

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université de Mohamed El-Bachir El-Ibrahimi - Bordj Bou Arreridj

Faculté des Sciences et de la technologie

Département d'électronique

Mémoire

Présenté pour obtenir

LE DIPLOME DE MASTER

FILIERE : Télécommunication

Spécialité : Systèmes de télécommunications

Par

➤ ZERROUGUI IBTISSEM

➤ HAMITOUCHE IMENE

Intitulé

***Réalisation d'un Système de Reconnaissance des Emotions Basé
sur des Images Faciales***

Devant le Jury composé de :

<i>Nom & Prénom</i>	<i>Grade</i>	<i>Qualité</i>	<i>Etablissement</i>
<i>M. BOUSSAHOUL Abdelkrim</i>	<i>MAA</i>	<i>Président</i>	<i>Univ-BBA</i>
<i>M. ROUABAH Khaled</i>	<i>Professeur</i>	<i>Encadreur</i>	<i>Univ-BBA</i>
<i>Mme. BOUKHOBZA Fatima Zohra</i>	<i>Doctorante</i>	<i>Co-Encadreur</i>	<i>Univ-Boumerdes</i>
<i>M. FLISSI Mustapha</i>	<i>MCA</i>	<i>Examineur</i>	<i>Univ-BBA</i>

Année Universitaire 2021/2022

Résumé : L'expression faciale est le moyen le plus couramment utilisé par les humains pour transmettre les états émotionnels. Ainsi, la reconnaissance automatique des émotions à partir des expressions faciales a gagné récemment une attention accrue. Un système capable de réaliser cette tâche peut être divisé en trois modules, à savoir la détection du visage, l'extraction des caractéristiques et la classification. Dans ce travail, nous nous sommes intéressés aux deux derniers modules. En effet, pour l'extraction des caractéristiques, nous avons appliqué deux approches, la première basée sur l'extraction de l'information de texture (LBP) et l'autre basée sur l'extraction de l'information de forme (HOG). Pour l'opération de classification, nous avons examiné deux différentes méthodes d'apprentissage automatique, notamment les machines à vecteurs de support (SVM) et K-plus proche voisin (KNN). Plusieurs expériences ont été menées afin de vérifier l'efficacité de chacune des approches. Les résultats expérimentaux, sur deux bases de données à savoir 'Japanese Female Facial Expression (JAFFE)' et 'Multimedia Understanding Group (MUG)', prouvent que les meilleures performances de reconnaissance, en termes de taux de reconnaissance et de temps de réponse, sont obtenues en utilisant les caractéristiques HOG avec un classifieur SVM.

Mot clé : SVM (Machine à vecteurs de support), KNN (K plus proches voisins), HOG (Histogramme de gradient orienté), LBP (Motif Binaire Local), JAFFE (Japanese Female Facial Expression), MUG (MultiMedia Understanding Group).

Abstract: Facial expression is the most common way humans used to convey emotional states. Thus, automatic emotion recognition from facial expressions has recently gained increased attention. A system capable of performing this task can be divided into three modules, namely face detection, feature extraction and classification. In this work, we are interested in the last two modules. In fact, for feature extraction, we applied two approaches. The first one is based on texture information extraction (LBP) and the second one is based on shape information extraction (HOG). For the classification task, we applied two different machine learning methods which are support vector machines (SVM) and K-nearest neighbor (KNN). Several experiments were conducted to confirm the effectiveness of each approach. The experimental results, based on 'Japanese Female Facial Expression (JAFFE)' and 'Multimedia Understanding Group (MUG)' databases, have shown that the best recognition performances, in terms of recognition rate and response time, are obtained by using HOG approach together with SVM classifier.

Key word: SVM (Support Vector Machine), KNN (K Nearest Neighbor), HOG (Oriented Gradient Histogram), LBP (Local Binary Pattern), JAFFE (Japanese Female Facial Expression), MUG (Multimedia Understanding Group).

ملخص: تعبيرات الوجه هي الطريقة الأكثر شيوعًا التي يستخدمها البشر للتعبير عن الحالات العاطفية. استفاد التعرف التلقائي على المشاعر من تعابير الوجه اهتمامًا متزايدًا مؤخرًا. يمكن تقسيم النظام القادر على أداء هذه المهمة إلى ثلاث وحدات، وهي اكتشاف الوجه واستخراج الميزات والتصنيف. في هذا العمل، نحن مهتمون بالوحدتين الأخيرتين. في الواقع، لاستخراج الميزات، طبقنا طريقتين. الأولى تعتمد على LBP والثانية تعتمد على HOG بالنسبة لمهمة التصنيف، طبقنا طريقتين مختلفتين للتعلم الآلي وهما SVM و KNN تم إجراء العديد من التجارب لتأكيد فعالية كل نهج. أظهرت النتائج التجريبية، المستندة إلى قواعد بيانات JAFFE و MUG ، أنه يتم الحصول على أفضل أداء التعرف ، من حيث معدل التعرف ووقت الاستجابة ، باستخدام خصائص نهج HOG مع مصنف SVM.

كلمات مفتاحية: SVM ; KNN ; HOG ; LBP ; JAFFE ; MUG



Remerciement

*Premièrement, nous remercions dieu, le tout puissant, de nous avoir
donné la santé et la volonté pour faire cet humble travail.*

*Nous tenons à exprimer nos vifs remerciements à nos encadreurs
Monsieur ROUABAH Khaled et Madame BOUKHOBZA Fatima
Zohra pour l'aide et pour les conseils précieux qu'ils nous ont donné
durant notre étude et durant la Réalisation de ce travail.*

*Nous remercions également nos enseignants, nos parents, nos frères
et sœurs et nos chers amis, pour leurs aides et*

Leurs sacrifices.



Dédicace

Tous les mots ne sauraient exprimer la gratitude, l'amour, le respect, la reconnaissance...

*Aussi, c'est tout simplement que Je dédie ce travail de fin d'étude à **mon père et ma mère**, mes piliers, mes exemples, mes premiers supporteurs et ma plus grande force.*

Merci pour votre présence, votre soutien, votre aide financière, et surtout votre amour que dieu vous protège. Je vous aime.

*A mes sœurs, Mon modèle dans cette vie, ce sont les plus belles sœurs du monde **Ahlem , Sarah, Assia**, Et surtout leurs enfants **Ishak, Takoua, Rayen, Taouba** que ce travail soit l'expression de mon estime pour vous et que Dieu vous protège, vous accorde santé, succès et plein de bonheur dans votre vie.*

*Mes âmes sœurs et mes meilleures amies **M.B,Imen, Haizia, Mouna, Lama** qui depuis des années m'encouragent, me comprennent et ont toujours été à mes côtés.*

*Toutes mes amies sans exception et surtout **Zayneb et Amel***

Tous les membres de ma promotion.

Tous ceux qui me sentent chers et que j'ai omis de citer.

Ibtissem



Dédicace

*Je dédie ce modeste travail en signe de respect, reconnaissance et de remerciement :
A mes chers parents : qui m'ont encouragé à aller de l'avant et qui m'ont donné tout
leur amour pour prendre mes études. Auxquels je dois ce que je suis. Que dieu les
protège.*

*Mes chers frères et sœurs pour leur dévouement, leur compréhension et leur grande
tendresse, qui en plus de m'avoir encouragé tout le long de mes études, m'ont consacré
beaucoup de temps et disponibilité, et qui par leur soutien, leurs conseils et leur
amour, M'ont permis d'arriver jusqu'à ici car ils ont toujours cru en moi.
En signe de l'affection et du grand amour que je vous porte, les mots sont
insuffisants pour exprimer ma profonde estime.*

*Je vous dédie ce travail en témoignage de ma profonde affection et de mon
attachement indéfectible.*

*A mes épouses des frères, mes nièces, mes neveux ; Pour leur amour, soutien et leur
présence dans ma vie.*

A mes copines Ibtissem, Lamia, Haizia et Mouna, tous mes amis

De la promo et tous ceux qui me sont chers.

Imene

Table des matières

Table des matières	
Liste des tableaux	
Liste des figures	
Introduction générale	1
Chapitre 1 Etat de l'art sur la reconnaissance des émotions	3
1.1 Définition des émotions.....	3
1.2 Types d'émotion	3
1.2.1 Émotions primaires	3
1.2.2 Émotions secondaires	4
1.2.3 Émotions sociales.....	5
1.3 Les théories des émotions.....	5
1.3.1 Théorie de l'universalité.....	5
1.3.2 Théorie physiologique.....	5
1.4 Représentation des émotions	5
1.4.1 L'état émotionnel à travers l'expression visuelle	6
1.4.2 L'état émotionnel à travers l'expression auditive	6
1.4.3 L'état émotionnel à travers la représentation physiologique.....	6
1.5 La reconnaissance des émotions par traitement des expressions faciales	6
1.5.1 Expression faciale	7
1.5.2 Approches de reconnaissances automatiques des expressions faciales	7
1.5.3 Représentation des expressions faciales émotionnelles	8
1.6 Conclusion	12
Chapitre 2 Système de reconnaissance des expressions faciale	14

Table des matières

2.1	Introduction.....	14
2.2	Modules du système d'analyse des expressions faciales.....	14
2.2.1	Acquisition d'images et détection du visage	15
2.2.2	Extraction de caractéristiques	17
2.2.3	Classification	31
2.3	Conclusion	36
Chapitre 3 Résultats expérimentaux et discussion		38
3.1	Introduction.....	38
3.2	Base de données et protocole d'évaluation.....	38
3.2.1	Description des bases de données	38
3.2.2	Protocole d'évaluation utilisé	39
3.3	Expérimentations et Résultats.....	39
3.3.1	Environnement du travail	39
3.3.2	Méthodologie	40
3.3.3	Résultats	41
3.3.3.1	Evolution sur la base de données JAFFE.....	41
3.3.3.2	Evolution sur la base de données HOG.....	46
3.3.3.3	Evolution sur la base de données MUG.....	49
3.4	Conclusion.....	51
Conclusion générale.....		52
Bibliographie		

Liste des tableaux

Tableau 1.1 les six émotions de bases	4
Tableau 1.2 les prototypes des six émotions de bases.....	9
Tableau 3.1 Résultats d'application du LBP et KNN/ SVM sur la base de données JAFFE (Taille de cellule= 8x8).....	41
Tableau 3.2 Résultats d'application du LBP et KNN/ SVM sur la base de données JAFFE (Taille de cellule= 16x 16).....	42
Tableau 3.3 Résultats d'application du LBP et KNN/ SVM sur la base de données JAFFE (Taille de cellule= 24 x 24).....	42
Tableau 3.4 Résultats d'application du HOG et KNN/ SVM sur la base de données JAFFE (Taille de cellule= 8 x 8).....	44
Tableau 3.5 Résultats d'application du HOG et KNN/ SVM sur la base de données JAFFE (Taille de cellule= 16 x 16).....	45
Tableau 3.6 Résultats d'application du HOG et KNN/ SVM sur la base de données JAFFE (Taille de cellule= 24 x 24).....	45
Tableau 3.7 Résultats d'application du LBP et KNN/ SVM sur la base de données MUG (Taille de cellule= 8 x 8).....	47
Tableau 3.8 Résultats d'application du LBP et KNN/ SVM sur la base de données MUG (Taille de cellule= 16x 16).....	48
Tableau 3.9 Résultats d'application du LBP et KNN/ SVM sur la base de données MUG (Taille de cellule= 24 x 24).....	48
Tableau 3.10 Résultats d'application du HOG et KNN/ SVM sur la base de données MUG (Taille de cellule= 8 x 8).....	49
Tableau 3.11 Résultats d'application du HOG et KNN/ SVM sur la base de données MUG (Taille de cellule= 16 x 16).....	49

Liste des tableaux

Tableau 3.12 Résultats d'application du HOG et KNN/ SVM sur la base de données MUG (Taille de cellule= 24x 24)	50
---	----

Liste des figures

Figure 1.1 Différents moyens pour communiquer les émotions	6
Figure 1.2 Expression prototypique de la tristesse.....	9
Figure 1.3 Modèle du visage par le standard MPEG-4 :	11
Figure 2.1 Modules principaux d'un système de reconnaissance automatique des expressions faciales. [26].....	15
Figure 2.2 Acquisition et détection du visage [6]	15
Figure 2.3 L'opérateur LBP [53]	22
Figure 2.4 Extraction des caractéristiques de l'histogramme LBP. [28]	23
Figure 2.5 Génération d'un motif binaire incohérent dans le processus d'encodage LBP [67]	24
Figure 2.6 Illustration de l'opérateur CLBP de base,Ici, le modèle binaire est 1011111110101010 [67].....	24
Figure 2.7 Génération des modèles de sous-CLBP a partir du code CLBP d'origine [69]....	25
Figure 2.8 L'opérateur CLBP et la génération de deux sous-CLBP [55]	25
Figure 2.9 exemple d'application de la méthode MBP	25
Figure 2.10 Organigramme de l'ensemble des étapes nécessaires à la génération du vecteur des caractéristiques par la méthode LPQ [56].....	27
Figure 2.11 gradients horizontaux et verticaux.....	29
Figure 2.12 Exemple de classification SVM	31
Figure 2.13 Séparation linéaire et non linéaire	33
Figure 2.14 Le passage par une description des données peut permettre une séparation linéaire des exemples	33
Figure 2.15 Exemple de classification KNN (K=3 et K=5)	35

Liste d'abréviation

Liste d'abréviation

SNA	ystème nerveux autonome.
FACS	Facial Action Coding System
AU	Action Units
MPEG	Moving Picture Experts Group.
FAPs	Facial Animation Parameters).
FAPU	Face Animation Parameter Units).
ASM	Active Shape Model.
AAM	Active Appearance Model.
LBP	Motif Binaire Local.
LPQ	Local phase Quantization.
HOG	L'histogramme de gradient orienté.
SIFT	scale-invariant Feature transforme 'La transformation de caractéristique invariante à l'échelle '.
CLBP	Modèle binaire local composé.
MBP	Modèle binaire Médian.
LTP	Motif ternaire local.
LPQ	Local phase Quantization.
SVM	Machines à vecteurs de support.
K-NN	K-plus proche voisin.
RF	Random Forest.
LDA	Analyse Discriminante Linéaire.
JAFFE	Japanese Female Facial Expression
MUG	Multimedia Understanding Group

Introduction générale

Au cours des dernières années, une énorme attention a été accordée à la reconnaissance automatique des émotions. Un système qui exécute cette opération trouvera de nombreux domaines d'applications, Nous citons à titre d'exemples : le marketing, la sécurité, la médecine et l'interaction Homme-Machine. Dans ce mémoire, nous traitons le cas de la reconnaissance des émotions par traitement d'images, en considérant le visage qui est le moyen le plus puissant pour communiquer les émotions à travers des expressions faciales.

Les systèmes de reconnaissance automatique des expressions faciales consistent généralement en trois phases principales : détection de visage, extraction de caractéristiques et classification. La plupart des méthodes d'analyse des expressions faciales classent les expressions faciales en sept catégories de base : la joie, la tristesse, le dégoût, la peur, la surprise, la colère et la neutralité. Une description bien définie de chaque émotion est indispensable pour une bonne reconnaissance. Ainsi, l'extraction des caractéristiques faciales caractérisant les différentes émotions est la phase la plus importante pour assurer une bonne reconnaissance. De même, un bon choix du classifieur est un critère primordial pour l'amélioration des performances.

Dans ce travail, nous proposons la mise en œuvre d'un système de reconnaissance des expressions faciales par considération de différentes approches d'extraction de caractéristiques faciales et de classification. En effet, pour l'extraction de caractéristiques, nous utiliserons deux types de descripteurs Le premier, qui consiste en le LBP (Motif Binaire Local), pour l'extraction de l'information de texture et le deuxième, qui est le HOG (Histogramme de gradient orienté), pour l'extraction de l'information de forme. Pour l'opération de la classification, nous utiliserons le SVM

Introduction générale

(les machines à vecteurs de support) et le KNN (K-plus proche voisin). Deux bases de données notamment la japonaise JAFFE (Japanese Female Facial Expression) et l'italienne MUG (Multimedia Understanding Group (MUG) seront utilisées pour tester les performances, en termes de taux de reconnaissance et de temps de réponse, des différentes approches.

Organisation du mémoire

Le chapitre 1 présente quelques notions de base sur les émotions plus particulièrement leurs définitions, leurs théories et les différents moyens les plus couramment utilisés par les humains pour les communiquer. Il donne ensuite un aperçu général sur la reconnaissance des émotions à travers les expressions faciales

Le chapitre 2 décrit les différentes étapes de traitement qui composent un système de reconnaissance automatique des expressions faciales, tout en discutant les différentes techniques existantes dans la littérature pour chaque étape.

Le chapitre 3 présente, dans une première partie, une description détaillée sur l'environnement du travail utilisé pour la mise en œuvre du système de reconnaissance des expressions faciales. Puis dans une deuxième partie, il montre les différentes expérimentations implémentées en considérant les différentes approches d'extraction de caractéristique et de classifications. Le chapitre se termine par discussion des résultats obtenus.

Le mémoire est clôturé par une conclusion d'ordre général.

Chapitre I

Etat de l'art

Sur la Reconnaissance

des Emotions

Chapitre 1
Etat de l'art sur la
reconnaissance des
émotions

1.1 Définition des émotions

D'après le professeur Timothy Judge et le docteur Stephen Robbins [1], l'émotion est un sentiment intense provoqué par un événement ou une interaction avec quelqu'un, puis diriger vers l'événement ou la personne responsable de l'apparition de l'émotion. Une émotion est habituellement accompagnée d'expressions du visage et de mouvements du corps. Elle est aussi d'une durée très courte.

1.2 Types d'émotion

Les émotions sont des séquences courantes et vives qui interviennent tout au long de nos journées et en donnent le ton. Il existe trois types d'émotions : émotions primaires ou dites de base, émotions secondaires et émotions sociales.

1.2.1 Émotions primaires

Sont à la base de nos réactions. Elles sont déclenchées par des événements particuliers ou bien elles se manifestent dans des circonstances précises en provoquant des comportements spécifiques [2].

Tableau 1.1 les six émotions de bases

Émotion	Déclencheurs et circonstances d'apparition	Comportement
Joie	Désir Réussite Bien-être Accomplissement	Approche
Tristesse	Perte Deuil	Repli sur soi
Colère	Obstacle Injustice Dommage Atteindre à son intégrité physique ou psychique Limites de la personne Atteinte au système de valeurs	Attaque
Peur	Menace Danger Inconnu	Fuite Sidération Évitement Parfois attaque
Dégoût	Substance ou personne nuisible Aversion physique ou psychique Contre quelqu'un Rejet	
Surprise	Danger immédiat Inconnu Imprévu	Retrait Sursaut

1.2.2 Émotions secondaires

Sont celles qui sont engendrées à l'évocation de souvenirs et arrivent à maturation à l'âge adulte [3].

1.2.3 Émotions sociales

Sont inhérentes à la relation aux autres, comme la culpabilité, la honte, la jalousie, la timidité, l'humiliation, etc.

1.3 Les théories des émotions

Dans cette section, nous aborderons les théories les plus connues des émotions

1.3.1 Théorie de l'universalité

Charles Darwin a été l'un des premiers à s'occuper de la mondialisation des expressions faciales et à les développer, car il a été l'un des participants et a proposé une théorie pour le développement et la formation de différentes expressions faciales. Son argument principal donné est que les expressions des enfants et des nouveaux nés existent aussi chez les adultes. D'après lui, l'expression des émotions est un processus nécessaire à la survie. Ainsi, les expressions non-verbales sont aussi importantes que les interactions verbales dans le processus de communication humaine. Selon Ekman [4], les émotions de base ont des caractéristiques uniques et leurs réponses sont préprogrammées. [5]

1.3.2 Théorie physiologique

La théorie de William James [7] stipule que les réponses du système nerveux périphérique soient à la base de l'expérience émotionnelle. A l'opposé, Cannon [8], pense que le système nerveux central est à l'origine des expériences émotionnelles qui déclenchent des changements physiques et physiologiques.

1.4 Représentation des émotions

L'état émotionnel d'un humain peut être obtenu à partir de différentes techniques disponibles par le biais d'une expression ou d'une présentation visuelle, auditive et physiologique.

1.4.1 L'état émotionnel à travers l'expression visuelle

Il est évalué en fonction de la modulation des expressions faciales, des gestes, des postures et plus généralement le langage corporel. [9]

1.4.2 L'état émotionnel à travers l'expression auditive

Il peut être estimé comme une modulation du signal vocal. [11]

1.4.3 L'état émotionnel à travers la représentation physiologique

Il est estimé par la modulation de l'activité du système nerveux autonome (SNA). [10]

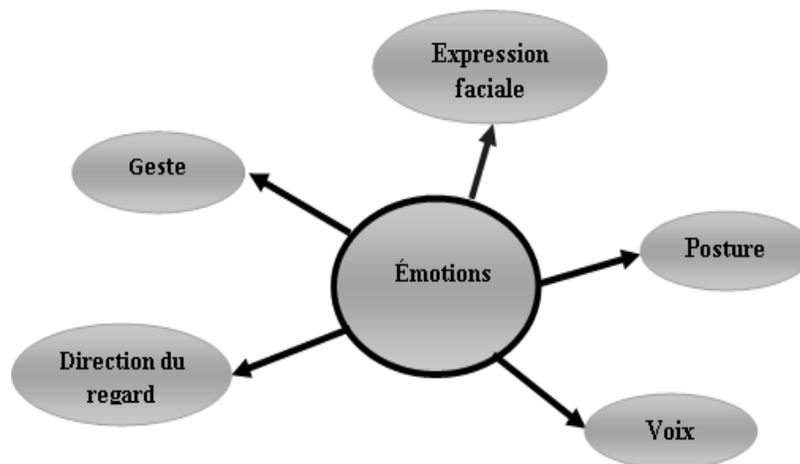


Figure 1.1 Différents moyens pour communiquer les émotions

La majorité des recherches dans ce domaine indiquent que les expressions faciales jouent le rôle le plus important dans la communication des émotions. Dans [12], Mehrabian a mis en évidence le fait que 55 % du message émotionnel est communiqué par l'expression faciale alors que 7 % seulement par le canal linguistique et 38% par le paralangage.

1.5 La reconnaissance des émotions par traitement des expressions faciales

Nous présentons dans cette section la définition de l'expression faciale et leur approche de reconnaissance automatique et leur représentation

1.5.1 Expression faciale

L'expression faciale est un ensemble de signes du visage qui traduisent un sentiment de changement dans le visage (mimique faciale chargée de sens), perceptible visuellement, dû à l'activation (volontaire ou non) de l'un ou de plusieurs muscles composant le visage.[11]

Les expressions faciales sont très importantes pour pouvoir connaître l'état de la personne. Grâce à ces expressions, il est possible de faire plusieurs déductions et récupération de plusieurs informations comme :

- L'état affectif que ce soit les émotions (peur, colère, joie, surprise, tristesse, dégoût) ou bien certaines humeurs ;
- L'activité cognitive comme la concentration, l'ennui ou la perplexité ;
- Le tempérament et la personnalité ; [13]

1.5.2 Approches de reconnaissances automatiques des expressions faciales

Ces approches peuvent être divisées globalement en trois approches principales qui sont: l'approche basée modèles, l'approche basée règles et l'approche basée géométrie. [13]

1.5.2.1 Approche basée modèles

Cette approche consiste à voir le visage comme un tout. L'analyse consiste alors à mesurer la ressemblance du visage observé à un modèle (connu ou appris). Les méthodes de cette approche font généralement appel à des méthodes de mise en correspondance de modèles. L'intérêt de ces méthodes est qu'elles peuvent être appliquées de manière plus locale.

1.5.2.2 Approche Géométrique

Depuis peu de temps, les informaticiens utilisent l'analyse géométrique pour représenter l'information faciale car elle donne des mesures précises qui ne peuvent pas être délivrées par les approches basées modèles. Dans cette approche, Les

distances entre les points distincts entre deux images, l'une neutre et l'autre expressive, sont calculées afin de détecter les mouvements. [15]

1.5.2.3 Approche basée règles

Cette approche réalise une codification automatique en unités d'action en appliquant un multi-détecteur automatique de composantes du visage. Elle extrait des modèles de composantes depuis les contours des composantes. Ensuite, elle calcule la différence entre le modèle détecté et le modèle à l'état neutre. La classification finale, dans l'une des six catégories correspondantes aux six expressions universelles, est réalisée à l'aide de la description linguistique donnée par Ekman. [11]

1.5.3 Représentation des expressions faciales émotionnelles

La section suivante montre trois catégories d'expressions faciales émotionnelles.

1.5.3.1 Représentation par le système FACS

En 1978, Ekman et Friesen [16] présentent un système de codification manuelle des expressions du visage. Ce système est connu sous le nom de FACS (Facial Action Coding System).

Le FACS est une méthode permettant de décrire les mouvements faciaux observables. Il est utilisé par de nombreux psychologues et il s'est imposé comme un outil puissant de description des mimiques du visage.

Il s'agit d'associer un code à chaque activation musculaire du visage qui peut être distinguée visuellement. Ces éléments atomiques sont appelés AU (Action Units) dont on distingue 46 unités d'Actions qui décrivent les mouvements élémentaires des muscles. Le manuel du codeur FACS contient ainsi la description visuelle des changements du visage lors de l'occurrence de chaque AU ou chaque combinaison d'AUs. De plus, chaque AU peut être affichée avec une amplitude différente.

De plus, la combinaison de plusieurs mouvements faciaux ne peut pas être décrite facilement par la combinaison visuelle de chacun des mouvements isolés. En

effet, il se produit un phénomène de coarticulation, de la même manière que la prononciation d'un mot ne peut pas être réduite à la concaténation de la prononciation de chacun de ses phonèmes. [13]

La figure 2 présente la description de la tristesse par l'un des prototypes formés des actions unitaires [14]. Le tableau 2 présente les différents prototypes des 6 émotions de bases [13].

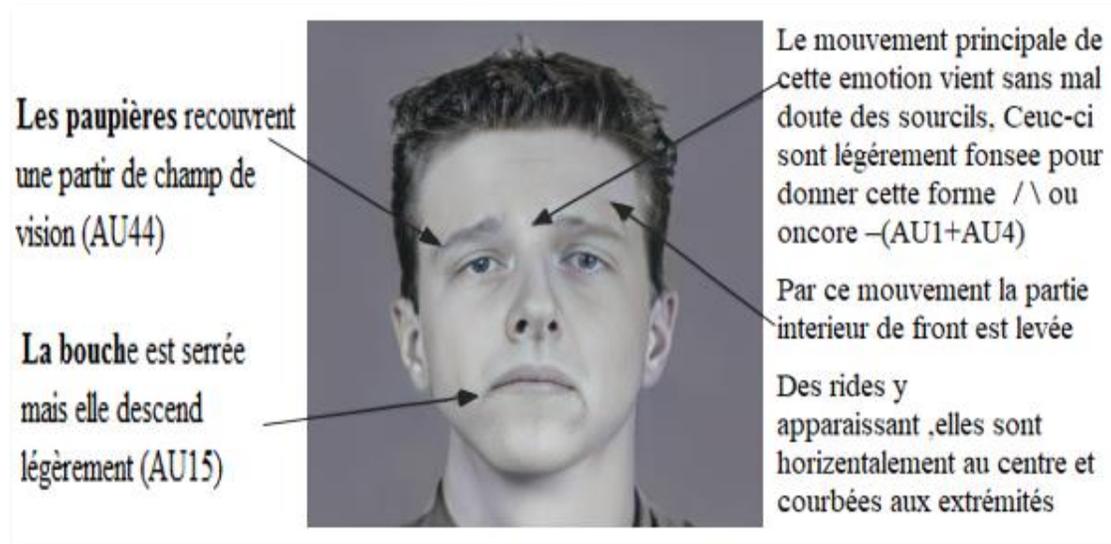


Figure 1.2 Expression prototypique de la tristesse

Tableau 1.2 les prototypes des six émotions de bases

Les émotions	Les Unités d'actions (AUs)
La joie	6+12
La tristesse	1+4+15
La surprise	1+2+5+26
La peur	1+2+4+5+20+26
La colère	4+5+7+23
Le dégoût	9+15+16

Bien que le système FACS soit un système de description bénéficiant d'une grande maturité (environ vingt années de développement), il souffre cependant de quelques inconvénients :

- **Complexité** : on estime qu'il faut 100 heures d'apprentissage pour en maîtriser les principaux concepts ;
- **Difficulté de manipulation par une machine** : FACS a d'abord été créé pour des psychologues, Certaines mesures restent floues et difficilement évaluables par une machine ;
- **Manque de précision** : les transitions entre deux états d'un muscle sont représentées de manière linéaire, ce qui représentées une approximation de la réalité. [9].

1.5.3.2 Représentation par le standard MPEG

Un deuxième codage, très utilisé pour la description des expressions faciales, est défini par les paramètres du standard de compression MPEG-4 [17]. Utilisé principalement pour l'animation et la synthèse du visage, leMPEG-4 définit un ensemble de paramètres FAPs (Facial Animation Parameters) nécessaire au codage de la déformation d'un visage à l'état neutre [18].

Les valeurs des paramètres d'animation FAPs sont proportionnelles à l'amplitude de la déformation faciale. Ces paramètres sont représentés par un total de 84 points caractéristiques FPs (Feature Points) définis sur un visage neutre (voir la figure 1.3 (a)). Le modèle ainsi défini doit être applicable sur des visages de différentes tailles. Pour ce faire, les paramètres d'animation faciale FAPs sont définis en fonction des unités de paramètres d'animation faciale FAPU (Face Animation Parameter Units). [20]

Les FAPU permettent de définir des mouvements élémentaires du visage de manière transposable. En effet, il est difficile de définir les mouvements élémentaires des muscles de manière absolue, mais on peut considérer l'intensité de leur déplacement, relative à certaines distances pertinentes, comme constante. C'est ce qui permet de donner des expressions humaines à des personnages non-humains [19]. Comme exemples des FAPU, on peut citer la largeur de la bouche, la distance de séparation entre la bouche et le nez, la distance de séparation entre les yeux et le nez,

etc. [13]. La figure 1.3 (b) illustre les rapports des distances entre les caractéristiques faciales clés d'un visage à l'état neutre. [20]

Comme le codage FACS, le codage par FAPs est très lié aux changements des muscles faciaux. Il définit également six expressions faciales émotionnelles qui coïncident avec les émotions de base définies par le système FACS.

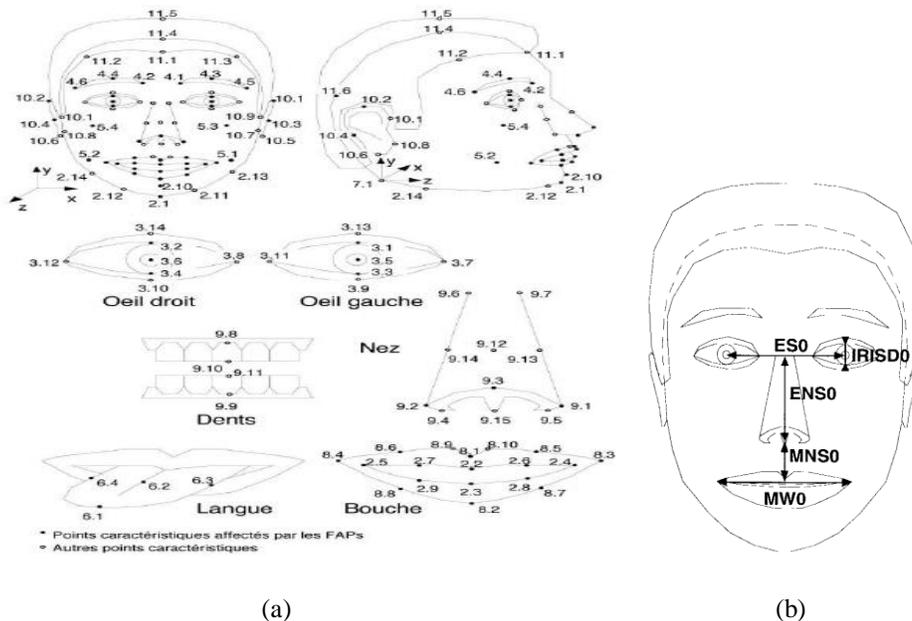


Figure 1.3 Modèle du visage par le standard MPEG :

(a) Représentation des points caractéristiques FAs.

(b) Modèle des paramètres unitaires FAPUs [20]

1.5.3.3 Représentation par catégorie

Dans ce mode de représentation, les descripteurs sont d'abord extraits dans des vecteurs caractéristiques. Par la suite, l'étape d'apprentissage se charge de construire un ou plusieurs modèles spécifiant le comportement des expressions émotionnelles. Lors des tests, le problème se réduit à une classification d'un vecteur de descripteurs dans des groupes d'émotions [20].

Les systèmes de reconnaissance des expressions, notamment les expressions émotionnelles, sont constituées de trois étapes principales [21] :

- a) **Acquisition du visage** : cette étape peut être réalisée soit par une détection du visage soit par une estimation de la pose de tête ;
- b) **Extraction des informations faciales liées à l'expression** : ces informations sont liées soit à l'apparence de l'expression, soit à la géométrie des déformations. Dans le cadre d'une séquence d'images, l'information liée à la dynamique de l'expression est également utile ;
- c) **Reconnaissance de l'expression** : cette étape est effectuée par une classification dans le cas d'une représentation discrète des émotions ou par une régression dans le cas d'une représentation continue des émotions.

1.6 Conclusion

Dans ce chapitre, nous avons d'abord présenté un aperçu sur les types et les théories des émotions. Nous avons ensuite exposé les différentes techniques pour les communiquer, avec un accent particulier sur la reconnaissance des émotions par traitement des expressions faciales. Dans le chapitre suivant, nous allons nous intéresser aux différentes étapes de conception d'un système d'analyse d'expression faciale avec description des méthodes existantes dans la littérature pour chaque étape.

Chapitre II

Systeme de Reconnaissance des Expressions Faciales

Chapitre 2***Système de reconnaissance
des expressions faciale*****2.1 Introduction**

Un système automatique d'analyse d'expressions faciales a comme caractère essentiel trois phases principales qui sont : la détection de visage à étudier, l'extraction de caractéristiques de visage à étudier et enfin la classification de l'expression faciale.

Dans ce chapitre, nous donnons un aperçu des différentes phases du système de reconnaissance automatique des expressions faciales ainsi que la description des différentes méthodes existantes pour chaque phase.

2.2 Modules du système d'analyse des expressions faciales

Comme l'illustre la figure 2.1, un système qui effectue la reconnaissance automatique des expressions faciales se compose généralement de trois modules principaux. Le premier module consiste à détecter et à enregistrer les régions du visage dans une image d'entrée ou une séquence d'images. Le deuxième module consiste à extraire et à afficher les changements faciaux causés par les expressions faciales. [23] Le dernier module détermine la similarité entre l'ensemble de caractéristiques extraites et l'ensemble de caractéristiques de référence. Des filtres supplémentaires ou des modules de prétraitement des données peuvent être utilisés entre ces modules

principaux pour améliorer les résultats de détection. Ils peuvent être également utilisés pour l'extraction des caractéristiques ou de classification [24]

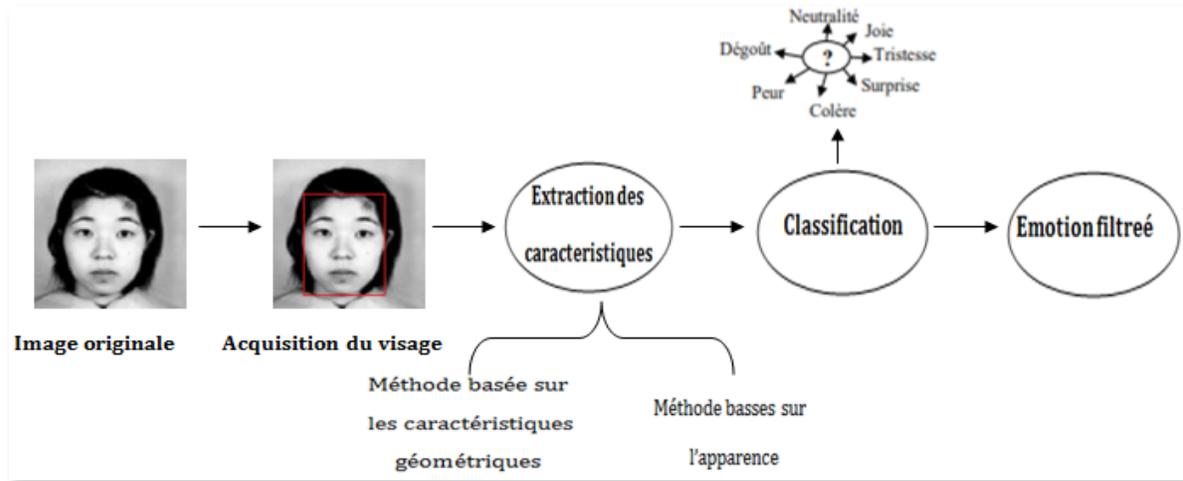


Figure 2.1 Modules principaux d'un système de reconnaissance automatique des expressions faciales. [26]

Dans ce qui suit, nous étudions en détail chacune des étapes du système de reconnaissance des expressions faciales.

2.2.1 Acquisition d'images et détection du visage

La Figure 2.2 montre la procédure d'acquisition et détection d'une image du visage japonaise.

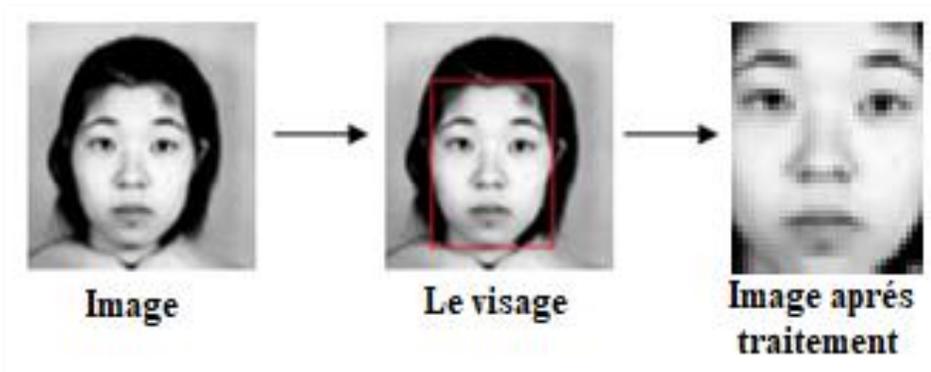


Figure 2.2 Acquisition et détection du visage [6]

2.2.1.1 Acquisition du visage

Les systèmes de capture sont généralement équipés d'une caméra ou d'un caméscope capable de capturer une image 2D du visage. Les problèmes de détection et d'enregistrement des visages incluent l'identification de la présence d'un visage dans une image et la détermination de la position et de l'échelle du visage [22]. Le système de localisation faciale doit faire face à certains des défis décrits ci-dessous.

- **Pose** : Traits du visage tels que les yeux, le nez et la bouche peuvent être partiellement invisibles ou déformés en raison des poses relatives du visage ou de la caméra ;
- **Obstruction** : Traits du visage peuvent être masqués par une barbe, une moustache et des lunettes. De même, le maquillage peut amener le visage à afficher des zones artificielles ou à masquer les limites normales du visage ;
- **Expressions faciales** : Traits du visage montrent de grands changements de forme avec différentes expressions faciales. Certaines fonctions peuvent être masquées, ou d'autres peuvent n'être visibles que dans d'autres expressions ;
- **Conditions de capture d'image** : Les changements d'éclairage et les caractéristiques de la caméra peuvent avoir un effet significatif sur la chromaticité de la zone du visage. Certaines fonctionnalités peuvent obscurcir, projeter des ombres, illuminer le visage et perdre des informations de couleur pertinentes.

2.2.1.2 Détection du visage

La première étape d'un système d'analyse de l'expression faciale consiste à trouver des zones du visage et leurs limites. [22].

Plusieurs méthodes de détection de visages ont été proposées. On peut les classer selon quatre catégories décrites ci-dessous [24] :

- a. Méthodes basées sur la connaissance ;
- b. Approches invariantes ;

- c. Méthodes d'appariement de modèles ;
- d. Méthodes basées sur l'apparence.

La méthode basée sur la connaissance reconnaît un visage à l'aide de règles prédéfinies basées sur la connaissance humaine

L'approche basée sur les caractéristiques immuables vise à trouver des caractéristiques structurelles du visage qui sont robustes aux poses et aux conditions d'éclairage.

Les méthodes basées sur des modèles utilisent des modèles de visage préenregistrés. Calculer les valeurs de corrélation entre le modèle et l'image d'entrée, et utiliser ces valeurs de corrélation pour déterminer la présence d'un visage,

Les méthodes basées sur l'apparence utilisent l'apprentissage automatique et des techniques statistiques pour modéliser les visages humains.

Une approche basée sur un modèle rigide, incluant des changements d'augmentation. Les principaux algorithmes de cette catégorie incluent l'algorithme de détection de visage Viola-Jones (VJ) et ses variantes.

2.2.2 Extraction de caractéristiques

Une fois la détection et l'enregistrement du visage terminés, l'étape suivante consiste à extraire et à rendre les changements faciaux causés par les expressions faciales. L'obtention des caractéristiques faciales est importante pour une reconnaissance réussie de l'expression faciale. Les expressions faciales sont principalement définies par la contraction des muscles faciaux et provoquent des changements dans l'apparence et la forme du visage [39]. Pour cette raison, les méthodes d'extraction de caractéristiques pour l'analyse d'expression peuvent être divisées en deux approches : les méthodes basées sur les caractéristiques géométriques et les méthodes basées sur l'apparence.

2.2.2.1 Méthode basée sur les caractéristiques géométriques

Nous présentons la Descripteurs géométriques et Reconnaissance des expressions faciales basée sur des caractéristiques géométriques

2.2.2.1.1 Définition

Des composants faciaux ou des traits faciaux sont extraits pour former une représentation vectorielle des caractéristiques géométriques faciales. Les caractéristiques géométriques représentent la forme et la position des composants du visage (bouche, yeux, sourcils, nez, etc.) [22].

2.2.2.1.2 Descripteurs géométriques

Les expressions faciales créent des distorsions géométriques sur le visage. Ces anomalies se manifestent par des changements dans la forme des caractéristiques permanentes telles que la bouche, le nez et les sourcils, et l'apparition de caractéristiques temporaires telles que les rides d'expression. L'apparence de la déformation peut également être décrite par des points caractéristiques ou des motifs géométriques qui révèlent les mouvements des muscles faciaux. [14]

Pour pouvoir déterminer les points caractéristiques du visage, il existe plusieurs algorithmes tel que :

- L'algorithme ASM (Active Shape Model) [24][25] qui est un modèle statistique permettant de placer les points caractéristiques sur un objet.
- L'algorithme AAM (Active Appearance Model) [26] qui est une version similaire à l'ASM mais qui utilise d'autres informations dans l'image
- L'algorithme EBG (Elastic Bunch Graph) [29] qui est un algorithme qui se base sur un graphe pour la reconnaissance des objets.

Ces points caractéristiques sont liés aux "Actions Units" définis par Paul Ekman. Autrement dit, ces points permettent de définir l'état de chaque muscle du visage et par la suite il est possible de suivre leurs mouvements et de déduire l'émotion exprimée à partir des expressions faciales.

2.2.2.1.2.1 Reconnaissance des expressions faciales basée sur des caractéristiques géométriques

Les caractéristiques géométriques représentent la forme d'une partie du visage et son emplacement. Il est représenté par les composants du visage ou les traits du visage qui constituent un vecteur de traits qui représentent la forme du visage. Les emplacements de ces références doivent être localisés avec précision pour que les techniques d'ingénierie soient efficaces. Les méthodes utilisant uniquement des traits géométriques sont principalement utilisées pour définir les traits du visage en tant qu'informations visuelles [34, 39, 44], les déplacements géométriques des traits du visage [29] ou les formes des composants du visage [31, 33]. Zhang etc. [50] ont utilisé 34 points de référence pour représenter l'image du visage. Les points de référence ont été sélectionnés manuellement et les coordonnées de ces points ont été utilisées comme caractéristiques pour générer un vecteur de caractéristiques à 68 dimensions. Tie et al. [42] ont proposé un modèle de composants faciaux multi-états pour détecter et suivre les changements des composants faciaux dans les images faciales. Ce modèle représente le mouvement facial en mesurant les transitions d'état des composants faciaux correspondants. Dans une série d'images, les mouvements du visage peuvent être modélisés en mesurant le déplacement géométrique des points caractéristiques du visage entre l'image actuelle et la première image. Le travail dans [39,40] décrit les expressions faciales en utilisant un ensemble de 20 points de caractéristiques faciales détectés à l'aide du détecteur de point de vue proposé dans [47]. De même, Val star et Pantic [41] ont utilisé les mêmes positions de points pour calculer des caractéristiques supplémentaires en fonction de la distance et de l'angle entre eux, et de la vitesse de déplacement des points dans l'image.

Ces caractéristiques ont été utilisées pour expliquer l'évolution de l'unité d'action (UA) au fil du temps. Dans [30], un modèle de forme, défini à partir de 58 points de vue, a été adopté et les catégories d'expression de base ont été analysées à partir de différents points de vue. Les positions des 68 sommets de l'ASM, qui font partie de l'AAM [40], ont été utilisées dans [31,35] pour tenir compte des changements

d'intensité de l'UA et des expressions faciales de la douleur. L'ASM a également été utilisée dans [43] pour appliquer des techniques d'ajustement ASM afin d'extraire des perspectives fiables. Les expressions faciales ont été évaluées à l'aide d'un décalage géométrique entre les coordonnées projetées du point caractéristique ASM et la forme ASM moyenne.

Dans [43,44], l'auteur a utilisé la SDM [48] pour suivre 49 points de vue pour faire correspondre la forme moyenne des points stables dans les coins des yeux et la zone du nez. Pour l'extraction des caractéristiques, ils ont calculé la différence entre les coordonnées du point de vue aligné et les coordonnées de la forme moyenne et la différence entre les coordonnées du point de vue aligné de l'image précédente et de l'image actuelle. Cette procédure a généré un total de 196 caractéristiques. Ensuite, ils ont divisé les points de vue en groupes selon trois zones différentes (œil gauche/sourcil gauche, œil droit/sourcil droit, bouche). Pour chacun de ces groupes, la distance euclidienne et l'angle entre les points sont calculés pour fournir 71 caractéristiques. Ils ont également calculé la distance euclidienne entre le point de vue stable médian et chaque point de vue aligné dans l'image vidéo. Au total, l'ensemble géométrique contient 316 éléments. L'extraction de caractéristiques géométriques, d'autre part, nécessite généralement une détection et un suivi précis et fiables des caractéristiques faciales. La reconnaissance et le suivi automatiques des traits du visage sont un problème ouvert dans de nombreuses situations de la vie réelle. Les caractéristiques géométriques ont l'avantage de la faible dimension et de la simplicité.

2.2.2.2 Méthodes basées sur l'apparence

A.Définition

Le principe de ces méthodes est d'utiliser toute la surface du visage ou des régions spécifiques du visage pour représenter les changements d'apparence. Parmi les approches basées sur l'apparence les plus communément utilisées pour l'analyse d'expression faciale, on donne :

- La méthode LBP (Motif Binaire Local) ;

- La méthode LPQ (Local phase Quantization) ;
- La méthode HOG (Histogram of Oriented Gradients) ;
- La transformation de caractéristique invariante à l'échelle (SIFT).

B. La méthode LBP

B.1 Définition

Les motifs binaires locaux (LBP) sont des caractéristiques utilisées en vision par ordinateur pour la classification des textures, la détection et le suivi des objets mobiles dans une séquence d'image [59].

Ce descripteur a été d'abord utilisé pour la première fois par Harwood en 1993 pour mesurer le contraste local d'une image mono spectrale, mais réellement popularisé en 1996 par Timo Ojala pour analyser les textures. Il a montré d'excellentes performances dans de nombreuses études comparatives, tant en termes de vitesse qu'en termes de discrimination des différentes textures [54].

B.2 Méthode LBP de base

Dans le concept LBP, le code binaire, décrivant la texture locale d'une région, est calculé en souillant un voisinage avec l'échelle de gris du pixel central. Afin de générer un motif binaire, tous les voisins prendront la valeur 1 si leur valeur est supérieure ou égale au pixel courant, et 0 sinon. Ensuite, les pixels de ce motif binaire sont alors multipliés par des poids et sommés afin d'obtenir un code LBP du pixel courant. Des pixels avec des intensités entre 0 et 255 seront finalement obtenus pour former une image LBP. [54]

LBP de base est défini par :

$$LBP_{P,R}(X_C, Y_C) = \sum_{p=1}^P s(g_p - g_C) 2^{p-1} \quad (1)$$

$$s(x) = \begin{cases} 0 & , x < 0 \\ 1 & , x \geq 0 \end{cases} \quad (2)$$

Où : g_C est le niveau de gris du pixel central de coordonnées (X_C, Y_C) .

($p = 0,1,\dots,7$) est le niveau de gris de chaque pixel voisin.

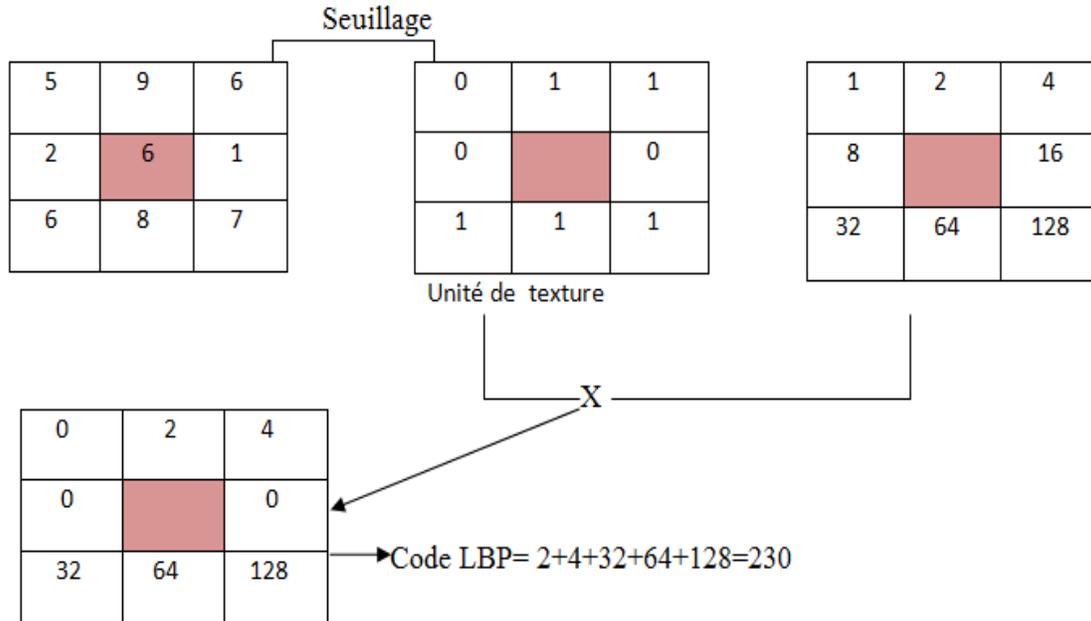


Figure 2.3 L'opérateur LBP [53]

'h (i)': représente la fréquence (nombre d'occurrences), L'histogramme un outil «visuel » qui permet de détecter certaines anomalies ou de faire un diagnostic avant d'engager une démarche .

$$h(i) = \sum_{x,y} B(LBP(x,y) = i) \tag{3}$$

$$\text{Avec : } i \in [0, \dots, 2^p - 1] \text{ et } B(v) = \begin{cases} 1, & \text{lorsque } v \text{ est vraie} \\ 0, & \text{autrement} \end{cases} \tag{4}$$

B.1.1 Histogramme

L'idée de l'opérateur Local Binary Pattern est d'attribuer à chaque pixel un code basé sur le niveau de gris de son voisinage [63]. Une fois que les codes LBP sont calculés pour tous les pixels de l'image, l'histogramme de l'image LBP est calculé pour former un vecteur caractéristique représentant l'image du visage [63].

En réalité, afin d'incorporer plus d'informations spatiales au vecteur représentant l'image (le visage), on divise tout d'abord l'image codée par l'opérateur LBP, en petites régions et l'histogramme sera intégré pour chaque région. Finalement,

on concatène tous les histogrammes des régions afin de former un grand histogramme qui représente l'image entière (voir la figure 2.4). L'efficacité du code LBP comme indice facial s'explique par le fait que le LBP permet de caractériser les détails fins d'un visage [54].

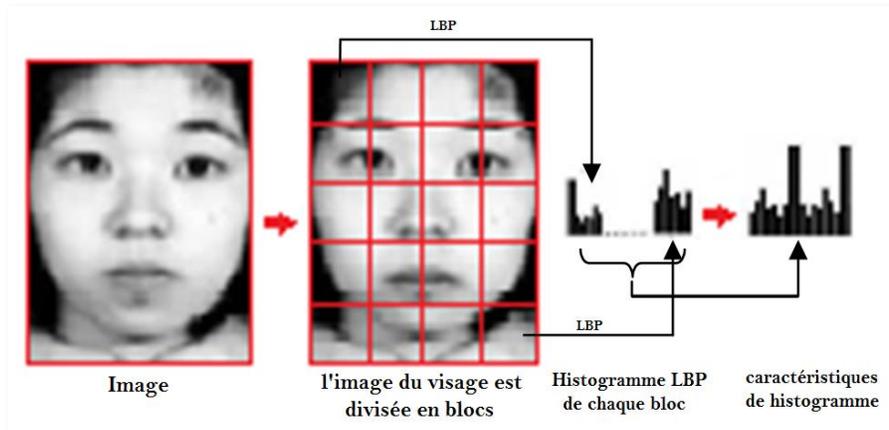


Figure 2.4 Extraction des caractéristiques de l'histogramme LBP. [28]

L'opérateur LBP (Local Binary Pattern) a été initialement proposé par Ojala, mais par la suite plusieurs variantes ont été proposées pour répondre aux limitations du LBP. Dans ce qui suit, nous décrivons brièvement quelques variantes.

B.2 Modèle binaire local composé (CLBP)

Comme l'opérateur LBP ne considère que le signe de la différence entre deux valeurs de gris, il échoue souvent à générer le code binaire approprié, la raison pour laquelle l'opérateur CLBP a été proposé pour incorporer des informations locales supplémentaires qui sont ignorées par l'opérateur original LBP.

Contrairement au LBP qui emploie un bit pour que chaque voisin pour exprimer le signe de la différence entre le pixel central et les niveaux de gris voisins, le CLBP utilise deux bits pour chaque voisin afin de représenter le signe ainsi qu'une information sur l'amplitude de la différence entre le pixel central et les niveaux de gris voisins [67].

Comme le montre la figure 2.9, le CLBP attribue un code 2 P-bit au pixel central basé sur les niveaux de gris d'un voisinage comprenant P voisins.

Le premier bit est utilisé pour représenter le signe de la différence entre le pixel central et les niveaux de gris voisins, tandis que l'autre bit est utilisé pour coder l'amplitude de la différence par rapport à une valeur seuil. L'opérateur CLBP définit le deuxième bit à 1 si l'amplitude de la différence entre le pixel central et le voisin correspondant est plus grande que du seuil M_{avg} . Sinon, il est mis à 0.

Une fonction d'indicateur $s(x)$ est définie comme suit :

$$s(i_p, i_c) = \begin{cases} 00 & i_p - i_c < 0. \quad |i_p - i_c| \leq M_{avg} \\ 01 & i_p - i_c < 0. \quad |i_p - i_c| > M_{avg} \\ 10 & i_p - i_c \geq 0. \quad |i_p - i_c| \leq M_{avg} \\ 11 & otherwise. \end{cases} \quad (5)$$

i_c : est la valeur de gris du pixel central, ' i_p ': est le niveau de gris d'un voisin p, Et " M_{avg} ": est la moyenne amplitude de la différence entre " i_p " et " i_c " dans le voisinage local. [47]

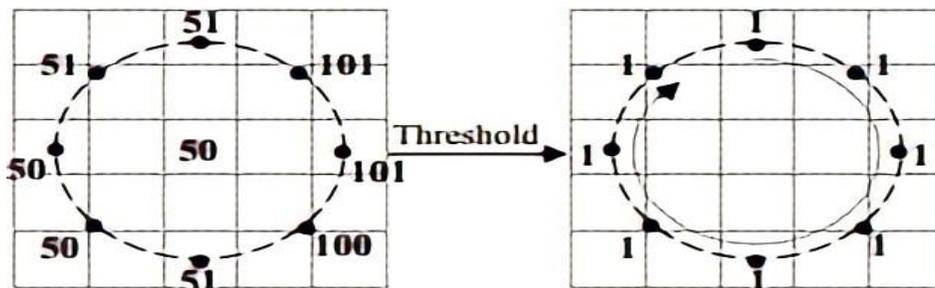


Figure 2.5 Génération d'un motif binaire incohérent dans le processus d'encodage LBP [67]

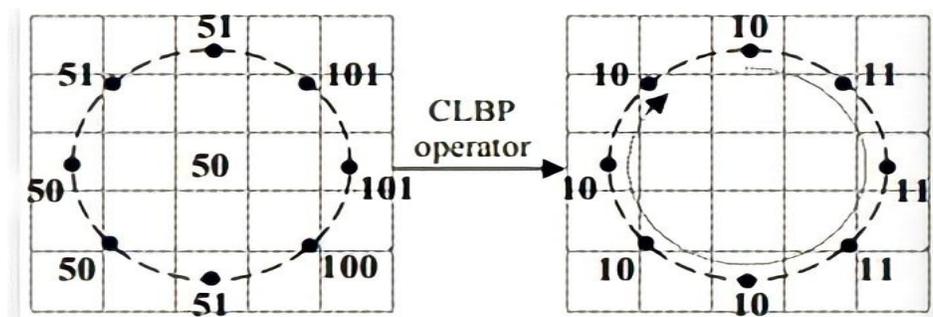


Figure 2.6 Illustration de l'opérateur CLBP de base,Ici, le modèle binaire est 1011111110101010 [67]

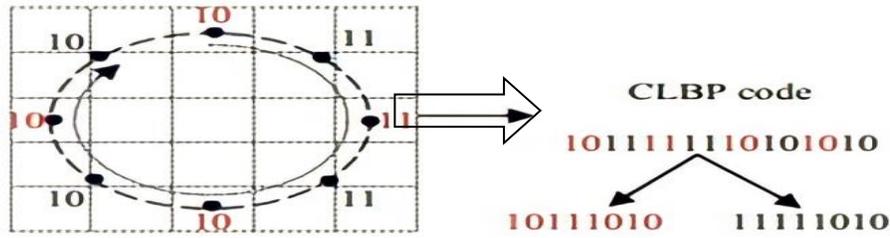


Figure 2.7 Génération des modèles de sous-CLBP a partir du code CLBP d’origine [69]

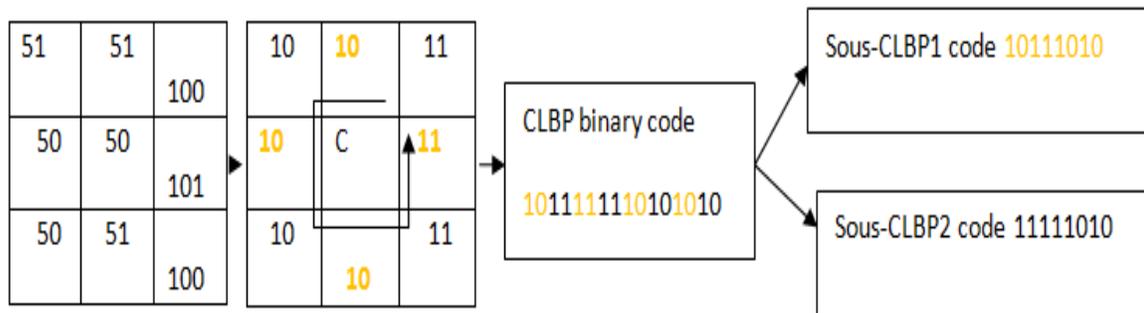


Figure 2.8 L’opérateur CLBP et la génération de deux sous-CLBP [55]

B.3 Modèle binaire Médian (MBP)

Hafiane et al [57] ont proposé le modèle binaire médian (MBP) pour la classification de texture. Cette méthode cherche à dériver le modèle binaire localisé par seuillage des pixels par rapport à leur valeur médiane sur un voisinage 3×3 . La figure 2.9 montre un exemple typique d’application de cette méthode.

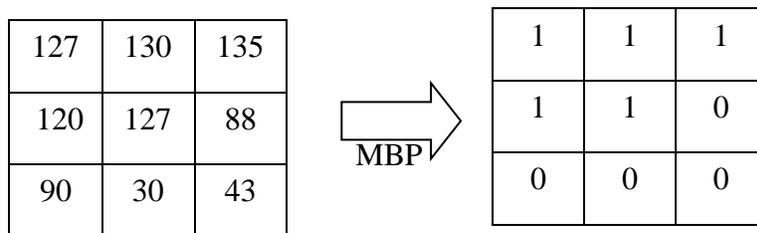


Figure 2.9 Exemple d’application de la méthode MBP

Dans l’exemple montré dans la figure 2.8, la valeur médiane est de 120. Le pixel central est inclus dans le processus de filtrage, nous obtenons donc 2^9 structures possibles. Alors,

$$\text{MBP} \sum_{i=1}^L \mathbf{f}(\mathbf{a}_i) \times 2^i \quad (6)$$

$$\mathbf{f}(\mathbf{a}_i) = \begin{cases} 1 & \text{Si } a_i \geq \text{Med} \\ 0 & \text{Sinon} \end{cases} \quad (7)$$

Où L est le nombre de voisins et a_i la valeur d'intensité. L'équation précédente produit une valeur dans l'intervalle fermé $[0 : 511]$.

Ici, le seuil ne dépend pas de la quantité d'intensité, ce qui rend l'opérateur MBP invariant aux changements d'échelle de gris. [58]

B.4 Motif ternaire local (LTP)

L'opérateur LTP (Local Ternary Pattern) proposé par Tan et Triggs [51] est une généralisation de la texture patchwise Local Binary Pattern (LBP). L'opérateur LTP étend la définition de l'opérateur LBP afin de générer un code pouvant prendre jusqu'à 3 valeurs (-1, 0 ou 1). Ceci est réalisé en modifiant la fonction échelon S de la manière suivante [64] :

$$LTP_{p,R} = \sum_{p=0}^{p-1} 2^p s(-i_c), \quad s(x) \begin{cases} 1, & x \geq t. \\ 0, & -t < x < t. \\ -1, & x < -t. \end{cases} \quad (8)$$

' t ': est le seuil a définie par l'utilisateur, L'avantage de cette méthode est qu'elle est plus robuste au bruit dans les régions uniformes [72].

B.5 Local phase Quantization (LPQ)

L'information de LPQ peut être extraite en utilisant la transformée discrète de Fourier à fenêtre à deux dimensions (2DWFT). [50]

$$\mathbf{F}_u(\mathbf{x}) = \sum_{\mathbf{m} \in N_x} \mathbf{h}(\mathbf{m}\mathbf{x}) \mathbf{f}(\mathbf{m}) e^{-j\pi \mathbf{u}^T \mathbf{m}} \mathbf{E}_u^T \mathbf{f}\mathbf{x} \quad (9)$$

Où \mathbf{E}_u^T , de taille = $1 \times M^2$, est un vecteur de base de 2DWFT avec la fréquence \mathbf{u} , et $\mathbf{f}\mathbf{x}$, de taille = $M^2 \times N$, est un vecteur contenant les valeurs des pixels d'image dans $N\mathbf{x}$ à chaque position de \mathbf{x} . La fonction fenêtre, $\mathbf{h}(\mathbf{x})$ est une fonction rectangulaire. La transformation est calculée à quatre valeurs de la fréquence :

$$\mathbf{u} = [\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] \text{ où } \mathbf{u}_0 = [\mathbf{a}, \mathbf{0}]^T, \mathbf{u}_1 = [\mathbf{0}, \mathbf{a}]^T, \mathbf{u}_2 = [\mathbf{a}, \mathbf{a}]^T \text{ et } \mathbf{u}_3 = [\mathbf{a}, -\mathbf{a}]^T \quad (10)$$

La méthode LPQ peut être résumée en quatre étapes distinctes. Dans un premier temps, l'opérateur (LPQ) est appliqué sur l'image d'entrée pour obtenir l'image labélisée. Ensuite, l'image obtenue est divisée en petites régions. Pour chacune d'entre elles, un histogramme des étiquettes est construit afin d'obtenir des vecteurs des caractéristiques (Template) locaux d'articulation de doigt (FKP). La représentation globale (vecteur des caractéristiques globales qui représente l'image entière) de l'articulation de doigt (FKP) est obtenue par combinaison de tous les vecteurs. La Figure (2.10) résume l'ensemble des étapes nécessaires à la génération de ce vecteur et du procédé de codage, en attribuant deux bits pour chaque pixel pour représenter le quadrant dans lequel et définit l'angle de phase. En fait, il fournit également la quantification de la fonction de phase de Fourier. En général, le LPQ est une chaîne binaire pour, chaque pixel, obtenue par la concaténation des codes quadrant bits réels et imaginaires des huit coefficients de Fourier de u_i . La chaîne binaire est convertie en nombre décimal pour produire une étiquette de LPQ.

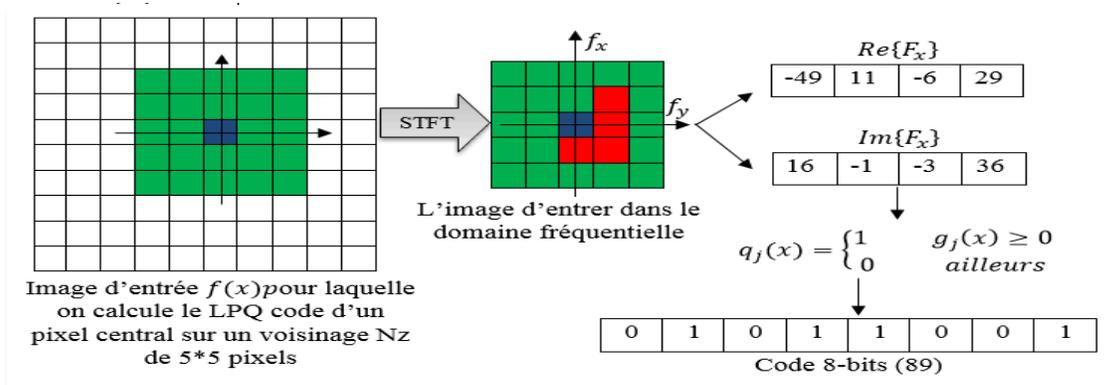


Figure 2.10 Organigramme de l'ensemble des étapes nécessaires à la génération du vecteur des caractéristiques par la méthode LPQ [56]

C. Méthode LBP uniforme

Le résultat du LBP traditionnel est classé en deux principaux types de valeurs. Le premier type « uniforme » contient moins de deux ou deux conversions au niveau du bit d'un à zéro ou vice versa, par exemple, zéro conversion (00000000) ou 2

conversions (01111110). Le deuxième type « non uniforme » a plus de deux conversions, par exemple, quatre conversions (11101101).

Le LBP uniforme, définis dans [56], contient au total $(P - 1) P + 2$ codes binaires. La mesure de l'uniformité enregistre le nombre de transitions spatiales dans la configuration binaire et le modèle uniforme qui contient aux plus deux transitions de bits. L'opérateur LBP uniforme ($LBP^{u2}_{P,R}$) est défini comme suit [57] :

$$LBP^{u2}_{P,R}(x,y) = \begin{cases} I(LBP_{P,R}(x,y)) & \text{si } U(LBP_{P,R}) \leq 2, I(z) \in [0, (P - 1)P + 2] \\ (P - 1)P + 2 & \text{sinon} \end{cases} \quad (11)$$

Où

$$U(LBP_{P,R}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^P |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (12)$$

U2 représenté dans l'expression de la mesure de l'uniformité $U(x)$ indique que la définition se rapporte à des modes uniformes avec une valeur $U \leq 2$. Si $U(x)$ est inférieur à 2, le pixel en cours est marqué par une fonction d'indexation $I(z)$. Sinon, il sera assigné à la valeur $(P - 1) P + 2$.

La fonction d'index, $I(z)$, contenant $(P - 1) P + 2$ indices, est utilisée pour attribuer un index particulier à chacune des configurations uniformes.

D. Histogram of Oriented Gradient (HOG)

L'histogramme de gradient orienté HOG proposé par Dalal et Triggs [75], est une méthode utilisée en vision par ordinateur pour la reconnaissance d'objets. Cette technique calcule un histogramme local dans la direction du gradient sur une grille dense, c'est-à-dire dans des zones régulièrement réparties dans l'image. L'idée importante derrière le descripteur HOG est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution d'intensité du gradient ou la direction des bords. Cela peut être fait en divisant l'image en petites zones continues appelées cellules et en calculant un gradient ou un histogramme de bord des pixels de cette cellule pour chaque cellule. La combinaison des histogrammes forme alors le descripteur HOG.

La procédure de mise en œuvre de l'algorithme de descripteur HOG est donnée comme suit :

Étape 1 : Calcul des gradients :

Le calcul des gradients horizontaux et verticaux se réalise en filtrant l'image avec les noyaux suivants :

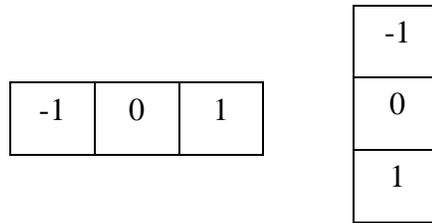


Figure 2.11 gradients horizontaux et verticaux

Ensuite, l'amplitude et la direction des gradients seront calculées en utilisant la formule suivante :

$$Grad_{Mag} = \sqrt{gx^2 + gy^2} \tag{13}$$

$$Grad_{Direction} = \tan\left(\frac{gy}{gx}\right) \tag{14}$$

Étape 2 : Calcul de l'histogramme des gradients dans des cellules de taille précise (soit 8x 8) ;

Dans cette étape, l'image est divisée en 8x8 cellules et un histogramme de gradients est calculé pour chaque 8x8 cellules.

L'étape suivante consiste à créer un histogramme de gradients dans ces 8 x 8 cellules. L'histogramme contient 9 bins correspondant aux angles 0, 20, 40 ... 160. Le bin est sélectionné en fonction de la direction et le vote est sélectionné en fonction de l'amplitude. Les contributions de tous les pixels dans les cellules 8 x 8 sont additionnées pour créer l'histogramme 9-bin ;

Étape 3 : Calcul du vecteur de caractéristiques de l'histogramme des gradients orientés ;

Pour calculer le vecteur de caractéristiques final qui représente l'image entière, les différents histogrammes obtenus pour chaque 8×8 cellules seront concaténés pour former un seul vecteur caractéristique.

E. La transformation de caractéristique invariante à l'échelle (SIFT)

Les descripteurs SIFT, développé en 1999 par le chercheur David Lowe [73], consistent à extraire d'une image des caractéristiques invariantes par changement d'échelle. Les caractéristiques sont calculées à partir du vecteur de gradient pour chaque pixel du voisinage afin de construire un histogramme normalisé des directions de gradient. Ces caractéristiques sont invariantes aux translations, aux rotations et aux transformations de mise à l'échelle.

Les étapes d'extraction des caractéristiques SIFT sont données comme suit :

1. Détection extrême de l'espace d'échelle

Le principe général de la détection extrême est basé sur la différence d'image floue gaussienne à différentes échelles (Différence of Gaussians) pour chercher des extrêmes locaux, les points correspondant à ces extrêmes trouvés sont appelés points clés.

2. Positionnement des clés

L'étape de détection d'extremums produit en général un grand nombre de points-clés, dont certains points clés peuvent être relativement difficiles à identifier. De ce fait, des traitements supplémentaires sont appliqués, pour éliminer les points de faible contraste ou ceux situés sur des arêtes de contour à faible courbure.

3. Orientation

Après calcul d'histogramme directionnel du voisinage local de la clé, la direction du maximum dans l'histogramme sera définie comme direction principale de la clé.

4. Descripteur de clés

Le descripteur SIFT est la représentation des résultats statistiques du gradient d'image gaussien du voisinage d'un point clé. Une région de 16×16 pixels, subdivisée en 4×4 zones de 4×4 pixels chacune est toujours considérée autour

d'un point-clé. Sur chaque zone est calculé un histogramme des orientations comportant 8 intervalles et en chaque point de la zone, l'orientation et l'amplitude du gradient sont calculés. Les 16 histogrammes à 8 intervalles chacun sont par la suite concaténés et normalisés. Dans le but de diminuer la sensibilité du descripteur aux changements de luminosité, les valeurs sont plafonnées à 0,2 et l'histogramme est de nouveau normalisé, pour finalement fournir le descripteur SIFT du point-clé, de dimension 128.

2.2.3 Classification

La dernière étape d'un système de reconnaissance automatique d'expressions faciales est la détermination de l'expression faciale en fonction des caractéristiques extraites. De nombreux classifieurs ont été appliqués pour la reconnaissance des expressions faciales plus particulièrement les Machines à vecteurs de support, le K-plus proche voisin, le Random Forest, l'Analyse Discriminante Linéaire, les Réseaux bayésiens, les Réseaux de neurone et d'autres.

- Les Machines à vecteurs de support ;
- Le K-plus proche voisin ;
- Le Random Forest ;
- L'Analyse Discriminante Linéaire ;

2.2.3.1 Machines à vecteurs de support

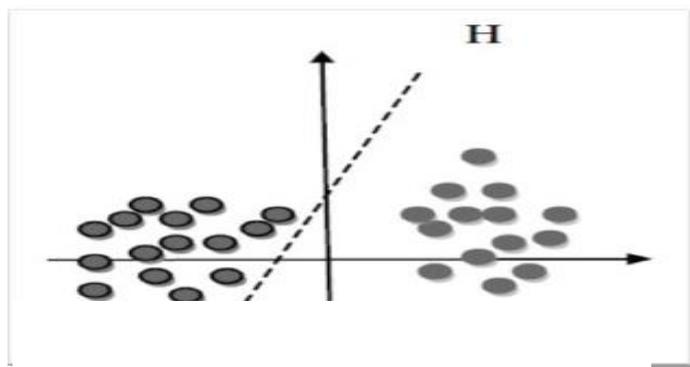


Figure 2.12 Exemple de classification SVM

Les SVM sont un ensemble de techniques d'apprentissage supervisé de nouvelle génération basées sur l'utilisation des fonctions capables de séparer les données de manière optimale. Lorsque l'on considère le cas de deux classes dont les données sont linéairement séparables, il existe une infinité d'hyperplans pour séparer les observations. L'objectif des SVM est de trouver l'hyperplan optimal qui sépare les données avec la maximisation de la distance entre les différentes classes. On appelle hyperplan séparateur l'hyperplan qui sépare deux classes.

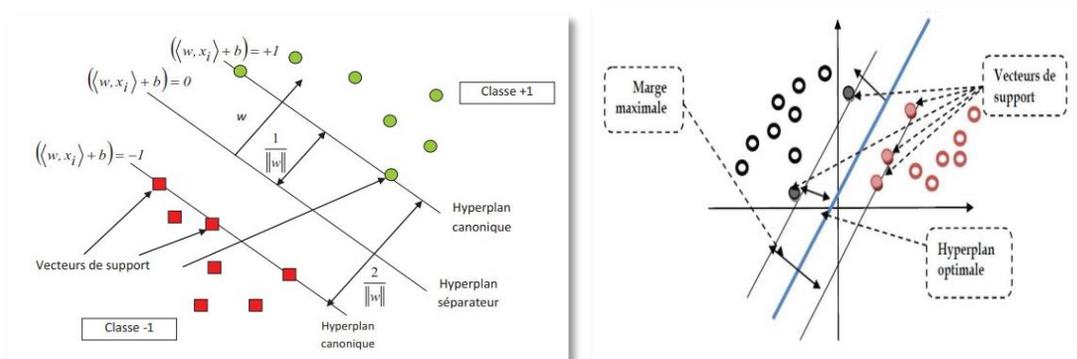
L'équation de l'hyperplan séparateur est donnée par :

$$h(x) = (w \cdot x) + b \quad \text{ou} \quad \sum \alpha_i * \mu_i(\phi(x_i), x_j) \quad (15)$$

Où $\phi(x_i)$ est une transformation non linéaire de l'espace d'entrée X en un espace de représentation $\phi(x)$.

Le principe des SVM est de choisir celle qui maximise la distance entre l'hyperplan et les échantillons d'apprentissage (c'est-à-dire hyperplan et vecteurs de support), cette distance est appelée la marge.

Parmi les modèles de SVM, il y a le Linéaire-SVM et le non-Linéaire-SVM. Les premiers sont les SVM les plus simples car ils peuvent séparer linéairement les données tandis que pour les seconds sont utilisées pour les données qui ne sont pas linéairement séparables. Dans ce dernier cas, les données sont transformées pour être représentées dans un grand espace où elles sont linéairement séparables. [85]



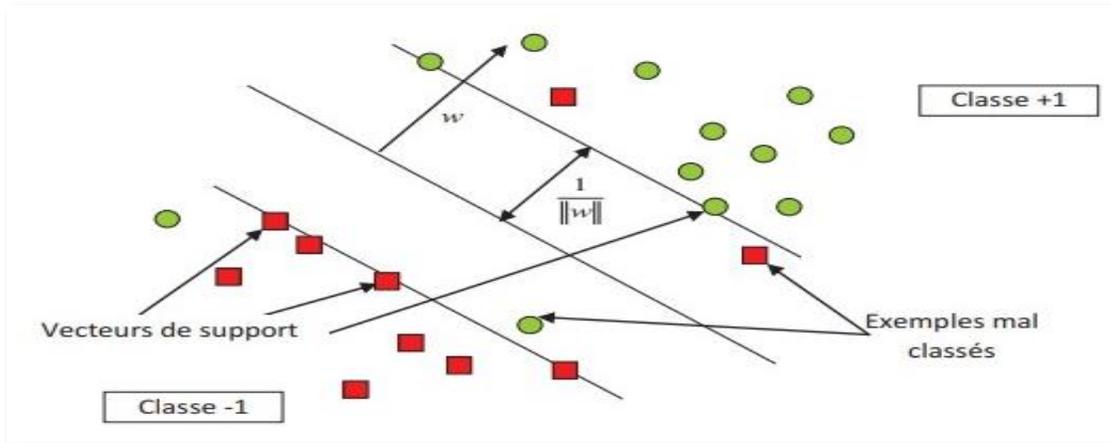


Figure 2.13 Séparation linéaire et non linéaire

Afin de pouvoir traiter des cas où les données ne sont pas linéairement séparables. Les fonctions du noyau sont utilisées pour cartographier efficacement les données d'entrée qui peuvent ne pas être séparées linéairement dans un espace de grande dimension, où des méthodes linéaires peuvent alors être appliquées.

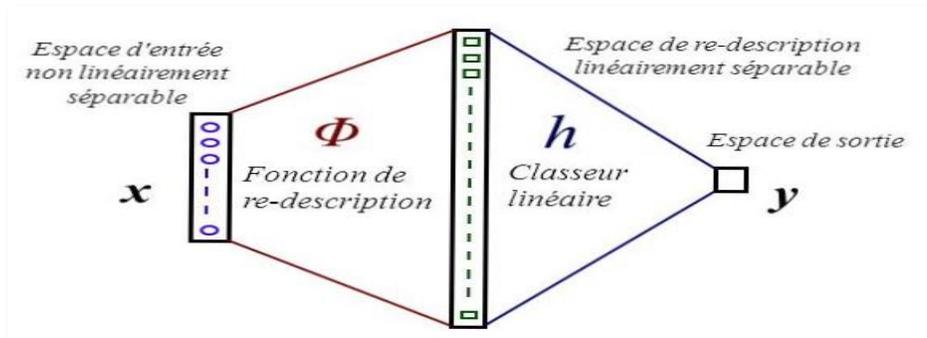


Figure 2.14 Le passage par une description des données peut permettre une séparation linéaire des exemples

Les SVM ont été conçues à l'origine pour résoudre des problèmes bi-classes. L'adaptation des SVM bi-classes au cas multi-classes peut se faire de deux façons différentes :

- ✓ **L'approche un contre tous** : elle consiste à entraîner un SVM bi-classes en utilisant les éléments d'une classe contre tous les autres. Il s'agit de résoudre de l'ordre de c problèmes SVM chacun de taille n ;

- ✓ **L'approche un contre un** : elle consiste à entraîner $c(c-1)/2$ SVM sur chacun des couples de classes, puis à décider la classe gagnante soit par un vote majoritaire soit en post-traitant les résultats grâce à l'estimation de probabilités a posteriori.

2.2.3.2 K-plus proche voisin

L'algorithme de K-plus proche voisin est un algorithme d'apprentissage supervisé qui permet de faire des prédictions sur des variables qualitative ou quantitative. Il consiste à trouver les K instances les plus proches en calculant la similarité entre l'instance à classer et les instances déjà classées (les instances de la base d'apprentissage).

Le principe de cette méthode est le suivant :

- On note x une nouvelle instance décrite par un vecteur de p attributs

$$x = \{x_1, x_2, \dots, x_p\}$$
- Chaque instance I_i de l'ensemble d'entraînement est sous forme d'un couple,

$I_i = \langle a_{1i}, a_{2i}, \dots, a_{pi}, c_i \rangle$ Tel que $(a_{1i}, a_{2i}, \dots, a_{pi})$ représente un vecteur de p attributs de l'instance I_i , c_i représente la classe de l'instance I_i .

L'algorithme consiste à trouver parmi l'ensemble d'instances de la base d'apprentissage, les k plus proches voisins de x et associe à x la classe majoritaire parmi ses k voisins (les plus proches).

Un exemple de classification KNN est illustré par la figure 2.15:

En effet, le point inconnu (étoile) appartient soit à la première classe (carré) ou à la deuxième classe (triangle).

Si $K = 3$, le point inconnu est classé en deuxième classe (triangle) parce qu'il y a deux triangles et un seul carré parmi les trois plus proches exemples à l'intérieur du cercle.

Si $K = 5$, il est classé dans la première classe (carré).

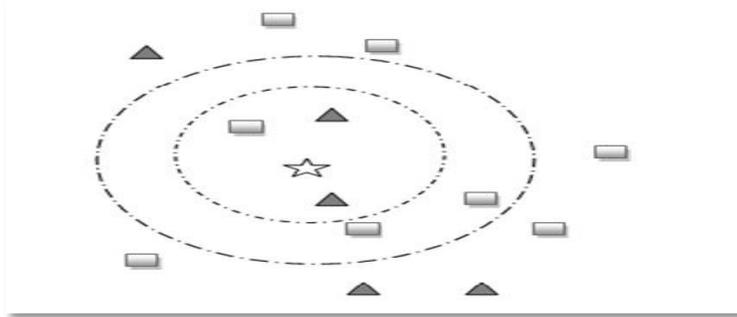


Figure 2.15 Exemple de classification KNN (K=3 et K=5)

2.2.3.3 Random Forest (RF)

Une forêt aléatoire est un algorithme d'apprentissage automatique supervisé qui est construit à partir d'algorithmes d'arbre de décision. Il est constitué de nombreux arbres de décision. Un arbre de décision se compose de trois composants : des nœuds de décision, des nœuds feuilles et un nœud racine. Le classifieur RF peut être défini comme suit :

$$H(x) = \arg \max_y \sum_1^K I(h_i(x) = y) \quad (16)$$

Où $H(x)$ est le classifieur combiné final, k est le nombre d'arbres de décision, $h_i(x)$ représente un arbre de décision, Y est l'étiquette de classe, $I(h_i(x) = Y)$ indique x appartient à la classe Y .

2.2.3.4 Analyse discriminante linéaire

L'algorithme d'analyse discriminante linéaire (LDA) est né des travaux de BELHUMEUR et AL. De Yale Université (États-Unis), en 1997 [67]. La LDA analyse les vecteurs propres de la matrice de dispersion des données, avec pour objectif de maximiser les variations entre les images d'individus différents (interclasses) tout en minimisant les variations entre les images d'un même individu (intra-classes).

$$\frac{|\phi^T S_b \phi|}{|\phi^T S_w \phi|} \quad (17)$$

$$a \text{ matrice de variance inter - classes } S_b : \sum_{i=1}^M (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (18)$$

$$a \text{ matrice de variance intra - classes } S_w \sum_{i=1}^M \sum_{x \in X_i} (x_i - \bar{x})(x_i - \bar{x})^T l. \quad (19)$$

M étant le nombre d'individus de la base de données pour construire l'espace de projection Φ [78].

Pour une classification supervisée. L'algorithme d'analyse discriminante linéaire peut être appréhendé de deux façons complémentaires :

- Une approche géométrique qui revient à chercher des hyperplans qui séparent au mieux les groupes
- Une approche modèle qui fait l'hypothèse que les lois des variables sont des vecteurs gaussiens avec des valeurs de paramètres différentes pour chaque groupe.

2.3 Conclusion

Dans ce chapitre, nous avons d'abord présenté un aperçu général du processus de conception des systèmes de reconnaissance des expressions faciales, et nous avons décrit brièvement quelques techniques de pointe proposées dans la littérature pour chaque phase de conception. Dans le chapitre suivant, nous allons nous intéresser à l'implémentation de différentes approches d'extraction de caractéristique et de classification pour la reconnaissance automatiquement des expressions faciales afin de faire une comparaison synthétique des résultats en termes de performance globale du système de reconnaissance.

Chapitre III

Résultats

Expérimentaux et

Discussion

Chapitre 3

Résultats expérimentaux et discussion

3.1 Introduction

Après avoir abordé l'aspect théorique de la reconnaissance des expressions faciales dans les deux chapitres précédents, nous illustrons dans ce dernier chapitre le côté pratique. Dans la première partie, nous présentons une description bien détaillée sur l'environnement matériel et logiciel utilisé pour mesurer les performances du système de reconnaissance faciale. Ici, les simulations sont effectuées sur les bases de données sélectionnées. Nous réservons la deuxième partie de ce chapitre pour les différents tests d'application des méthodes d'extraction des caractéristiques et de classification afin d'évaluer les performances de reconnaissance. Nous finissons notre chapitre par une étude comparative et discussion des résultats obtenus.

3.2 Base de données et protocole d'évaluation

3.2.1 Description des bases de données

Pour l'évaluation des différentes approches implémentées dans ce travail, deux bases de données ont été utilisées :

- ✓ **Base de données (JAFEE)**

La base de données JAFEE (Japanese Female Facial Expression) [1] se compose de 213 images d'expressions faciales de dix femmes japonaises. Pour chaque femme,

simulé 7 expressions faciales (6 expressions faciales de base et l'état neutre) ont été simulées. La résolution des images est de 256×256 pixels.

✓ **Base de données (MUG)**

La base de données MUG (MultiMedia Under standing Group) comprend des images de 86 sujets exécutant des expressions faciales. Les sujets étaient 35 femmes et 51 hommes âgés entre 20 et 35 ans. Chaque image de cette base a été enregistrée avec un format JPG, 896×896 pixels et une taille allant de 240 à 340 Ko [2].

3.2.2 Protocol d'évaluation utilisé

Pour assurer les deux phases d'apprentissage et de test, chaque base de données a été divisée en deux parties. Pour chaque expérience, 70 % des images ont été sélectionnées pour l'apprentissage, tandis que les 30 % restantes ont été sélectionnées pour les tests.

3.3 Expérimentations et Résultats

3.3.1 Environnement du travail

Dans cette section, nous présenterons les environnements matériels et logiciels de notre travail.

3.3.1.1 Environnement matériel

Un ordinateur LENOVO avec les caractéristiques suivantes :

- Processeur : Intel (R) Core (TM) i5-4300U CPU @ 1.90GHz 2.49 GHz
- Mémoire RAM : 4,00 Go.
- OS : Microsoft Windows 10 Professionnel 64bits.

3.3.1.2 Environnement logiciel

✓ **L'environnement de programmation utilisé est MATLAB R2016b**

La plate-forme MATLAB est optimisée pour résoudre les problèmes scientifiques et techniques. Le langage MATLAB, basé sur le calcul des matrices, est le moyen le

plus naturel au monde pour exprimer les mathématiques computationnelles. Il dispose de plusieurs boîtes à outils en particulier celle du traitement d'images « Image Processing Tool Box » qui propose un ensemble d'algorithmes et d'outils graphiques de référence permettant de traiter, analyser, visualiser et développer des algorithmes de traitement d'images.

3.3.2 Méthodologie

Dans une première étape, la technique de redimensionnement d'images est appliquée pour obtenir les mêmes dimensions pour toutes les images de chaque base de données. Les images de la base de données JAFFE ont été redimensionnées en 160X160 pixels et les images de la base de données MUG en 240 x 240 pixels. Ensuite, afin d'extraire les traits des images faciales, les deux approches HOG et LBP sont appliquées. Nous avons utilisé le descripteur HOG pour extraire l'information de forme et le descripteur LBP pour extraire l'information de texture. Pour le descripteur LBP, nous avons choisi d'appliquer la variante LBP uniforme car elle présente l'avantage du gain en mémoire et en temps de calcul. Finalement, nous avons examiné deux différentes méthodes d'apprentissage automatique : notamment la SVM (l'approche un contre tous) et le K-plus proche voisin 'KNN' pour la reconnaissance de sept expressions faciales (la joie, la tristesse, la surprise, la peur, la colère, le dégoût et la neutralité) sur les deux bases de données.

Notre objectif est de tester les différentes approches dans le but d'atteindre une meilleure performance de reconnaissance. Pour l'évaluation de la performance de reconnaissance, nous avons utilisé deux critères qui sont :

La mesure du taux de reconnaissance Calculé de la manière suivante :

$$\text{Taux de reconnaissance} = \frac{(\text{nombre d'images de test correctement classées})}{(\text{nombre total des images de test})} \quad (20)$$

Le temps de réponse du système de reconnaissance : testé en fonction du nombre de caractéristiques extraites en utilisant les deux approches LBP et HOG. Pour cela, nous avons pris en charge le traitement d'image par blocs afin de voir l'effet du

changement de la taille des vecteurs caractéristiques sur le temps de réponse du système de reconnaissance.

3.3.3 Résultats

3.3.3.1 Evaluation sur la base de données JAFFE

➤ **Dans une première section** : le LBP est utilisé pour extraire les caractéristiques des images faciales. D'autre part, les classificateurs KNN et SVM sont utilisés pour la classification.

Remarque : différentes tailles de cellule sont utilisées dans chaque base de données pour montrer l'efficacité de la taille des cellules sur les performances du classifieur.

✓ **Taille de la cellule= 8x8 :**

Tableau 3.1Résultats d'application du LBP et KNN/ SVM sur la base de données JAFFE

(Taille de cellule= 8x8)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Précision Colère	0.00%	100.00%
Précision Dégoût	66.67%	0.00%
Précision Peur	0.00%	0.00%
Précision Joie	100.00%	100.00%
Précision Neutre	100.00%	0.00%
Précision Tristesse	66.67%	0.00%
Précision Surprise	33.33%	100.00%
Précision globale	52.28%	42.85%
Nombre de caractéristiques de visage	23600	23600
Temps de réponse	6.818	13.121

✓ **Taille de cellule= 16x 16 :**

Tableau 3.2Résultats d'application du LBP et KNN/ SVM sur la base de données JAFFE
(Taille de cellule= 16x 16)

Précision	En utilisant KNN	En utilisant SVM
Expressions		
Précision Colère	0.00%	100.00%
Précision Dégout	33.33%	33.33%
Précision Peur	75.00%	0.00%
Précision Joie	33.33%	100.00%
Précision Neutre	66.67%	100.00%
Précision Tristesse	33.33%	0.00%
Précision Surprise	33.33%	100.00%
Précision globale	39.28%	59.09%
Nombre de caractéristiques de visage	5900	5900
Temps de réponse	5.998 s	7.681 s

✓ **Taille de cellule = 24x24**

Tableau 3.3Résultats d'application du LBP et KNN/ SVM sur la base de données JAFFE
(Taille de cellule= 24 x 24)

Précision	En utilisant KNN	En utilisant SVM
Expressions		
Précision Colère	33.33%	100.00%
Précision Dégout	33.33%	33.33%
Précision Peur	100.00%	50.00%
Précision Joie	33.33%	66.67%
Précision Neutre	100.00%	100.00%
Précision Tristesse	66.67%	0.00%

Précision Surprise	33.33%	100.00%
Précision globale	57.14%	64.28%
Nombre de caractéristiques de visage	2124	2124
Temps de réponse	5.458 s	6.706 s

En analysant les résultats obtenus dans cette première section (tableau 3.1, tableau 3.2 et tableau 3.3) en terme de taux de reconnaissance, on trouve qu'en utilisant la méthode LBP avec le KNN et lorsque la taille de cellule égale à 8, la meilleure précision obtenue est de 100 % avec l'expression joie et l'état neutre, et pour la taille de cellule égale à 16, la précision obtenue est de 75% avec l'expression peur. De plus, pour la taille de cellule égale à 24, la précision obtenue est de 100 % avec l'expression peur et l'état neutre. D'autre part, les résultats obtenus par la méthode LBP avec le SVM et lorsque la taille de la cellule est égale respectivement 8 et 24, la précision obtenue atteint 100 % avec les expressions colère, joie et surprise. En outre, pour la taille de cellule égale à 16, la précision obtenue atteint 100 % avec les expressions colère, joie, surprise et l'état neutre. Cependant, la méthode LBP avec le KNN atteint une précision globale de 57,14% en utilisant une taille de cellule égale à 24, tandis que la méthode LBP avec le SVM atteint une précision globale de 64.28% en utilisant une taille de cellule égale à 24.

Concernant l'analyse des résultats en termes de temps de réponse, les résultats obtenus montrent que plus le nombre de caractéristiques extraites pour assurer les deux phases d'apprentissage et de test est grand plus la réponse de système est lente. Pour une taille de cellule égale à 8, et avec une taille d'image égale à 160 x160, le nombre de blocs est alors $(160/8 \times 160/8) 20 \times 20 = 400$ blocs. Chaque bloc est représenté par un histogramme de 59 valeurs (LBP uniforme) et dans ce cas le vecteur final résultant de la concaténation des différents histogrammes des différents blocs est alors de taille $400 \times 59 = 23600$ valeurs. Tandis que, le nombre de caractéristiques obtenu, pour une taille de cellule égale à 16, est égal à 5900 (nombre de blocs égale à 100). Pour une taille de cellule égale à 24 (nombre de blocs égale à 36), ce nombre est égal à 2124.

Bien que ce nombre ait parfois un impact négatif sur les performances du système, cependant comme le montre les résultats, il permet dans certain cas d'augmenter le taux de reconnaissance.

3.3.3.2 Evaluation sur la méthode HOG

➤ **Dans une deuxième section :** le HOG est utilisé pour extraire les caractéristiques des images faciales. D'autre part, les classificateurs KNN et SVM sont utilisés pour la classification.

De même, différentes tailles de cellule sont utilisées dans chaque base de données pour montrer l'efficacité de la taille des cellules sur les performances du classifieur.

✓ **Taille de cellule = 8 x8 :**

Tableau 3.4 Résultats d'application du HOG et KNN/ SVM sur la base de données JAFFE
(Taille de cellule= 8 x 8)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Précision Colère	100.00%	100.00%
Précision Dégoût	100.00%	33.33%
Précision Peur	0.00%	100.00%
Précision Joie	100.00%	100.00%
Précision Neutre	100.00%	100.00%
Précision Tristesse	33.33%	100.00%
Précision Surprise	66.67%	100.00%
Précision globale	71.42%	90.47%
Nombre de caractéristiques de visage	12996	12996
Temps de réponse	6.481 s	10.540

✓ **Taille de cellule = 16x 16 :**

Tableau 3.5 Résultats d'application du HOG et KNN/ SVM sur la base de données JAFFE
(Taille de cellule= 16 x 16)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Colère	100.00%	100.00%
Dégoût	66.67%	66.67%
Peur	0.00%	100.00%
Joie	100.00%	100.00%
Neutre	66.67%	100.00%
Tristesse	66.67%	100.00%
Surprise	66.67%	66.67%
Précision globale	66.66%	90.47%
Nombre de caractéristiques de visage	2916	2916
Temps de réponse	6.218 s	7.700 s

✓ **Taille de cellule = 24 x24 :**

Tableau 3.6 Résultats d'application du HOG et KNN/ SVM sur la base de données JAFFE
(Taille de cellule= 24 x 24)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Précision Colère	100.00%	100.00%
Précision Dégoût	66.67%	33.33%
Précision Peur	0.00%	100.00%
Précision Joie	100.00%	100.00%
Précision Neutre	100.00%	100.00%
Précision Tristesse	100.00%	66.67%

Précision Surprise	66.67%	66.67%
Précision globale	76.19%	80.85%
Nombre de caractéristiques de visage	900	900
Temps de réponse	6.146 s	7.199 s

En analysant les résultats obtenus dans cette section (tableau 3.4, tableau 3.5 et tableau 3.6) en terme de taux de reconnaissance, on trouve qu'en utilisant la méthode HOG avec le KNN et lorsque la taille de cellule = 8, la meilleure précision obtenue était de 100 % avec les expressions colère, dégoût et l'état neutre, et pour la taille de cellule = 16, la précision obtenue était aussi 100% avec les deux expressions colère et joie, de plus; pour la taille de cellule = 24, la précision obtenue était de 100 % avec les expressions colère, joie, tristesse et l'état neutre. D'autre part, les résultats obtenus par la méthode HOG avec le SVM et lorsque la taille de cellule = 8, la précision obtenue atteignait 100 % avec toutes les expressions, sauf l'expression dégoût, de plus ; pour la taille de cellule = 16, la précision obtenue à atteignait 100 % avec les expressions peur, joie, tristesse et l'état neutre, et pour la taille de cellule = 24, la précision obtenue était aussi 100% avec les expressions colère, peur, joie et l'état neutre. Cependant, la méthode HOG avec le KNN a atteint une précision globale de 76,19% en utilisant une taille de cellule = 24, tandis que la méthode HOG avec le SVM a atteint une précision globale de 90.47% en utilisant une taille de cellule = 8 ou 16.

En outre, l'analyse des résultats en termes de temps de réponse, montre comme les résultats obtenus dans la section précédente, que plus le nombre de caractéristiques extraites pour assurer les deux phases d'apprentissage et de test est grand plus la réponse de système est lente.

Pour une taille de cellule= 8 et avec une taille d'image égale à 160 x160, le HOG divise l'image en 20×20 blocs. Avec un chevauchement de 50%, Il en résulte un total de $19 \times 19 = 361$ blocs (Chaque bloc se compose de 2×2 cellules chacune de taille 8×8). La plage d'orientation des angles est définie entre $[0-180]$ et quantifiée à l'aide d'histogrammes de 9 Bins. De cette façon, le nombre total de caractéristiques

HOG pour une image d'entrée devient $19 \times 19 \times 2 \times 2 \times 9 = 12996$. En suivant le même concept, Pour une taille de cellule= 16, le nombre total de caractéristiques HOG égale à $9 \times 9 \times 2 \times 2 \times 9 = 2916$ et pour une taille de cellule= 24, le nombre total de caractéristiques HOG égale à $5 \times 5 \times 2 \times 2 \times 9 = 900$. Ainsi, plus la taille de cellule est élevée plus le nombre de caractéristiques HOG résultantes est faible, ce qui permet une réponse rapide du système.

3.3.3.3 Evaluation sur la base de données MUG

Afin de vérifier et confirmer les résultats obtenus sur la base de données JAFFE, nous avons refait le même travail sur la base de données MUG.

➤ **Dans cette section** le LBP est utilisé pour extraire les caractéristiques des images faciales et les classificateurs KNN et SVM sont utilisés pour la classification.

✓ **Taille de la cellule= 8x8**

Tableau 3.7Résultats d'application du LBP et KNN/ SVM sur la base de données MUG
(Taille de cellule= 8 x 8)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Colère	20.00%	100.00%
Dégoût	80.00%	100.00%
Peur	100.00%	0.00%
Joie	100.00%	100.00%
Neutre	0.00%	0.00%
Tristesse	100.00%	100.00%
Surprise	100.00%	100.00%
Précision globale	71.43%	71.43%
Nombre de caractéristiques de visage	36875	36875
Temps de réponse	108.823 s	3.350

✓ Taille de cellule= 16x 16 :

Tableau 3.8 Résultats d'application du LBP et KNN/ SVM sur la base de données MUG
(Taille de cellule= 16x 16)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Précision Colère	20.00%	100.00%
Précision Dégoût	80.00%	100.00%
Précision Peur	100.00%	20.00%
Précision Joie	100.00%	100.00%
Précision Neutre	20.00%	60.00%
Précision Tristesse	100.00%	100.00%
Précision Surprise	100.00%	100.00%
Précision globale	74.29%	82.86%
Nombre de caractéristiques de visage	8496	8496
Temps de réponse	29.970 s	36.757 s

✓ Taille de cellule = 24x24

Tableau 3.9 Résultats d'application du LBP et KNN/ SVM sur la base de données MUG
(Taille de cellule= 24 x 24)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Colère	0.00%	100.00%
Dégoût	40.00%	100.00%
Peur	80.00%	60.00%
Joie	100.00%	100.00%
Neutre	80.00%	80.00%
Tristesse	100.00%	100.00%
Surprise	100.00%	100.00%

Précision globale	71.43%	91.43%
Nombre de caractéristiques de visage	3776	3776
Temps de réponse	25.800 s	32.145 s

➤ **Dans cette section :** le HOG est utilisé pour extraire les caractéristiques des images faciales et les classificateurs KNN et SVM sont utilisés pour la classification.

✓ **Taille de cellule = 8 x8 :**

Tableau 3.10 Résultats d'application du HOG et KNN/ SVM sur la base de données MUG
(Taille de cellule= 8 x 8)

Précision	En utilisant KNN	En utilisant SVM
Expressions		
Colère	0.00%	100.00%
Dégoût	100.00%	100.00%
Peur	40.00%	100.00%
Joie	100.00%	100.00%
Neutre	60.00%	100.00%
Tristesse	100.00%	100.00%
Surprise	60.00%	100.00%
Précision globale	65.71%	100.00%
Nombre de caractéristiques de visage	20736	20736
Temps de réponse	36.875 s	3.257

✓ **Taille de cellule = 16x 16 :**

Tableau 3.11 Résultats d'application du HOG et KNN/ SVM sur la base de données MUG
(Taille de cellule= 16 x 16)

Précision	En utilisant KNN	En utilisant SVM
Expressions		
Colère	60.00%	100.00%

Dégoût	20.00%	100.00%
Peur	100.00%	100.00%
Joie	100.00%	100.00%
Neutre	100.00%	100.00%
Tristesse	100.00%	100.00%
Surprise	60.00%	100.00%
Précision globale	77.14%	100.00%
Nombre de caractéristiques de visage	4356	4356
Temps de réponse	22.950 s	26.469 s

✓ **Taille de cellule = 24 x24 :**

Tableau 3.12 Résultats d'application du HOG et KNN/ SVM sur la base de données MUG
(Taille de cellule= 24x 24)

Expressions \ Précision	En utilisant KNN	En utilisant SVM
Colère	40.00%	80.00%
Dégoût	100.00%	100.00%
Peur	100.00%	100.00%
Joie	100.00%	100.00%
Neutre	100.00%	100.00%
Tristesse	100.00%	100.00%
Surprise	20.00%	100.00%
Précision globale	80.00%	97.14%
Nombre de caractéristiques de visage	1764	1764
Temps de réponse	20.808 s	23.427 s

En analysant les résultats obtenus sur la base de données MUG, nous avons constaté que cela ressemble à ce que nous avons obtenu sur la base de données JAFFE :

- ✓ Certaines approches sont parfaitement capables de reconnaître certaines émotions mais pas capables d'en reconnaître d'autres. Cela est dû à la nature des changements faciaux causé par chaque émotion et qui peuvent être bien détectés (ou ignorés) par le descripteur utilisé.
- ✓ Le temps de réponse du système augmente avec l'augmentation du nombre de caractéristiques faciales utilisées pour l'entraînement du classificateur et les tests.
- ✓ Avec un même nombre de caractéristiques faciales, un classificateur peut être plus rapide qu'un autre (le KNN plus rapide que le SVM).
- ✓ La meilleure performance de reconnaissance est obtenue en utilisant le classifieur SVM avec des caractéristiques HOG.

3.4 Conclusion

Ce dernier chapitre a été consacré à l'évaluation des performances d'un système de reconnaissance de sept expressions faciales (la joie, la tristesse, la surprise, la peur, la colère, le dégoût et la neutralité) par évaluation de l'efficacité de deux méthodes utilisées pour l'extraction des caractéristiques faciales. Il s'agit LBP et du HOG, ainsi que deux classificateurs notamment le SVM et le KNN. L'étude a été menée sur les deux bases de données JAFFE et MUG.

Plusieurs tests ont été faits dans le but de trouver la meilleure performance de reconnaissance en fonction de deux critères qui sont le taux de reconnaissance et le temps de réponse. Au travers les différentes expériences, nous avons trouvé que la meilleure performance de reconnaissance est obtenue en utilisant le classifieur SVM avec des caractéristiques HOG extraites en fixant la taille de cellule à 8 ou à 16

Conclusion générale

Dans le cadre de la reconnaissance des émotions à partir des expressions faciales, nous nous sommes intéressés dans ce travail à l'évaluation des performances d'un système réalisant cette tâche en termes de deux critères : le taux de reconnaissance et le temps de réponse.

L'approche globale de l'analyse automatique des expressions faciales comprend généralement trois étapes: la détection du visage, l'extraction de caractéristiques et la classification. Durant cette étude, une évaluation de deux méthodes de reconnaissance utilisant deux types de descripteurs différents à savoir un descripteur pour l'extraction de l'information de forme (HOG) et un autre pour l'extraction de l'information de texture (LBP) est présentée. Les résultats obtenus en utilisant deux différentes méthodes d'apprentissage automatique, à savoir le SVM et le KNN pour la reconnaissance de sept expressions faciales (la joie, la tristesse, la surprise, la peur, la colère, le dégoût et la neutralité) montrent que la méthode basée sur le descripteur de l'information de forme (le HOG) est plus efficace en termes de taux de reconnaissance, plus précisément avec le classifieur SVM, en atteignant des taux de reconnaissance de 90.47% pour la base de données JAFFE, et 100.00% pour la base de données MUG. En outre, l'analyse des résultats obtenus en termes de temps de réponse montre que ce temps dépend de deux facteurs, le premier lié directement au nombre de caractéristiques extraites en utilisant le descripteur considéré (l'augmentation du nombre de caractéristiques provoque l'augmentation de temps de réponse du système), tandis que le deuxième facteur dépend du temps nécessaire pour assurer les deux phases d'apprentissage et de tests par le classifieur.

Conclusion générale

Par conséquent, on constate que le succès d'un système de reconnaissance des expressions faciales dépend de succès de chaque phase de conception, en particulier la phase d'extraction de caractéristique qui se considère le cœur du system (une bonne représentation des changements faciaux causés par une expression émotionnelle est la clé de la réussite de reconnaissance). Ainsi, un bon choix de méthodes est crucial pour l'amélioration des Performances de ces systèmes.

Bibliographie

- [1] : Stephen Robbins, Timothy Judge et Véronique Tran. Comportements organisationnels. Pearson Education, 2011. ISBN : 2744074845 (cf. p. 7).
- [2] : E. Couzon et F. Dorn : Les émotions : développer son intelligence émotionnelle. Issyles-Moulineaux : ESF éditeur, 2009.
- [3] : R. Damasio : Le sentiment même de soi, corps, émotions, conscience. Editions Odile Jacob, Paris, 1999.
- [4]: P. Ekman. Are there basic emotions? *Psychological Review*, 99 :550–553, 1992.
- [5] ZUFFEREY, Cyril. EMOLIN–EMOtional Learning INterations robot-student. 2019. Thèse de doctorat. Haute Ecole de Gestion & Tourisme.
- [6] GHARSALLI, Sonia. Reconnaissance des émotions par traitement d’images. 2016. Thèse de doctorat. Université d’Orléans.
- [7]: J. William. What is emotion? *Mind*, 9 :188–205, 1884.
- [8]: W. B. Cannon. The James Lange theory of emotion: A critical examination and an alternative
- [9] ABDAT, Faiza. Reconnaissance automatique des émotions par données multimodales: expressions faciales et signaux physiologiques. Université de Metz, France, 2010.
- [10] : M. Paleari, R. Benmokhtar et B. Huet: Evidence theory-based multimodal emotion recognition. Springer-Verlag Berlin Heidelberg 2009, (Eds.): MMM 2009, LNCS 5371, 2009.
- [11] OUAKKAF, Hassiba et BERKANE, M. Reconnaissance automatique des expressions faciales par support vector machine. 2017.theory. *American Journal of Psychology*, 39 :10–124, 1927.
-

Bibliographie

- [12]: A. Mehrabian: Communication without words. *Psychology Today*, 2.4:53–56, 1968.
- [13]: Ouakkaf, H., & Berkane, M. (2017). Reconnaissance automatique des expressions faciales par support vector machine.
- [14] : P. Philippot. *Emotion et psychothérapie*. Mardaga, 2007.
- [15] : LEROY, Bertrand, CHOUAKRIA, Ahlame, HERLIN, Isabelle, et al. Approche géométrique et classification pour la reconnaissance de visage. In : *Congrès de Reconnaissance des Formes et Intelligence Artificielle*. 1996. p. 548-557.
- [16] : DERBAIX, Christian et PHAM, Michel T. Pour un développement des mesures de l'affectif en marketing: synthèse des prérequis. *Recherche et Applications en Marketing (French Edition)*, 1989, vol. 4, no 4, p. 71-87.
- [17]: Standard mpeg-4: ISO/IEC 14496-2, Information Technology – Coding of Audio-Visual Objects, 2001.
- [18]: A. M. Tekalp and J. Ostermann. Face and 2-d mesh animation in mpeg-4. *Image Communication Journal*, pages 387–421, 1999.
- [19] : Franck Davoine, Bouchra Abboud et Van Mô Dang, Analyse de visages et d'expressions faciales par modèle actif d'apparence, Université de Technologie de Compiègne février 2004.
- [20] : Gharsalli, S. (2016). Reconnaissance des émotions par traitement d'images (Doctoral dissertation, Université d'Orléans).
- [21]: Y. L. Tian, T.C. Kanade, and Jeffrey F., editors. *Facial expression analysis*. Springer-Verlag, 2004.
- [22] Reconnaissance de visages, suivi de trajectoire, multiples représentations de visage, adaptation de domaine, unique image de référence, contrôle d'accès par vidéosurveillance
- [23] Ekman, P. and Friesen, W. (1978a). *Facial action coding system: a technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists.
- [24] Tim Cootes. "An Introduction to Active Shape Models". In: () (cf. p. 16).
-

Bibliographie

- [25] T.F. Cootes, C.J. Taylor, D.H. Cooper et J. Graham. “Active Shape Models - Their Training and Application”. In: *Computer and Image Understanding* Vol. 61, No 1 January, pp 38-59 (1995) (cf. p. 16, 35, 38).
- [26] T.F. Cootes, G.J. Edwards et C.J. Taylor. “Active Appearance Models”. In: (1998) (cf. p. 16).
- [27] CHETTOUH, HADJER. Montage d’un Système de Reconnaissance des Expressions Faciales avec le Deep Learning.
- [28] Laurenz Wiskott, Jean-Marc Fellous, Norbert Kruger et Christoph von der Malsburg. “Face Recognition by Elastic Bunch Graph Matching”. In: *Intelligent Biometric Techniques in Fingerprint and Face Recognition* (1999) (cf. p. 16)
- [29] Chang, Y., Hu, C., Feris, R., and Turk, M. (2006). Manifold based analysis of facial expression. *Image and Vision Computing*, 24(6) :605–614
- [30] Cohen, I., Sebe, N., Garg, A., Chen, L. S., and Huang, T. S. (2003a). Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and image understanding*, 91(1-2) :160–187.
- [31] Kaltwang, S., Rudovic, O., and Pantic, M. (2012). Continuous pain intensity estimation from facial expressions. In *International Symposium on Visual Computing*, pages 368–377. Springer.
- [32] Kapoor, A., Qi, Y., and Picard, R. W. (2003). Fully automatic upper facial action recognition. In *Analysis and Modeling of Faces and Gestures, 2003.AMFG 2003. IEEE International Workshop on*, pages 195–202. IEEE.
- [33] Kotsia, I. and Pitas, I. (2007). Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE transactions on image processing*, 16(1) :172–187.
- [34] Lei, G., Li, X.-h., Zhou, J.-l., and Gong, X.-g. (2009). Geometric feature based facial expression recognition using multiclass support vector machines. In *Granular Computing, 2009, GRC’09. IEEE International Conference on*, pages 318–321. IEEE.
- [35] Lucey, P., Cohn, J. F., Prkachin, K. M., Solomon, P. E., and Matthews, I. (2011). Painful data: The unbc-mcmaster shoulder pain expression archive database. In
-

Bibliographie

Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on, pages 57–64. IEEE.

[36] Matthews, I. and Baker, S. (2004). Active appearance models revisited. *International journal of computer vision*, 60(2) :135–164.

[37] Pantic, M. (2005). Face for interface. *The Encyclopedia of Multimedia Technology and Networking*, 1 :308–314.

[38] Pantic, M. and Patras, I. (2006). Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 36(2) :433–449.

[39] Pantic, M. and Rothkrantz, L. J. (2004). Facial action recognition for facial expression analysis from static face images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(3) :1449–1461.

[40] Shbib, R. and Zhou, S. (2015). Facial expression analysis using active shape model. *Int. J. Signal Process. Image Process. Pattern Recognit*, 8(1) :9–22.

[41] Tian, Y.-I., Kanade, T., and Cohn, J. F. (2001). Recognizing action units for facial expression analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 23(2) :97–115.

[42] Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Lalanne, D., Torres Torres, M., Scherer, S., Stratou, G., Cowie, R., and Pantic, M. (2016). Avec 2016: Depression, mood, and emotion recognition workshop and challenge. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*, pages 3–10. ACM.

[43] Valstar, M. and Pantic, M. (2006). Fully automatic facial action unit detection and temporal analysis. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pages 149–149. IEEE.

[44] Valstar, M. F., Almaev, T., Girard, J. M., McKeown, G., Mehu, M., Yin, L., Pantic, M., and Cohn, J. F. (2015). Fera 2015-second facial expression recognition and analysis challenge. In *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, volume 6, pages 1–8. IEEE.

Bibliographie

- [45] Valstar, M. F. and Pantic, M. (2012). Fully automatic recognition of the temporal phases of facial actions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(1) :28–43.
- [46] Vukadinovic, D. and Pantic, M. (2005). Fully automatic facial feature point detection using gabor feature based boosted classifiers. In *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, volume 2, pages 1692–1698. IEEE.
- [47] Xiong, X. and De la Torre, F. (2013). Supervised descent method and its applications to face alignment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 532–539.
- [48] Zhang, Z., Lyons, M., Schuster, M., and Akamatsu, S. (1998). Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 454–459. IEEE.
- [49] adaptation du détecteur de harris pour l’indexation de textures. Jérôme da rugna, Hubertkonik
- [50] Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M. Martinez, and Seth D. Pollak, “Emotional Expressions [Reconsidered]: Challenges to Inferring Emotion from Human Facial Movements”. *Psychological Science in the Public Interest*, July 2019, Vol. 20(1) 1-68.
- [51]: Bouzit, D. (2019). Reconnaissance de visage basée sur une approche triangulaire.
- [52]: Aghiles, G., & Samira, S. (2014). Contribution de l’approche LBP à la classification des images multispectrales. Application aux données TM de la région du Hoggar (Doctoral dissertation, Université Mouloud Mammeri).
- [53] https://www.researchgate.net/figure/Construction-dun-motif-binaire-et-calcul-du-code-LBP_fig2_50371611
- [54] : Lekdioui, K. (2018). Reconnaissance d’états émotionnels par analyse visuelle du visage et apprentissage machine (Doctoral dissertation, Université Bourgogne Franche-Comté ; Université Ibn Tofail. Faculté des sciences de Kénitra).
-

Bibliographie

- [55] T. Mäenpää, T. Ojala, M. Pietikäinen, and S. Maricor. Robust texture classification by subsets of local binary patterns. In Pattern Recognition, International Conference on, vol 3, pp. 3947, 2000.
- [56] Ouamane, A. (2015). Reconnaissance Biométrique par Fusion Multimodale du Visage 2D et 3D (Doctoral dissertation, Université Mohamed Khider-Biskra).
- [57] Hafiane, A., Seetharaman, G., & Zavidovique, B. (2007, August). Median binary pattern for textures classification. In International Conference Image Analysis and Recognition (pp. 387-398). Springer, Berlin, Heidelberg.
- [58] : Eddine, C. A. (2019). Identification Biométriques des Personnes par les Empreintes d'Articulation du Doigt. Mémoire de master, Université Badji Mokhtar Annaba.
- [59] : Bouafia, N. (2020). Classification efficace des vêtements de mode basée sur les approches : apprentissage automatique ML et apprentissage profond DL (Doctoral dissertation, FACULTE MATHEMATIQUES ET INFORMATIQUE-DEPARTEMENT INFORMATIQUE-OPTION : Informatique Décisionnelle Et Optimisation).
- [60] : Ouakkaf, H., & Berkane, M. (2017). Reconnaissance automatique des expressions faciales par support vector machine.
- [61] : Belahcen, M. (2013). Authentification et identification en biométrie (Doctoral dissertation, Université Mohamed Khider Biskra).
- [62] : Mekkani, S., & Berkane, M. (2014). Reconnaissance de visage.
- [63] : Rami, H. (2016). Contribution au développement de descripteurs LBP pour le suivi d'objet en mouvement.
- [64]: Bensoltana, W., Porebski, A., Vandenbroucke, N., Ahmad, A., & Hamad, D. (2014, November). Contribution des descripteurs de texture LBP à la classification d'images de dentelles. In 17ème édition du colloque CORESA (Compression et Représentation des Signaux Audiovisuels), Image, Groupement de recherche Informatique géométrique et graphique, réalité virtuelle et visualisation (GdR IG-RV).
-

Bibliographie

- [65] : Djebbar, S., & Houari, K. (2011). Système de vérification de l'identité de personne par reconnaissance de l'iris.
- [66] Ahmed, F., Hossain, E., Bari, A. H., & Shihavuddin, A. S. M. (2011, November). Compound local binary pattern (CLBP) for robust facial expression recognition. In 2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI) (pp. 391-395). IEEE.
- [67]: Ahmed, F., Bari, H., & Hossain, E. (2014). Person-independent facial expression recognition based on compound local binary pattern (CLBP). *Int. Arab J. Inf. Technol.*, 11(2), 195-203.
- [68]: Ahmed, F., Hossain, E., Bari, A. S. M. H., & Hossen, M. S. (2011). Compound local binary pattern (clbp) for rotation invariant texture classification. *International Journal of Computer Applications*, 33(6), 5-10.
- [69] : Josse, J., & Husson, F. (2009). Gestion des données manquantes en analyse en composantes principales. *Journal de la société française de statistique*, 150(2), 28-51.
- [70] : Mishra, S. P., Sarkar, U., Taraphder, S., Datta, S., Swain, D., Saikhom, R., ... & Laishram, M. (2017). Multivariate statistical data analysis-principal component analysis (PCA). *International Journal of Livestock Research*, 7(5), 60-78.
- [71] Tan, X. and Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, 19(6) :1635– 1650.
- [72] Yang, M. H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on pattern analysis and machine intelligence*, 24(1), 34-58.
- [73] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), volume 1, pages 886–893. IEEE.
- [74]: Freitas, P. G., Akamine, W. Y., & Farias, M. C. (2016, June). No-reference image quality assessment based on statistics of local ternary pattern. In 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX) (pp. 1-6). IEEE.
-

Bibliographie

- [75]: Khawla, B. E. L. L. O. U. A. D., & Abdallah, O. U. H. A. N. N. A. C. H. E. (2017). La mise en place d'un système biométrique de reconnaissance faciale (Doctoral dissertation, Abdelhafid Boussouf University centre-Mila).
- [76] : Akhloufi, M. A. (2013). Reconnaissance des visages par imagerie multispectrale : du visible à l'infrarouge lointain.
- [77] : Nabila, M. E. R. A. M. R. I. A. (2016). Reconnaissance de visages par Analyse Discriminante Linéaire (LDA). Mémoire de master, Université Badji Mokhtar Annaba.
- [78] : Chaari, A., Lelandais, S., SAAIDIA, M., Vigneron, V., & Ahmed, M. B. (2008, May). Comparaison d'approches pour la détection et l'identification de visages 2D. In 14ème Colloque National sur la Recherche en IUT 2008.
- [79] : Lienou, M. L. (2009). Apprentissage automatique des classes d'occupation du sol et représentation en mots visuels des images satellitaires (Doctoral dissertation, Télécom ParisTech).
- [80]: Guillaumin, M., Verbeek, J., & Schmid, C. (2010, June). Multimodal semi-supervised learning for image classification. In 2010 IEEE Computer society conference on computer vision and pattern recognition (pp. 902-909). IEEE.
- [81] : GAFOUR, Y. (2020). Apprentissage automatique pour une classification des images basée sur les descripteurs (Doctoral dissertation).
- [82] : Sublime, J. (2022). L'apprentissage non-supervisé et ses contradictions. 1024 : Bulletin de la Société Informatique de France, (19), 145-156.
- [83] : Chapel, L. (2007). Maintenir la viabilité ou la résilience d'un système : les machines à vecteurs de support pour rompre la malédiction de la dimensionnalité ? (Doctoral dissertation, Université Blaise Pascal-Clermont-Ferrand II).
- [84] : Maalej, A. (2012). Reconnaissance d'Expressions Faciale 3D Basée sur l'Analyse de Forme et l'Apprentissage Automatique (Doctoral dissertation, Université des Sciences et Technologie de Lille-Lille I).
- [85] Lyons, M., Akamatsu, S., Kamachi, M., and Gyoba, J. (1998). Coding facial expressions with gabor wavelets. In Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on, pages 200–205. IEEE.
-

Bibliographie

- [86] Aifanti, N., C. Papachristou, and A. Delopoulos. The MUG facial expression database. in 11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10. 2010. IEEE.