Jones, nous avons obtenu une image complète du visage, y compris les oreilles et la tête, bien que nous nous intéressons aux zones de les sourcils, la bouche et les yeux uniquement. Alors, on a coupé l'image avec des coordonnées bien déterminé, on a obtenu une forme de visage qui nous permet d'extraire ses propriétés avec précision. Un autre problème que nous avons rencontré dans cette étape, concernait les effets de l'environnement externe sur la clarté des détails des expressions faciales, pour cela nous avons essayé de réduire les effets de la lumière et du bruit, et améliorer le contraste d'image en appliquant l'égalisation de l'histogramme.

Dans la phase d'implémentation nous avons effectué plusieurs expérience, en se basant sur les différents types de descripteurs d'une part, et sur le type de classifieur d'autre part, en tenant compte les modifications apportées aux paramètres de prétraitement. Toutes ces expériences ont été menées sur la base de données JAFFE, comme décrit dans le 3ème chapitre

Conclusion générale

D'après les résultats obtenus et présentés dans le 3^{ème} chapitre , les meilleurs résultats obtenus étaient basés sur le classifieur SVM, où le score de reconnaissance des expressions faciales a atteint 92.06% et 90.48% , pour les descripteurs HOG et LBP successivement. Aussi nous avons également obtenu un score de 90.48% avec le descripteur HOG et classifieur euclidienne.

Nous avons observé que le choix de type du classifieur, et le prétraitement d'images telles que l'égalisation de l'histogramme avant l'étape de l'extraction des caractéristiques, jouent un rôle très important dans l'amélioration des résultats et augmentent de l'efficacité du système d'identification des expressions faciales, et ce qui explique la variabilité des résultats en fonction des conditions expérimentales.

Comme perspectives, on propose d'amélioré ce travail par :

- › Essayer une autre base de données universelle d'autant plus que la base JAFFE qui est une collection de visages asiatiques très similaires.
- › Essayez d'améliorer le prétraitement de l'image, par exemple en utilisant des filtres.
- › Utiliser d'autre descripteurs ou fusionner plusieurs descripteurs ensemble.

Bibliographie

- [1] AHMED ZAID, Aghiles. Segmentation d'images appliquée en biométrie : Cas des images de l'Iris. Thèse de doctorat. Université Mouloud Mammeri, 2013.
- [2] CHOUCANE, Ammar. Analyse d'images d'expressions faciales et orientation de la tête basée sur la profondeur. Thèse de doctorat. Université Mohamed Khider-Biskra, 2016.
- [3] BALLIHI, Lahoucine. Biométrie faciale 3D par apprentissage des caractéristiques géométriques : application à la reconnaissance des visages et à la classification du genre. Thèse de doctorat. Université des Sciences et Technologie de Lille-Lille I, 2012.
- [4] <https://www.aware.com/portfolio-items/quest-ce-que-la-biometrie-livre-blanc/>
- [5] ZITOUNI, Sif Eddine et SACI, Abdelmoumen. Authentification et Identification biométrique des personnes par les empreintes palmaires. Université Kasdi Merbah, Ouargla, 2016.
- [6] <http://www.linflux.com/litterature/piques-de-polars/>
- [7] <https://www.guide-vue.fr/glossaire/ophtalmie>
- [8] [généralité sur la biométrie]. (s.d.). Récupéré 24 juin, 2019, de http://www.univ-tebessa.dz/fichiers/master/master_2168.
- [9] <https://www.frank-tueren.ch/fr/biometrie-du-reseau-veineux-de-la-paume-de-la-main>.
- [10] <http://www.mysti2d.net/polynesie2/ETT/C052/23/SerruresBioMeca/files/Documents/Biom%C3%A9trie>
- [11] https://fr.wikipedia.org/wiki/Expression_faciale.
- [12] DANALET, Antonin. Modèles de choix discrets pour la reconnaissance des expressions faciales statiques. 2007.
- [13] ABDAT, Faiza. Reconnaissance automatique des émotions par données multimodales : expressions faciales et signaux physiologiques. Université de Metz, France, 2010.
- [14] <http://www.lilhoot.eu/quelles-sont-nos-emotions-de-base-et-les-expressions-faciales-permettent-elles,a70>.
- [15] <https://psychologycompass.com/blog/scientific-tactics-that-boost-non-verbal-communication-and-body-language/>

- [16] https://fr.wikipedia.org/wiki/Robert_Plutchik#Th%C3%A9orie_des_%C3%A9motions.
- [17] CHITRA, Nivedita et NIJHAWAN, Geeta. Facial expression recognition using local binary pattern and support vector machine. *Int J Innovative Res Adv Eng*, 2016, p. 103-108.
- [18] KAUSER, Nazima et SHARMA, Jitendra. Facial expression recognition using LBP template of facial parts and multilayer neural network. In: 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). IEEE, 2017. p. 445-449.
- [19] WANG, Zhen et YING, Zilu. Facial expression recognition based on local phase quantization and sparse representation. In: 2012 eighth International Conference on Natural Computation. IEEE, 2012. p. 222-225.
- [20] DHALL, Abhinav, ASTHANA, Akshay, GOECKE, Roland, et al. Emotion recognition using PHOG and LPQ features. In: *Face and Gesture 2011*. IEEE, 2011. p. 878-883.
- [21] ANE, Tanjea et PATWARY, Md Fazlul Karim. Performance Analysis of Similarity Coefficient Feature Vector on Facial Expression Recognition. *Procedia engineering*, 2016, vol. 144, p. 444-451.
- [22] CARCAGNÌ, Pierluigi, DEL COCO, Marco, LEO, Marco, et al. Facial expression recognition and histograms of oriented gradients: a comprehensive study. SpringerPlus, 2015, vol. 4, no 1, p. 645.
- [23] ORRITE, Carlos, GAÑÁN, Andrés, et ROGEZ, Grégory. Hog-based decision tree for facial expression classification. In: *Iberian Conference on Pattern Recognition and Image Analysis*. Springer, Berlin, Heidelberg, 2009. p. 176-183.
- [24] RUSSELL, Brian. Dennis Gabor: The father of seismic attribute analysis. CREWES Research Report, 2012.
- [25] YE, Jingfu, ZHAN, Yongzhao, et SONG, Shunlin. Facial expression features extraction based on Gabor wavelet transformation. In: 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583). IEEE, 2004. p. 2215-2219.
- [26] https://en.wikipedia.org/wiki/Gabor_wavelet.

- [27] ZHAN, Yongzhao, YE, Jingfu, NIU, Dejiao, et al. Facial expression recognition based on Gabor wavelet transformation and elastic templates matching. *International Journal of Image and Graphics*, 2006, vol. 6, no 01, p. 125-138.
- [28] COOTES, Timothy F., TAYLOR, Christopher J., COOPER, David H., et al. Active shape models-their training and application. *Computer vision and image understanding*, 1995, vol. 61, no 1, p. 38-59.
- [29] LEI, Gang, LI, Xiao-hua, ZHOU, Ji-liu, et al. Geometric feature based facial expression recognition using multiclass support vector machines. In: 2009 IEEE International Conference on Granular Computing. IEEE, 2009. p. 318-321.
- [30] SHBIB, Reda et ZHOU, Shikun. Facial expression analysis using active shape model. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2015, vol. 8, no 1, p. 9-22.
- [31] DANG-NGUYEN, Duc-Tien, BOATO, Giulia, et DE NATALE, Francesco GB. Identify computer-generated characters by analysing facial expressions variation. In: 2012 IEEE International Workshop on Information Forensics and Security (WIFS). IEEE, 2012. p. 252-257.
- [32] EDWARDS, Gareth J., TAYLOR, Christopher J., et COOTES, Timothy F. Interpreting face images using active appearance models. In: *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 1998. p. 300-305.
- [33] MARTINS, Pedro, SAMPAIO, Joana, et BATISTA, Jorge. Facial Expression Recognition Using Active Appearance Models. In: *VISAPP (2)*. 2008. p. 123-129.
- [34] https://en.wikipedia.org/wiki/Active_appearance_model.
- [35] HUANG, Zhong et REN, Fuji. Facial expression recognition based on active appearance model & scale-invariant feature transform. In: *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration*. IEEE, 2013. p. 94-99.
- [36] MANGAT, Amapreet Kaur et RANADE, Sukhjeet Kaur. Facial Expression Recognition Based on Local Binary Patterns. *International Journal of Advanced Research in Computer Science*, 2016, vol. 7, no 4.
- [37] NGUYEN, Huu-Tuan, VU, Ngoc-Son, et CAPLIER, Alice. How far we can improve micro features based face recognition systems? In: 2012 third International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE, 2012. p. 350-353.

- [38] HUANG, Di, SHAN, Caifeng, ARDABILIAN, Mohsen, et al. Local binary patterns and its application to facial image analysis: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2011, vol. 41, no 6, p. 765-781.
- [39] HADID, Abdenour, ZHAO, Guoying, AHONEN, Timo, et al. Face analysis using local binary patterns. In: *Handbook of Texture Analysis*. 2008. p. 347-373.
- [40] NAZIR, Muhammad, JAN, Zahoor, et SAJJAD, Muhammad. Facial expression recognition using histogram of oriented gradients based transformed features. *Cluster Computing*, 2018, vol. 21, no 1, p. 539-548.
- [41] https://fr.wikipedia.org/wiki/Histogramme_de_gradient_orient%C3%A9.
- [42] MIGNIOT, Cyrille. Segmentation de personnes dans les images et les vidéos. Thèse de doctorat. Grenoble, 2012.
- [43] DHALL, Abhinav, ASTHANA, Akshay, GOECKE, Roland, et al. Emotion recognition using PHOG and LPQ features. In : *Face and Gesture 2011*. IEEE, 2011. p. 878-883.
- [44] KHAN, Rizwan, MEYER, Alexandre, KONIK, Hubert, et al. Une méthode de reconnaissance des expressions du visage basée sur la perception. In : *RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle)*. 2012. p. 978-2-9539515-2-3.
- [45] DAS, Amlan Jyoti et SAIKIA, Navajit. Pedestrian detection using dense LDB descriptor combined with HOG. In: *2016 International Conference on Information Technology (InCITe)-The Next Generation IT Summit on the Theme-Internet of Things: Connect your Worlds*. IEEE, 2016. p. 299-304.
- [46] NABILA, MERAMRIA Reconnaissance de visages par Analyse Discriminante Linéaire (LDA).Thèse de doctorat. Université Badji Mokhtar Annaba, 2016.
- [47] https://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support#R%C3%A9sum%C3%A9_intuitif.
- [48] https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Viola_et_Jones.
- [49] LIEW, Chun Fui et YAIRI, Takehisa. Facial expression recognition and analysis: a comparison study of feature descriptors. *IPSN transactions on computer vision and applications*, 2015, vol. 7, p. 104-120.
- [50] <https://www.lebigdata.fr/confusion-matrix-definition>

- [51] https://blogs.msdn.microsoft.com/big_data_france/2014/06/17/evaluer-un-modle-en-apprentissage-automatique/
- [52] EYRAUD, Rémi. Classification, Apprentissage, Décision.

Annexe

1. Matrice de confusion

Une Confusion Matrix est un résumé des résultats de prédictions sur un problème de classification. Les prédictions correctes et incorrectes sont mises en lumière et réparties par classe. Les résultats sont ainsi comparés avec les valeurs réelles. Cette matrice permet de comprendre de quelle façon le modèle de classification est confus lorsqu'il effectue des prédictions. Ceci permet non seulement de savoir quelles sont les erreurs commises, mais surtout le type d'erreurs commises. Les utilisateurs peuvent les analyser pour déterminer quels résultats indiquent comment les erreurs sont commises [50].

•Calcul de matrice de confusion

Pour calculer une matrice de confusion, il est nécessaire de disposer d'un ensemble de données de test ou d'un ensemble de données de validation avec les valeurs de résultat attendues. On fait ensuite une prédiction pour chaque ligne du 'données de test'.

A partir des résultats escomptés et des prédictions, la matrice indique le nombre de prédictions correctes pour chaque classe et le nombre de prédictions incorrectes pour chaque classe organisées en fonction de la classe prédite. Chaque ligne du tableau correspond à une classe prédite, et chaque colonne correspond à une classe réelle [50].

Pour mesurer les performances de ce classifieur, il est d'usage de distinguer 4 types d'éléments classés pour la classe voulue [51] :

Vrai positif VP : les cas où la prédiction est positive, et où la valeur réelle est effectivement positive.

Vrai négatif VN : les cas où la prédiction est négative, et où la valeur réelle est effectivement négative.

Faux positif FP : les cas où la prédiction est positive, mais où la valeur réelle est négative.

Faux négatif FN : les cas où la prédiction est négative, mais où la valeur réelle est positive

		Classe prédite	
		Classe +	Classe -
Classe réelle	Classe +	Vp	Fp
	Classe -	Fn	Vn

Tableau A.1 : Matrice de confusion.

En particulier, si la matrice de confusion est diagonale, le classifieur est parfait.

2. Score

Le score peut être exprimé comme suite [52] :

$$\bullet \text{Précision} = \frac{Vp}{Vp+Fp} \quad (\text{A.1})$$

$$\bullet \text{Rappel} = \frac{Vp}{Vp+Fn} \quad (\text{A.2})$$

$$\bullet \text{Score} = 2 * \frac{\text{précision} * \text{rappel}}{\text{précision} + \text{rappel}} \quad (\text{A.3})$$