



**Université Sidi Mohammed Ben Abdellah  
Faculté des Sciences Dhar El Mahraz- Fès  
Centre d'Etudes Doctorales  
"Sciences et Technologies"**

**Formation Doctorale : STIC**

**Discipline : Informatique**

**Spécialité : Informatique**

**Laboratoire : LIAN**

**THESE DE DOCTORAT**

Présentée par: Abdellatif DAHMOUNI

**Optimisation de la biométrie faciale via des méthodes locales  
probabilistes, statistiques et structurelles**

Soutenue le 21 Mars 2019 devant le jury composé de :

<b>Pr. Mohamed MEKNASSI</b>	<b>FSDM, Fès</b>	<b>Président</b>
<b>Pr. Rachid OULAD HAJ THAMI</b>	<b>ENSIAS, Rabat</b>	<b>Rapporteur</b>
<b>Pr. Samir MBARKI</b>	<b>FS, Kénitra</b>	<b>Rapporteur</b>
<b>Pr. Ali YAHYAOUY</b>	<b>FSDM, Fès</b>	<b>Rapporteur</b>
<b>Pr. Hamid TAIRI</b>	<b>FSDM, Fès</b>	<b>Examineur</b>
<b>Pr. Jaouad BOUMHIDI</b>	<b>FSDM, Fès</b>	<b>Examineur</b>
<b>Pr. Mohammed Chakib SOSSE ALAOUI</b>	<b>CRMEF, Fès</b>	<b>Examineur</b>
<b>Pr. Khalid SATORI</b>	<b>FSDM, Fès</b>	<b>Directeur de thèse</b>
<b>Pr. Karim EL MOUTAOUAKIL</b>	<b>ENSA, Al-Hoceima</b>	<b>Co-directeur de thèse</b>

Année universitaire : 2018-2019

# Dédicace

---

*À ma mère.*

*À l'âme de mon père,*

*À toute ma famille.*

*À tous mes professeurs.*

*À tous mes collègues.*

## Remerciements

---

Mes remerciements et ma profonde gratitude sans limites pour **Dieu**, pour sa clémence et son aide, c'est grâce à lui que j'ai pu achever ce travail.

Mes remerciements et ma profonde gratitude sans limites pour notre **Prophète Mohammed**, pour ses recommandations et ses conseils.

Je tiens essentiellement et particulièrement à exprimer ma sincère et profonde gratitude à mon directeur de thèse : **Mr. Khalid Satori**, Professeur à la Faculté des Sciences Dhar El Mahraz de Fès. Sa confiance, ses conseils permanents et ses grandes qualités humaines étaient les points clés d'aboutissement de ce travail.

Je tiens à témoigner ma profonde gratitude à **Mr. Rachid OULAD HAJ THAMI**, Professeur à École nationale supérieure d'informatique et d'analyse des systèmes de Rabat, à **Mr. Samir MBARKI**, Professeur à la Faculté des Sciences de Kenitra, et à **Mr. Ali YAHYAOUY**, Professeur à la Faculté des Sciences de Dhar El Mahraz de Fès, qui m'ont fait l'honneur d'être rapporteurs de ma thèse. Leurs commentaires et leurs conseils m'ont notamment permis d'améliorer ce manuscrit.

Je tiens à exprimer ma haute considération à **Mr. Mohamed MEKNASSI**, Professeur à la Faculté des Sciences de Dhar El Mahraz de Fès, pour l'honneur qu'il m'a fait en présidant mon jury de thèse. Mes remerciements s'adressent également à **Mr. Hamid TAIRI**, Professeur à la Faculté des Sciences de Dhar El Mahraz de Fès, à **Mr. Jaouad BOUMHIDI**, Professeur à la Faculté des Sciences de Dhar El Mahraz de Fès, à **Mr. Mohammed Chakib SOSSE ALAOUI**, Professeur au Centre Régional des Métiers de l'Education et de la Formation de Fès, pour le grand honneur qu'ils m'ont fait en acceptant de participer à mon jury de thèse.

Je tiens à exprimer ma profonde reconnaissance et ma haute considération à mon co-directeur de thèse: **Mr. Karim EL MOUTAOUAKIL**, Professeur à professeur à l'École Nationale des Sciences Appliquées à Al-Houcima. Son aménité, sa patience et son esprit d'équipe m'ont vraiment facilité ce parcours doctoral.

Je remercie également tous les membres de notre équipe de recherche « LIIAN », Professeurs et Doctorants, pour leur esprit de groupe.

Mes profonds remerciements sont adressés à ma chère mère, à ma chère femme, à mes enfants, à mes frères, à mes sœurs et tous mes amis qui n'ont jamais cessé de m'encourager.

---

La description des visages via les méthodes locales reste l'outil le plus recouru pour générer des représentations faciales réduites et discriminantes. Parmi les modèles locaux les plus utilisés sont ceux basés sur le modèle binaire local (Local Binary Pattern, LBP). Bien qu'ils soient simples et moins complexes, les pertes d'informations dues principalement à leur aspect binaire est l'une des handicaps majeure de ces modèles. Dans le présent travail, nous avons remédié à ce problème en développant quatre nouveaux descripteurs basés sur le concept de la localité de LBP, l'estimation par intervalle de confiance, l'effet du gradient, la dualité « pixel -charge électrique virtuelle » et enfin l'optimisation primale-duale. Ces modèles sont les suivants :

- Modèle probabiliste binaire local (Local Binary Probabilistic Pattern, LBPP).
- Modèle probabiliste gradient local (Local Gradient Probabilistic Pattern, LGPP).
- Modèle binaire à champs électrique virtuel (Electric Virtual Binary Pattern, EVBP).
- Modèle optimal binaire local (Optimal Local Binary Pattern, Optim-LBP).

Ainsi, LBPP génère une nouvelle description de l'image, particulièrement celle du visage avec utilisation des intervalles de confiance qui sont déterminés en fonction de la normalité asymptotique des distributions empiriques qui régissent les voisinages des pixels.

LGPP est une extension de LBPP qui engendre une représentation améliorée de l'image, tout en exploitant les avantages de la description par gradient et intervalle de confiance. Dans ce sens, l'intervalle de confiance d'un voisinage des gradients est également déterminé en fonction du degré de normalité de la distribution empirique qui les régit.

EVBP est un descripteur qui adopte la dualité « pixel -charge électrique virtuelle », pour engendrer une nouvelle représentation de l'image où chaque voisinage est assimilé à une grille de charges électriques virtuelles. Ainsi, cette vision permet de décrire la topologie locale de chaque élément de surface réel par l'ensemble des interactions entre les charges virtuelles de la grille qui lui correspond.

Enfin, Optim-LBP qui détermine les poids optimaux des contributions des pixels voisins dans le recodage du pixel courant, en résolvant un problème de minimisation sous contraintes. En outre la fonction objective de ce problème est définie comme mesure du degré d'homogénéité de chaque voisinage en termes d'erreur quadratique de l'entropie de Shannon.

Afin de mettre en évidence les apports et l'efficacité des descripteurs que nous avons proposés, nous les avons combinés avec des algorithmes d'apprentissage statistiques, notamment des méthodes de réduction de la dimension et des algorithmes de classification. La comparaison de nos approches avec des méthodes connexes de l'état d'art a montré la qualité des descriptions engendrées ainsi que la prépondérance des approches proposées.

**Mots Clefs :** Biométrie, Réduction de la dimension, Algorithmes d'apprentissage statistique, Réseaux de neurones, Machine à vecteurs de support, Validation croisée, Recherche profonde, Extraction de caractéristiques, Optimisation, LBPP, LGPP, EVBP et Optim-LBP.

Face description using local methods remains the most used tool to generate reduced and discriminating facial representations. The local patterns are the most used; they are based on the Local Binary Pattern (LBP). Although they are simple and less complex, the information loss due to their binary aspect remains one of the main handicaps of these models. In the present work, we have remedied this problem by developing four new descriptors based on the LBP locality concept, the confidence interval estimation, the gradient effect, the « pixel-'virtual electric charge'» duality and the primal-dual optimization. These models are:

- Local Binary Probabilistic Pattern (LBPP).
- Local Gradient Probabilistic Pattern (LGPP).
- Electric Virtual Binary Pattern (EVBP).
- Optimal Local Binary Pattern (Optim-LBP).

Hence, LBPP generates a new image description, particularly face image, by using confidence intervals determined according to the asymptotic normality of the neighborhoods empirical distributions.

LGPP is an LBPP extension that generates an improved image representation, while exploiting gradient and confidence interval advantages. In this sense, the confidence interval of the gradients' neighborhood is also determined by the normality degree of the empirical gradients' distribution.

EVBP is a descriptor that adopts a new concept called «pixel-' virtual electric charge '» duality, to generate a new image representation. In this descriptor, each neighborhood is assimilated to a grid of virtual electrical charges. Indeed, this vision allows describing the local topology of each real surface element by the set of the virtual charges' interactions of the corresponding grid.

Finally, Optim-LBP determines the optimal contributions' weights of the neighboring pixels by solving a constrained minimization problem. Accordingly, the objective function of this problem is defined as a measure of the homogeneity degree of each neighborhood by using the quadratic error of Shannon's entropy.

In order to highlight the contributions and the effectiveness of the proposed descriptors, we combine them with many machine learning algorithms including the subspace methods and the classification algorithms. The comparison of our approaches with related state-of-art methods showed the quality of the generated descriptions as well as the preponderance of the proposed approaches .

**Key words:** Biometrics, Subspace Methods, Machine-Learning algorithms, Neuronal networks, Machine support vectors, Cross-validation, Deep search, Features extraction, optimization, LBPP, LGPP, EVBP and Optim-LBP.

## *Table des matières*

<b><i>Table des matières</i></b>	<b><i>v</i></b>
<b><i>Liste des tableaux</i></b>	<b><i>ix</i></b>
<b><i>Liste des figures</i></b>	<b><i>xi</i></b>
<b><i>Liste des sigles</i></b>	<b><i>xiii</i></b>
<b><i>Introduction générale</i></b>	<b><i>1</i></b>
<b><i>Chapitre I. Systèmes Biométriques et Biométrie Faciale</i></b>	<b><i>6</i></b>
<b>I. Introduction</b>	<b><i>7</i></b>
<b>II. Systèmes biométriques</b>	<b><i>7</i></b>
II.1. Définition	<i>7</i>
II.2. Conception d'un système biométrique	<i>8</i>
<b>III. Mesure de performance d'un système biométrique</b>	<b><i>10</i></b>
<b>IV. Biométrie faciale</b>	<b><i>11</i></b>
IV.1. Principaux défis de la biométrie faciale	<i>11</i>
IV.2. État de l'art de la détection des visages	<i>13</i>
IV.2.1. Méthodes basées sur les caractéristiques invariantes	<i>13</i>
IV.2.2. Méthodes basées sur l'apparence globale	<i>15</i>
IV.3. État de l'art de la reconnaissance de visages	<i>15</i>
IV.3.1. Approche globale ou Holistique	<i>16</i>
IV.3.2. Approche locale	<i>18</i>
IV.3.3. Approche hybride	<i>20</i>
IV.4. Base de données de visages	<i>20</i>
IV.4.1. Base de données ORL	<i>21</i>
IV.4.2. Base de données Yale	<i>21</i>
IV.4.3. Base de données AR	<i>21</i>
<b>V. Conclusion</b>	<b><i>23</i></b>
<b><i>Chapitre II. Extraction des caractéristiques, réduction de la dimension et classification</i></b>	<b><i>24</i></b>
<b>I. Introduction</b>	<b><i>25</i></b>
<b>II. Méthodes de réduction de la dimension</b>	<b><i>25</i></b>
II.1. Méthodes basées sur PCA	<i>25</i>
II.1.1. Méthode des visages propres	<i>26</i>
II.1.2. Méthode PCA à noyau	<i>28</i>
II.1.3. Méthode PCA à deux dimensions (2DPCA)	<i>29</i>
II.2. Méthodes basées sur LDA	<i>30</i>
II.2.1. Méthode des Fisher-faces	<i>30</i>

II.2.2.	Méthode LDA à noyau	31
II.2.3.	Problème de singularité de LDA	32
II.2.4.	Méthode LDA à deux dimensions (2DLDA)	32
II.3.	Méthodes spectrales	33
II.3.1.	DCT à deux dimensions (2DDCT)	33
II.3.2.	DWT à deux dimensions (2DDWT)	33
<b>III.</b>	<b>Classification</b>	<b>34</b>
III.1.	Classification par mesure de similarité	34
III.2.	Classification par apprentissage artificiel	35
III.2.1.	Réseaux de neurones artificiels	35
III.2.2.	Machine à vecteurs de support (SVM)	37
III.2.3.	Les K-Plus Proches Voisins (KNN)	40
III.2.4.	Les critères de mesure des performances des classificateurs	40
III.2.5.	Combinaison des classificateurs	41
<b>IV.</b>	<b>Extraction des Caractéristiques</b>	<b>42</b>
IV.1.	LBP d'origine et variétés de base	42
IV.1.1.	LBP d'origine	42
IV.1.2.	LBP uniforme et invariant par rotation	42
IV.2.	Nouvelles variétés de LBP	43
<b>V.</b>	<b>Conclusion</b>	<b>46</b>
<b>Chapitre III. Modèle Probabiliste Binaire Local (Local Binary Probabilistic Pattern, LBPP) : Application à la Biométrie Faciale</b>		<b>47</b>
<b>I.</b>	<b>Introduction</b>	<b>48</b>
<b>II.</b>	<b>Test de normalité et moments statistiques</b>	<b>48</b>
II.1.	Moments Statistiques	49
II.1.1.	Tendance Centrale et Dispersion	49
II.1.2.	Asymétrie et Aplatissement	50
II.2.	Test de Normalité de Jarque–Bera (JB-test)	51
<b>III.</b>	<b>Modèle Probabiliste Binaire Local (LBPP)</b>	<b>52</b>
III.1.	Présentation de LBPP	52
III.2.	Évaluation de l'intervalle de confiance	53
III.3.	Formulation de LBPP	54
III.4.	Les avantages de LBPP	55
<b>IV.</b>	<b>Reconnaissance faciale par LBPP</b>	<b>57</b>
IV.1.	Système de reconnaissance faciale proposé	57
IV.1.1.	Prétraitement	57
IV.1.2.	Traitement par LBPP	58
IV.1.3.	Réduction de la dimension	59
IV.1.4.	Classification	60
IV.2.	Résultats expérimentaux	61
IV.2.1.	Valeurs optimales des paramètres	61
IV.2.2.	Evaluation des classificateurs	67

---

IV.3. Comparaison avec les méthodes de l'état de l'art	70
<b>V. Conclusion</b>	<b>71</b>
<b>Chapitre IV. Modèle Probabiliste Gradient Local (Local Gradient Probabilistic Pattern, LGPP) : Application à la Biométrie Faciale</b>	<b>73</b>
<b>I. Introduction</b>	<b>74</b>
<b>II. Modèle Probabiliste Gradient Local (LGPP)</b>	<b>74</b>
II.1. Présentation de LGPP	74
II.2. Évaluation de l'intervalle de confiance	75
II.3. Formulation de LGPP	76
II.4. Avantages de LGPP	77
<b>III. Reconnaissance de visage par LGPP</b>	<b>80</b>
III.1. Système de reconnaissance faciale proposé	80
III.1.1. Prétraitement	81
III.1.2. Traitement par LGPP	81
III.1.3. Réduction de dimension et vecteurs caractéristiques	81
III.1.4. Classification	83
III.2. Les résultats expérimentaux	83
III.2.1. Valeurs optimales des paramètres	83
III.2.2. Evaluation par classification artificielle	87
III.2.3. Combinaison des classificateurs	91
III.3. Comparaison avec d'autres méthodes	93
<b>IV. Conclusion</b>	<b>94</b>
<b>Chapitre V. Modèle Binaire à Champ Électrique Virtuel (Electric Virtual Binary Pattern, EVBP) : Application à la Biométrie Faciale</b>	<b>95</b>
<b>I. Introduction</b>	<b>96</b>
<b>II. Principe du descripteur EVBP</b>	<b>97</b>
II.1. Présentation du descripteur EVBP	97
II.2. Modélisation du descripteur EVBP	97
II.2.1. Modèle champ électrique virtuel	97
II.2.2. Formulation du descripteur EVBP	99
II.2.3. Seuil du descripteur EVBP	100
II.3. Les avantages du descripteur EVBP	102
<b>III. Reconnaissance de visage par EVBP</b>	<b>106</b>
III.1. Le système de reconnaissance faciale proposé	106
III.1.1. Prétraitement	106
III.1.2. Traitement par le descripteur EVBP	106
III.1.1. Vecteurs caractéristiques (eLBPH)	107
III.1.2. Classification	108
III.2. Les résultats expérimentaux	108
III.2.1. Taille optimale des blocs	109
III.2.2. Métrique adéquate	110

---



III.2.3. Effet de l'ensemble d'apprentissage	112
III.2.4. Evaluation par apprentissage artificiel	113
III.2.5. EVBP et méthodes de réduction de la dimension	114
III.3. Comparaison avec les méthodes de l'état de l'art	115
<b>IV. Conclusion</b>	<b>117</b>
<b>Chapitre VI. Modèle Optimal Binaire Local (Optimal Local Binary Pattern, Optim-LBP) : Application à la Biométrie Faciale</b>	<b>118</b>
<b>I. Introduction</b>	<b>119</b>
<b>II. Méthodes d'optimisation</b>	<b>120</b>
II.1. Optimisation d'une fonction régulière	120
II.1.1. Optimisation sans contraintes	120
II.1.2. Optimisation avec contraintes	122
II.2. Les méta-heuristiques	123
<b>III. Modèle optimal binaire local (Optim-LBP)</b>	<b>124</b>
III.1. Modélisation du descripteur Optim-LBP	124
III.1.1. Présentation	124
III.1.2. Formulation de Optim-LBP	125
III.1.3. Analyse et discussion	127
III.2. Reconnaissance de visage par Optim-LBP	129
III.2.1. Présentation du système	129
III.2.1. Expérimentations	130
III.2.2. Comparaison avec les travaux antérieurs	133
<b>IV. Conclusion</b>	<b>134</b>
<b>Conclusion générale et perspectives</b>	<b>135</b>
<b>Bibliographie de l'auteur</b>	<b>139</b>
<b>Références</b>	<b>141</b>

## *Liste des tableaux*

---

---

Tableau I-1: Comparaison de différentes modalités biométriques.....	8
Tableau I-2: Aperçu des principales bases de données de reconnaissance du visage.....	22
Tableau II-1: Matrice de confusion à deux classes.....	40
Tableau II-2: Aperçu des principales variétés de LBP.....	45
Tableau III-1: Valeurs critiques $T_{JB}$ pour différentes tailles $n$ et différents risque $\alpha$ .....	52
Tableau III-2 : Dimension optimale de la base ORL.....	62
Tableau III-3 : Dimension optimale de la base Yale.....	63
Tableau III-4 : Dimension optimale de la base AR.....	63
Tableau III-5 : Intervalle de confiance optimal de la base ORL.....	64
Tableau III-6 : Intervalle de confiance optimal de la base Yale.....	64
Tableau III-7 : Intervalle de confiance optimal de la base AR.....	65
Tableau III-8 : Paramètres optimaux de 2DDCT pour ORL.....	66
Tableau III-9 : Paramètres optimaux de 2DDCT pour Yale.....	66
Tableau III-10 : Paramètres optimaux de 2DDCT pour AR.....	67
Tableau III-11 : Taille optimale des vecteurs descripteurs.....	67
Tableau III-12 : Comparaison de différents classificateurs pour ORL.....	68
Tableau III-13 : Comparaison des différents modes de vote sur ORL.....	68
Tableau III-14 : Comparaison de différents classificateurs pour Yale.....	69
Tableau III-15 : Comparaison des différents modes de vote sur Yale.....	69
Tableau III-16 : Comparaison de différents classificateurs sur AR.....	70
Tableau III-17 : Comparaison de différents modes de vote sur AR.....	70
Tableau III-18 : Comparaison avec des méthodes de l'état de l'art.....	71
Tableau IV-1 : Dimension optimale pour ORL.....	84
Tableau IV-2 : Effet de 2DDWTsur l'efficacité du système de reconnaissance.....	85
Tableau IV-3 : Intervalle de confiance optimal.....	86
Tableau IV-4 : Résultats du classificateur LIBSVM.....	87
Tableau IV-5 : Résultats du classificateur SMO.....	88
Tableau IV-6 : Résultats des classificateurs F-MLP.....	88
Tableau IV-7 : Résultats des classificateurs CS-MLP.....	89
Tableau IV-8 : Résultats de la fusion des SVMs par règle de décision AP.....	92
Tableau IV-9 : Résultats de la fusion des SVMs par règle de décision MV.....	92
Tableau IV-10 : Comparaison avec des méthodes de l'état de l'art.....	93
Tableau V-1 : Taille du vecteur eLBPH pour des variétés de LBP (56 blocs).....	106
Tableau V-2 : Taille optimal des blocs.....	110
Tableau V-3 : Métrique adéquate pour ORL.....	111

---

---

Tableau V-4 : Métrique adéquate pour Yale .....	111
Tableau V-5 : Métrique adéquate pour AR.....	112
Tableau V-6 : Résultats de classification parKNN ( $L^1$ ).....	114
Tableau V-7 : Combinaison d'EVBP et les méthodes de réduction de la dimension.....	115
Tableau V-8 : Comparaison avec des méthodes de l'état de l'art .....	116
Tableau VI-1 : Dimensions optimales des bases de données ORL, Yale et AR .....	131
Tableau VI-2 : Classification artificielle de la base de données ORL.....	132
Tableau VI-3 : Classification artificielle de la base de données Yale .....	132
Tableau VI-4 : Classification artificielle de la base de données AR.....	133
Tableau VI-5 : Comparaison avec des méthodes de l'état de l'art .....	133

## Liste des figures

---

---

Figure I.1: Différentes Modalités Biométriques .....	7
Figure I.2 : Architecture Générique d'un Système Biométrique.....	9
Figure I.3 : Exemples de changements d'illumination. ....	12
Figure I.4 : Exemples de variations de pose.....	12
Figure I.5 : Exemples de variations des expressions faciales.....	12
Figure I.6 : Exemples d'occlusions partielles.....	13
Figure I.7: Exemple de méthodes de reconnaissance de visages.....	16
Figure I.8 : Exemples des images extraites de la base ORL.....	21
Figure I.9 : Exemples des images extraites de la base Yale .....	21
Figure I.10 : Exemples des images extraites de la base AR.....	22
Figure II.1: Reconstruction des images par 2DPCA et PCA .....	29
Figure II.2: Représentation du réseau MLP .....	36
Figure II.3: Description de l'hyperplan optimal de SVM .....	38
Figure II.4: Les 58 modèles uniformes de LBP <sub>8,1</sub> .....	43
Figure III.1 : Différents intervalles de confiance de la loi normale .....	49
Figure III.2 : Skewness et Kurtosis de quelques distributions.....	51
Figure III.3 : Histogrammes des zones de l'image LBPP .....	53
Figure III.4 : Description d'un pixel par la méthode LBPP .....	55
Figure III.5 : Comparaison entre LBP et LBPP pour des changements d'éclairage .....	56
Figure III.6 : Résultats LBPP(k=4) avec changements d'éclairage et bruits .....	57
Figure III.7 : Diagramme blocs du système de reconnaissance .....	58
Figure III.8 : Images LBPP(k = 1,2,3,4) des bases de données ORL, Yale et AR .....	58
Figure III.9 : Différentes étapes de la méthode 2DDCT .....	60
Figure III.10 : Description de différentes zones de visage par 2DDCT.....	60
Figure IV.1 : Histogrammes des zones de l'image LGPP .....	75
Figure IV.2 : Description d'un pixel par la méthode LGPP .....	77
Figure IV.3 : Images obtenues par LGPP pour différents intervalles de confiance. ....	78
Figure IV.4 : Comportements de LGPP vis-à-vis différents d'éclairages.....	78
Figure IV.5 : Comportements de LGPP vis-à-vis différents changements et bruits.....	79
Figure IV-6 : Domaines d'application de LGPP.....	79
Figure IV-7 : Diagramme blocs du système de reconnaissance proposé .....	80
Figure IV.8 : Images LGPP(k= 4) pour ORL, Yale et AR. ....	82
Figure IV.9 : Sub-bandes de décomposition au premier niveau de 2DDWT. ....	82
Figure IV.10 : Les courbes ROC relatives aux MLPs .....	90
Figure IV.11 : Les courbes ROC relatives aux SVMs.....	91

---

---

Figure V-1 : Analogie entre pixel et charge électrique virtuelle .....	98
Figure V.2 : Exemple de codage du descripteur EVBP .....	102
Figure V.3 : Comportement d'EVBP vis-à-vis les changements d'illumination .....	103
Figure V.4 : Comportement d'EVBP vis-à-vis des bruits Gaussien et flou .....	104
Figure V-5 : Autres perspectives d'application d'EVBP .....	105
Figure V.6 : Diagramme blocs du système de reconnaissance proposé.....	107
Figure V.7 : Exemples d'images obtenues par EVBP pour ORL, Yale et AR. ....	107
Figure V.8 : Les étapes d'extraction du vecteur caractéristique eLBPH. ....	108
Figure V.9 : Taux de reconnaissance de différents couples ( $n_{tr}$ , $n_{ts}$ ).....	113
Figure VI.1 : Les méta-heuristiques les plus connues pour l'optimisation .....	123
Figure VI.2 : Effet de l'aspect binaire de la fonction de décision.....	125
Figure VI-3 : Exemples de comportement d'Optim-LBP via différents voisinages .....	128
Figure VI-4 : Comportement de Optim-LBP vis-à-vis changement d'illumination.....	128
Figure VI-5: Exemples d'histogrammes d'Optim-LBP .....	129
Figure VI.6 : Images Optim-LBP pour ORL, Yale et AR. ....	130

*Liste des sigles*

---

---

<b>ICA</b>	Analyse en composantes indépendantes
<b>LPP</b>	Préservation de la localisation de la projection
<b>JPEG</b>	Groupe mixte d'experts en photographie
<b>PCA</b>	Analyse en composante principale
<b>KPCA</b>	Analyse en composante principale à noyaux
<b>2DPCA</b>	Analyse en composante principale à deux dimensions
<b>LDA</b>	Analyse linéaire discriminante
<b>KLDA</b>	Analyse linéaire discriminante à noyaux
<b>2DLDA</b>	Analyse linéaire discriminante à deux dimensions
<b>DLDA</b>	Analyse linéaire discriminante direct
<b>DCT</b>	Transformée en cosinus diserte
<b>2DDCT</b>	Transformée en cosinus diserte à deux dimensions
<b>DWT</b>	Transformée en ondelettes diserte
<b>2DDWT</b>	Transformée en ondelettes diserte à deux dimensions
<b>RBF</b>	Fonction de base radiale
<b>MLP</b>	Réseaux de neurones multicouches
<b>F-MLP</b>	MLP à apprentissage par Back-propagation
<b>CS</b>	Algorithme de recherche des oiseaux
<b>CS-MLP</b>	MLP à apprentissage par la recherche des oiseaux
<b>SVM</b>	Machine à vecteurs de support
<b>LIBSVM</b>	Bibliothèques des méthodes basées SVM
<b>SMO</b>	Optimisation Minimale Séquentielle
<b>LBP</b>	Modèle binaire local
<b>LGP</b>	Modèle local gradient
<b>LBPP</b>	Modèle probabiliste binaire local
<b>LGPP</b>	Modèle probabiliste gradient local
<b>EVBP</b>	Modèle binaire à champ électrique virtuel
<b>Optim-LBP</b>	Modèle optimal binaire local
<b>eLBPH</b>	Histogramme LBP spatialement amélioré
<b>CMC</b>	Caractéristique du match cumulatif
<b>ROC</b>	Courbe caractéristique d'exploitation du récepteur
<b>HOG</b>	Histogramme de gradient orienté
<b>POEM</b>	Modèle de magnitudes de bord orientées
<b>EBGM</b>	Graphe élastique de correspondance graphique
<b>MMC</b>	Modèles de Markov cachés
<b>BD</b>	Base de données

---

---

## Introduction générale

---

---

La dernière décennie a été marquée par plusieurs événements dramatiques, tels que des attaques terroristes à grande échelle et des exodes massifs des zones de conflit. À cela s'ajoute la pénétration des systèmes d'informations les plus sophistiqués et les opérations récurrentes de fraude et de criminalité. La couverture médiatique des informations sur ces actes a probablement contribué au sentiment croissant d'insécurité de la population mondiale. La conséquence immédiate de ces événements est la révolution scientifique et technique du domaine de la biométrie. En effet, cette dernière a mis en disposition des systèmes biométriques pratiques, permanents et difficilement falsifiables pour la vérification ou l'identification automatique des personnes. Ces systèmes utilisent des caractéristiques intrinsèques, comportementales ou physiologiques, appelées modalités biométriques notamment la voix, la signature, l'iris, l'empreinte digitale et le visage [1-3].

Les applications potentielles utilisant la biométrie sont nombreuses et peuvent être regroupées en trois catégories : des applications de sécurité de l'information (systèmes d'accès logique), des applications de contrôle d'accès et de vidéosurveillance (systèmes d'accès physique) et des applications officielles couvrant principalement la sécurité des territoires, du transport, et des secteurs publics [4-6]. Le choix d'une modalité biométrique repose principalement sur les objectifs de l'application mise en œuvre et la nature des utilisateurs ciblés. A cet égard, le degré *d'intrusivité*, mesurant l'interaction *utilisateurs-capteur*, reste un critère primordial qui influence fortement le choix de la modalité adéquate lors de la mise en place d'un système biométrique. Cependant, la majorité des modalités biométriques sont intrusives, à l'image l'empreinte digitale et l'iris qui nécessite la coopération de l'utilisateur.

En outre, grâce aux travaux de recherches investis dans le domaine de l'extraction des primitives et à la performance des modèles de *Machine Learning* notamment en reconnaissance faciale, la biométrie est devenue de plus en plus automatique et temps réel permettant la gestion des flux de données massifs [7-8]. La reconnaissance faciale s'est initiée dès les années 60 à l'époque où les approches géométriques semi-automatiques étaient les seules outils développées, par inspiration des empreintes digitales, qui se basent sur les distances calculées entre des points caractéristiques du visage [9]. Cette vision restrictive a montré ses limites vis-à-vis les différentes variabilités liées à l'apparence globale du visage et

aux différents changements d'illumination, des expressions faciales, de pose et etc., c'est le dilemme « interclasses et intra-classes ». Ce problème multicritères np-complet a fait l'objet de plusieurs recherches réparties en trois approches: globale, locale et hybride. Les méthodes globales sont des techniques de réduction de la dimension basées sur des processus d'analyse statistique ou spectrale. Parmi ces méthodes nous citons l'analyse en composantes principales (PCA) [10], l'analyse linéaire discriminante (LDA) [11], la transformées en cosinus discrètes (DCT) [12] et la transformée en ondelettes discrètes (DWT) [13]. Malgré leurs succès, ces méthodes restent très sensibles aux changements d'éclairage d'illumination, de pose et d'expression faciale. Afin de remédier à ces défis, les méthodes locales utilisent nos connaissances a priori sur la morphologie du visage pour extraire des caractéristiques pertinentes de celui-ci. Les ondelettes de Gabor [14], l'histogramme à gradient orienté (HOG) [15] et le modèle binaire local (LBP) [16] sont les plus connues de cette approche. Enfin, les méthodes de la dernière approche visent augmenter la performance de reconnaissance en combinant, d'une manière complémentaire, les méthodes des deux approches précédentes. Parmi ces méthodes hybrides sont: Log-PCA [17].et EBGM [18].

Nos travaux de recherche s'inscrivent dans le cadre des méthodes locales utilisant le LBP d'Ojala pour engendrer de nouvelles descriptions du visage. En effet, LBP a été adapté, pour la première fois, à la reconnaissance faciale par Ahonen qui a décrit le visage selon trois niveaux : motifs LBP (pixels), histogrammes locaux (régions) et histogramme global (image) [16]. Suite à son succès, plusieurs variétés de LBP ont été proposées dont les principales sont regroupées dans les deux livres [19-20]. Cependant, le caractère binaire de cette famille de descripteurs cause de significatives pertes d'information qui persiste, davantage, avec la taille des voisinages adoptés autour du pixel courant. Pour remédier à ce problème, nous avons élaboré quatre nouvelles variétés de LBP en se basant sur des avancées théoriques en probabilité, en statistique et en optimisation. Il s'agit, en effet, du modèle probabiliste binaire local (Local Binary Probabilistic Pattern, LBPP), du modèle probabiliste gradient local (Local Gradient Probabilistic Pattern, LGPP), du modèle binaire à champs électrique virtuel (Electric Virtual Binary Pattern, EVBP), et du modèle optimal binaire local (Optimal Local Binary Pattern, Optim-LBP).

Ce mémoire de thèse présente des contributions variées aux problèmes de la reconnaissance faciale, en particulier nos quatre descripteurs *LBPP*, *LGPP*, *EVBP* et *Optim-LBP*. Dans notre

---



première contribution, nous avons mis en évidence le descripteur LBPP qui consiste à générer une nouvelle description de l'image en utilisant des résultats importants de la théorie des probabilités notamment la loi des grands nombres et le théorème central limite. Dans ce sens, l'étude de la convergence en loi de la distribution empirique qui gouverne chaque voisinage de l'image vers une loi normale permet d'estimer l'intervalle de confiance qui régit d'une manière probabiliste les contributions des pixels voisins lors du recodage du pixel courant. Un résultat marquant de cette description réside dans son pouvoir de localisation des zones de fortes variations du gradient caractérisant, en particulier, les principaux traits d'un visage. Dans un premier temps nous nous sommes focalisés sur la formulation de LBPP ainsi que son utilité dans la description du visage. Ensuite, nous avons mis en évidence son comportement vis-à-vis des changements d'illumination et du bruit. Enfin, nous l'avons utilisé pour concevoir un système de reconnaissance faciale robuste en termes de taux de reconnaissance et temps d'exécution. Dans ce contexte, nous l'avons hybridé avec des méthodes de réduction de la dimension, statistiques et spectrales, ensuite nous avons amélioré l'efficacité de classification en faisant appel aux algorithmes de *Machine Learning*.

Dans notre deuxième contribution, nous avons visé améliorer l'extraction des caractéristiques faciales en mettons en œuvre notre descripteur LGPP qui utilise des voisinages formés par des gradients. Il consiste à déterminer l'intervalle de confiance qui régit chaque voisinage en étudiant la distribution empirique qui gouverne ses gradients..Nous avons commencé par la formulation du descripteur LGPP, ensuite nous avons mis en évidence l'effet de la taille de l'intervalle de confiance sur la description faciale et le comportement de ce descripteur vis-à-vis les différents changements affectant l'image. Enfin, nous avons intégré la description faciale engendrée par LGPP dans notre système de reconnaissance. Nous l'avons hybridé dans un premier temps avec des méthodes de réduction statistiques. Ensuite, nous avons amélioré les performances globales de notre système en utilisant les ondelettes de Haar et les algorithmes de *Machine Learning*.

Dans notre troisième contribution, nous avons utilisé la dualité « pixel-charge électrique virtuelle » pour concevoir notre descripteur EVBP. Cette robuste variété de LBP encode la structure locale de chaque élément de la surface réelle en assimilant son voisinage, correspond, au niveau de l'image à une grille élémentaire de charges électriques virtuelles. Ce mode de codage utilise l'interaction entre chaque couple de charges situées suivant chacune

des quatre directions usuelles (verticale, horizontale et diagonales) pour réduire le nombre d'étiquettes d'histogramme de 93.75%. D'abord, nous avons présenté les étapes formulant ce descripteur, puis nous avons mis en évidence ses avantages et ses domaines d'application. Ensuite, nous l'avons intégré dans notre système de reconnaissance. Dans ce contexte, nous l'avons testé avec différentes métriques de mesure de similarité entre histogrammes. Ensuite, nous avons montré que son hybridation avec des méthodes de réduction statistiques et du Machine Learning a aboutie à de bons résultats.

Dans notre dernière contribution, nous avons proposé le descripteur Optim-LBP qui génère une nouvelle description de l'image en formulant le processus d'encodage de cette dernière comme un problème d'optimisation où chaque pixel courant dépend implicitement des contributions de ses plus proches voisins. Dans ce problème, la fonction objectif régissant chaque voisinage est défini comme une mesure de son degré d'homogénéité basée sur l'entropie de Shannon. En effet, Optim-LBP utilise des méthodes d'optimisation performantes, primales et duales, pour déterminer les variables représentant le degré de participation de chaque pixel voisin dans la constitution du pixel courant.

Cette thèse est organisée selon les six chapitres suivants :

Le premier chapitre présente la problématique générale des systèmes biométriques et focalise la biométrie faciale et plus précisément la reconnaissance des visages.

Le deuxième chapitre est consacré à l'étude détaillée des méthodes et des algorithmes de l'état de l'art concernant l'extraction de caractéristiques, la réduction de la dimension, et la classification. Cette étude a focalisé les méthodes basées sur la PCA et la LDA, ainsi que celles basées sur l'analyse spectrale. En plus, elle a mis l'accent sur les réseaux de neurones artificiels et les machines à vecteurs de support. Enfin, elle a présenté une taxonomie du descripteur local LBP où des variétés récentes de ce dernier sont brièvement exposées.

Le troisième chapitre est réservé à la présentation de notre descripteur LBPP, ainsi qu'à l'étude du système de reconnaissance faciale basé sur celui-ci. Dans ce cadre, plusieurs expériences d'évaluation de performances sont réalisées à l'aide des bases de données très connues. Les résultats satisfaisants ainsi obtenus sont analysés et discutés, des conclusions sont tirées.

Pareillement, dans le quatrième chapitre nous présentons notre descripteur LGPP, nous mettons en évidence ses avantages et nous étudions sa contribution à la conception de notre système de reconnaissance faciale. Les résultats obtenus par des expériences réalisées à l'aide des bases de données standards sont aussi analysés et discutés, des conclusions sont tirées.

Le cinquième chapitre présente notre troisième contribution qui concerne notre descripteur EVBP. Après sa présentation détaillée, ce descripteur est adopté pour la phase d'extraction de caractéristiques faciale de notre système de reconnaissance. Des expériences d'évaluation de performances sont réalisées et des résultats confirmant l'efficacité de l'approche proposée sont analysés et discutés, des conclusions sont tirées.

Dans le dernier chapitre, nous présentons notre descripteur Optim-LBP. Ensuite, nous l'utilisons pour l'extraction de caractéristiques faciale de notre système de reconnaissance. Des expériences d'évaluation de performance du système de reconnaissance faciale sont réalisées et les résultats obtenus sont analysés puis discutés; des conclusions sont tirées.

Enfin, nous concluons cette thèse et nous présentons quelques perspectives pour des travaux futurs envisageables.

---

---

## **Chapitre I. Systèmes Biométriques et Biométrie Faciale**

---

---

La grande évolution des techniques liées à l'intelligence artificielle a fait de la vision par ordinateur une tâche omniprésente dans différents secteurs de notre société. Elle a, en fait, poussé en avant les recherches dans plusieurs domaines d'analyse et de traitement d'image y compris la biométrie faciale. Cette dernière est devenue la ligne de recherche la plus abordé par la communauté scientifique.

## I. Introduction

L'identification automatique des personnes est devenue un outil indispensable dans une multitude de contextes, notamment la sécurité des personnes, le contrôle d'accès physique et logique et etc. Elle a en fait attiré l'attention de la communauté scientifique qui a mis en place une nouvelle ligne de recherche scientifique appelée biométrie. Plusieurs travaux de recherche portant sur les modalités biométrique ont été réalisés partant des années 70. Ces travaux représentent un état de l'art de la biométrie riche, diversifiée et accessible.

Dans ce chapitre, nous commençons par la présentation des meilleurs travaux qui ont été effectués dans le domaine de la biométrie. Ainsi, nous exposons les composantes des systèmes biométriques, leur conception, leurs modes de fonctionnement et les critères d'évaluation de leurs performances. Ensuite, nous soulignons les caractéristiques d'un système de reconnaissance faciale, nous évoquons ses avantages et ses défis. Nous présentons aussi d'autres méthodes de biométrie faciale et des bases de données des images de visage 2D.

## II. Systèmes biométriques

### II.1. Définition

La biométrie désigne l'ensemble des processus de vérification de l'identité d'un individu à partir de ses caractéristiques physiques ou comportementales. Elle regroupe l'ensemble de moyens universels appelés modalités biométriques ; Voir, Figure I.1.

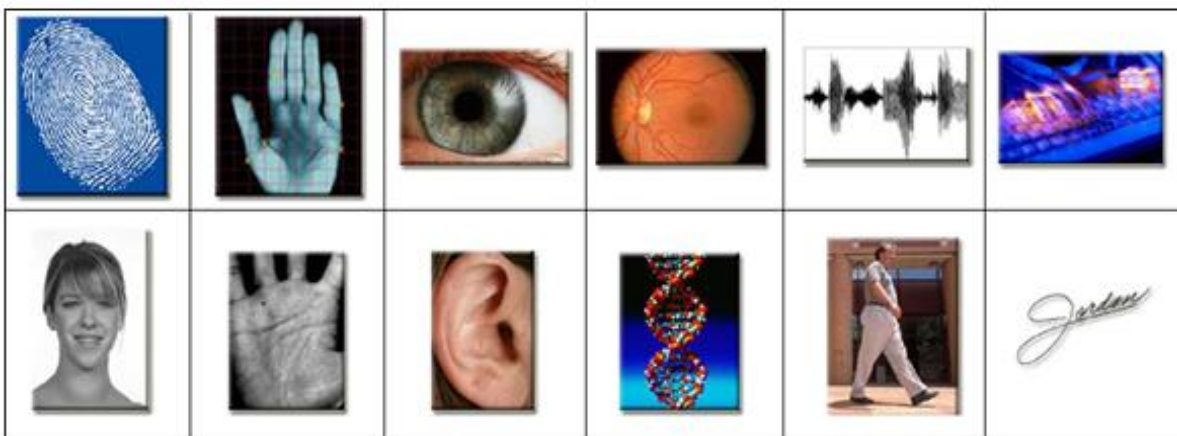


Figure I.1: Différentes Modalités Biométriques

Pratiquement, pour qu'une caractéristique personnelle puisse être qualifiée de modalité biométrique, elle doit être universelle, unique, permanente, enregistrable, acceptable, performante et difficilement falsifiable [1-3]. Cependant, les différentes modalités biométriques du Tableau I-1 ne peuvent garantir le même degré de critères de quantification, le choix de l'une à l'autre doit répondre aux besoins de chaque application [5-6,21].

Tableau I-1: Comparaison de différentes modalités biométriques.

Modalités	Universelle	Unique	Permanente	Enregistrable	Performance	Acceptable	Infalsifiable
ADN	Fort	Fort	Fort	Faible	Fort	Faible	Faible
Voix	Moyen	Faible	Faible	Moyen	Faible	Fort	Faible
Iris	Fort	Fort	Fort	Moyen	Fort	Faible	Fort
Visage	Fort	Faible	Moyen	Fort	Faible	Fort	Faible
Rétine	Fort	Fort	Moyen	Faible	Fort	Faible	Fort
Empreintes	Moyen	Fort	Fort	Moyen	Fort	Moyen	Fort
Signature	Faible	Faible	Faible	Fort	Faible	Fort	Fort

## II.2. Conception d'un système biométrique

Comme systèmes informatiques, les systèmes biométriques doivent avoir une structure globale commune à quelques différences près et doivent obéir à un ensemble de critères d'évaluation de performance. En effet, les systèmes biométriques peuvent fonctionner soit en mode d'enrôlement, soit en mode d'authentification ou bien en mode d'identification [21].

- Mode d'enrôlement : phase préliminaire de tout système biométrique où les caractéristiques biométriques des individus, souhaitant se faire enrôler, sont extraites et ensuite enregistrées dans le module base de données.
- Mode d'authentification : un test de correspondance « 1:1 » dans lequel le système vérifie l'approbation de l'identité déclarée par l'utilisateur (accès à des données sécurisées ou ressources informatiques).

$$f(V_p, V_v) = \begin{cases} 1, & \text{si } S(V_p, V_v) > \alpha \\ 0, & \text{sinon} \end{cases} \quad (I.1)$$

Où :  $V_p$ ,  $V_v$ ,  $\alpha$  et  $S$  sont respectivement le vecteur proclamé, le vecteur validé, le seuil de décision et la fonction de similarité.

- Mode d'identification : un test de correspondance « 1:N » dans lequel le système doit déterminer l'identité d'une personne en comparant sa signature biométrique à celles des N utilisateurs préalablement enrôlés (sécurité et contrôle d'accès aux lieux publics).

$$f(V_p) = \begin{cases} I_k, & \text{si } \max_{1 < k < N} S(V_p, V_k) > I_0 \\ I_0, & \text{sinon} \end{cases} \quad (1.2)$$

Où :  $I_p$ ,  $I_k$ , et  $V_k$  sont respectivement l'identité proclamée, l'identité enrôlée, le vecteur caractéristique de  $I_k$ .

Étant un système de reconnaissance de forme, un système biométrique peut être représenté par les quatre principaux modules illustrés dans la Figure I.2. Soient,

- Module capture : dans lequel le système acquiert les données biométriques d'un individu via un terminal de capture (appareils photo, caméras, scanners ...).
- Module traitement de signal : consiste à déterminer des nouvelles représentations des données acquises. Il est généralement composé de trois phases : la phase de prétraitement et de normalisation des données acquises, la phase d'extraction de l'information pertinente et la phase d'optimisation de l'information extraites et de création des signatures propres. Ces signatures sont mesurables et enregistrables.

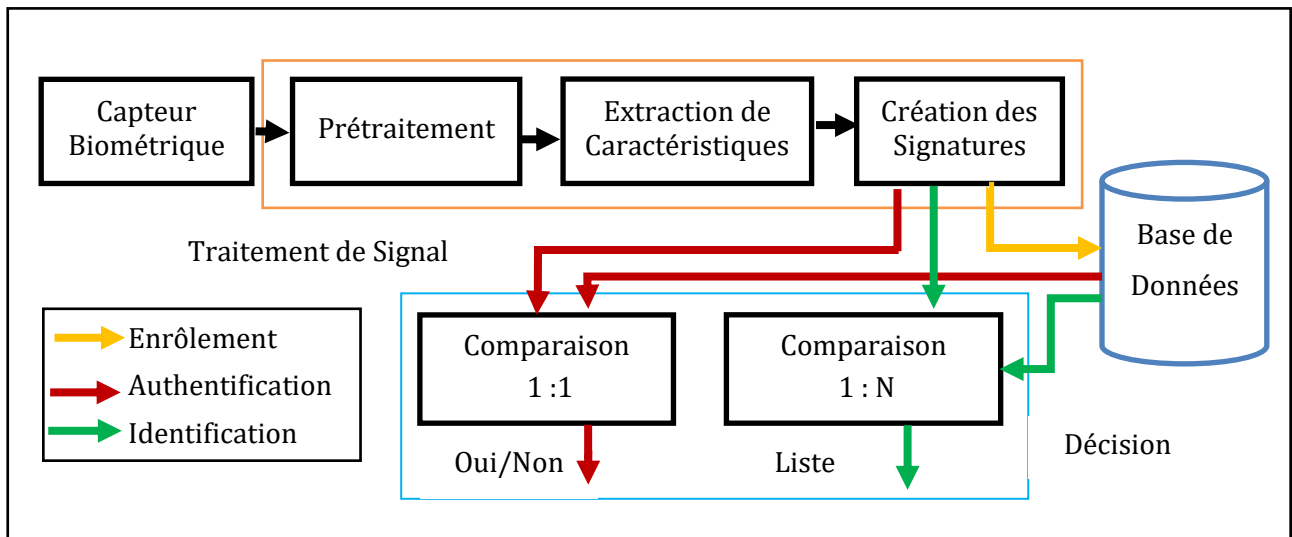


Figure I.2 : Architecture Générique d'un Système Biométrique.

- Module base de données : ce module utilise des supports de sauvegarde pour préserver les signatures biométriques des individus préalablement enrôlés.

- Module de décision : phase de classification de l'ensemble des signatures biométriques qui détermine la décision finale du système.

### III. Mesure de performance d'un système biométrique

Le déploiement d'un système biométrique réel et fonctionnel doit mettre en place les critères d'évaluation de la performance suivants : les taux d'erreur, les courbes de performance et les points de fonctionnement.

- Taux d'erreur : ce sont des variables de quantification des erreurs de prise de la décision dont les deux principaux sont :
  - Taux de faux rejet (*False Rejection Rate*, FRR) : représente la proportion d'individus censés être reconnus par le système mais qui sont malheureusement rejetés.
  - Taux de fausse acceptation (*False Acceptation Rate*, FAR) : représente la proportion d'individus censés être rejetés par le système mais qui sont tout de même acceptés.
- Courbes de performance : ce sont des illustrations graphiques qui résument les résultats de l'étude de la performance en fonction d'un certain seuil. Deux courbes sont les plus utilisées.
  - Courbe ROC (*Receiver Operating Characteristic Curve*) [22] : représente la relation entre les FAR et les FRR pour différentes valeurs du seuil de décision et sert généralement à évaluer un système d'authentification.
  - Courbe CMC (*Cumulative Match Characteristic Curve*) [23] : représente la probabilité que la personne est parmi les  $n$  premiers individus du système (appelé rang). On dit qu'un système reconnaît au rang 2 quand il choisit le plus proche des deux individus à la personne recherchée. Cette courbe sert souvent à évaluer un système d'identification.
- Points de fonctionnement : points critiques des courbes de performance qui représentent des compromis entre les FAR et les FRR. Les quatre points les plus utilisés sont des points de la courbe ROC.
  - FAR fixé : c'est un point qui correspond au seuil tel que le FAR est fixe.
  - FRR fixé : c'est un point qui correspond au seuil tel que le FRR est fixe.



- Taux d'égale erreur (*Equal Error Rate*, EER) : point qui correspond au seuil qui vérifie l'égalité : FAR = FRR.
- Taux d'erreur pondéré (*Weight Error Rate*, WER) : point qui correspond au seuil qui donne des taux FAR et FRR proportionnels.

## IV. Biométrie faciale

Compte tenu de l'immense révolution qu'a connue notre société en termes de volume et de diversité des services fournis par différents organismes (transactions financières, accès aux services et etc.), l'identification automatique des personnes est devenue une tâche indispensable. La question clé qui se pose est : « quelle modalité biométrique pour telle application ? ». Actuellement, la majorité des thèmes de recherche se focalise sur les modalités non intrusives dont la biométrie faciale est la plus populaire. En effet, demander à chaque utilisateur de coopérer avec le capteur biométrique chaque fois qu'il se présente au système biométrique reste un acte contraignant et mal perçu. Telle situation est facilement évitée par la biométrie faciale où le mode d'acquisition est socialement acceptable. Ainsi, le facial-scan des personnes est assuré par des caméras de qualité sans aucune perturbation. Cependant, concevoir un système de reconnaissance faciale aussi efficace reste un vrai défi. Ce défi est généralement lié à différents facteurs affectant l'apparence du visage tels que l'éclairage, la pose, les expressions faciales et etc.

### IV.1.Principaux défis de la biométrie faciale

La reconnaissance d'une personne à partir des images reste une tâche difficile suivant les propriétés du visage et les conditions d'acquisitions. Ces changements sont en fait les principales causes d'une possible fausse identification.

- Changement d'illumination : l'illumination affecte l'apparence d'un visage selon deux types d'éclairage : éclairage global et éclairage local. La Figure I.3 illustre des exemples où le changement d'éclairage affecte totalement ou partialement la surface faciale [24].
- Variation de pose : l'orientation du visage par rapport au plan (OXY), contenant le capteur, affecte amplement son apparence. La Figure I.4 illustre des exemples de trois types de variations: de pose par rotation autour de (OZ), de profil par rotation autour de (OY), et en profondeur par rotation autour de (OX) [25].



Figure I.3 : Exemples de changements d'illumination.

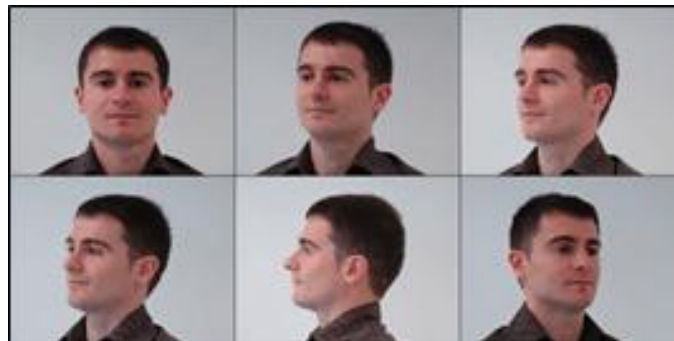


Figure I.4 : Exemples de variations de pose

- Variation des expressions faciales : comme système dynamique, le visage est le siège d'un ensemble de déformations appelées expressions faciales. Comme elles sont illustrées dans la Figure I.5, ces déformations affectent amplement la géométrie des principaux traits de visage en conduisant à un changement de l'aspect local de celui-ci [26].



Figure I.5 : Exemples de variations des expressions faciales

- Oclusions partielles: dans une application réelle (Figure I.6), le visage peut être partiellement occulté par des composantes structurelles ou par des objets de la scène. Ce qui peut causer une forte défaillance du système de reconnaissance [27].

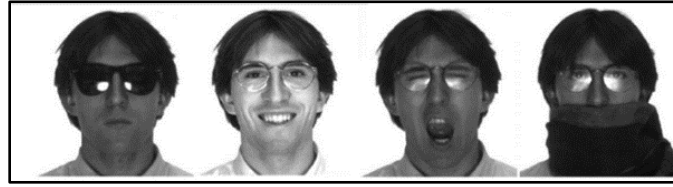


Figure I.6 : Exemples d'occlusions partielles

- Autres limitations : il existe d'autres types de limitations qui peuvent causer une séreuse défaillance du système de reconnaissance de visages. Ainsi, l'existence de différents types de bruits, de maquillages, des effets de vieillissement et des effets de résolutions ont un grand impact sur l'apparence du visage.

## IV.2.État de l'art de la détection des visages

Dans un système de biométrie faciale, la détection des visages est une phase décisive au cours de laquelle s'exécutent des processus de localisation des visages existant dans une image fixe ou dans une séquence vidéo. En effet, les méthodes de détection de visages peuvent être regroupées en deux approches : des méthodes basées sur les caractéristiques invariantes et des méthodes basées sur l'apparence globale.

### IV.2.1.Méthodes basées sur les caractéristiques invariantes

Ces méthodes se basent sur la recherche des caractéristiques structurelles légèrement variables avec les conditions de capture. La couleur de la peau est devenue un principal argument pour la détection des visages où la décision d'appartenance à la classe peau est prise selon l'un des modèles suivants : modèle à seuils fixes, modèle paramétrique et modèle non-paramétrique.

#### IV.2.1.1. Modélisation à frontières de décision fixes

Généralement, la modélisation à seuils fixe prend en compte l'espace colorimétrique et les règles de décisions qui définissent les frontières d'appartenance de la couleur de la peau. En effet, les trois espaces colorimétriques les plus étudiés sont: RGB, HSV et  $YCbCr$  [28].

- Espace RGB : cet espace utilise les trois rayonnements de base : le rouge (R), le vert (G) et le bleu (B). Les règles de décision de cet espace sont :

$$\text{Éclairage fort : } \begin{cases} ((R > 95) \wedge (G > 40) \wedge (B > 20)) \wedge \\ (\max(R, G, B) - \min(R, G, B) > 15) \wedge \\ (\text{abs}(R - G) > 15) \wedge ((R > G) \wedge (R > B)) \end{cases} \quad (1.3)$$

$$\text{Éclairage faible : } \begin{cases} ((R > B) \wedge (G > B)) \vee ((R > 220) \wedge (G > 210)) \vee \\ (\text{abs}(R - G) < 15) \wedge (B > 170) \end{cases} \quad (1.4)$$

- Espace HSV : déduit de l'espace RGB, HSV utilise les trois canaux suivants : le canal (V) représentant l'intensité, et les deux canaux (H et S) représentant la chromatique. La règle de décision de ce système est :

$$0 \leq H \leq 50 \wedge 0.23 \leq S \leq 0.68 \quad (1.5)$$

- Espace  $YC_bC_r$  : déduit de l'espace RGB,  $YC_bC_r$  utilise les trois canaux suivant : le canal (Y) représentant l'intensité, et les deux canaux ( $C_b$  et  $C_r$ ) représentant la chromatique. La règle de décision de ce système est :

$$((C_b > 105) \vee (C_b < 130)) \wedge ((C_r > 135) \vee (C_r < 160)) \quad (1.6)$$

Par ailleurs, les règles de décisions précédentes sont fusionnées soit en parallèle soit en cascade pour augmenter la fiabilité du processus de détection.

#### IV.2.1.2. Modélisation paramétrique

La modélisation paramétrique consiste à ajuster la distribution de la couleur de peau avec des fonctions paramétrées. L'une des plus populaires de ces fonctions est la loi gaussienne. Cette dernière permet d'estimer la densité de probabilité de la distribution de la couleur de peau en tenant compte des deux paramètres la moyenne  $\mu_{\text{peau}}$  et la matrice de covariance  $\sigma_{\text{peau}}$ . Deux modèles sont proposés, le modèle simple Gaussienne, et le mélange de gaussiennes [29].

$$p(c/\text{peau}) = \frac{1}{2\pi \sqrt{|\sigma_{\text{peau}}|}} \exp\left(-\frac{1}{2}(c - \mu_{\text{peau}})^t \sigma_{\text{peau}}^{-1} (c - \mu_{\text{peau}})\right) \quad (1.7)$$

$$\mu_{\text{peau}} = \frac{1}{n_{\text{peau}}} \sum_c n_{\text{peau}}(c) \quad (1.8)$$

$$\sigma_{\text{peau}} = \frac{1}{n_{\text{peau}} - 1} \sum_c n_{\text{peau}}(c) (c - \mu_{\text{peau}})^t (c - \mu_{\text{peau}}) \quad (1.9)$$

Avec,  $c$  : un vecteur de couleur et  $n$  : taille de l'échantillon

### IV.2.1.3. Modélisation non-paramétrique

Cette modélisation consiste à estimer la distribution de la couleur de la peau en utilisant les données d'apprentissage et sans aucune hypothèse sur les paramètres de cette distribution. Plusieurs méthodes sont proposées pour créer un tel modèle dont les plus utilisés sont : modèle de Bayés, réseaux de neurones artificiels, et machines à vecteurs de support. Tous ces modèles traitent le problème de détection de la couleur de la peau comme étant un problème de classification à deux classes (peau/non-peau).

### IV.2.2. Méthodes basées sur l'apparence globale

Le principe de ces méthodes consiste à modéliser le problème de détection de visages comme un problème de classification par apprentissage artificielle à deux classes (visage/non-visage). En se basant sur les travaux de Rowley et al [30], Viola et Jones ont introduit une méthode de référence pour la détection du visage [31]. Cette méthode comporte les quatre innovations suivantes :

- Filtre de Haar : un masque qui détermine les caractéristiques de *Haar* comme étant la différence des sommes de pixels des zones rectangulaires adjacentes.
- Image intégrale : une image dont la valeur de chacun de ses pixels est la somme des pixels situés au-dessus et à gauche de celui-ci.
- Algorithme Adaboost : construit un classificateur fort comme somme pondérée de classificateurs faibles suivant le principe de *Boosting*.
- Cascade de classificateurs : afin de réduire au maximum le temps de calcul, Viola et Jones proposent d'appliquer séquentiellement les classificateurs forts. Seules les zones de l'image classées positives par le classificateur courant sont traitées par le classificateur suivant.

## IV.3. État de l'art de la reconnaissance de visages

La création d'un système de reconnaissance faciale intelligent et autonome, nécessite le développement des algorithmes capables de simuler artificiellement l'aptitude naturelle des humains à reconnaître les visages. Plusieurs travaux de recherches ont bénéficié des avancées théoriques en algèbre linéaire, en analyse spectrale, en statistique, et en géométrie pour développer de nombreux algorithmes dédiés pour la reconnaissance faciale 2D. Ces

algorithmes (Figure I.7) sont regroupés en trois principales approches : approche globale, approche locale et approche hybride.

### IV.3.1. Approche globale ou Holistique

Les méthodes de cette approche se basent principalement sur l'apparence globale des visages et consistent à déterminer des bases propres de projection par optimisation de certains critères de décision. Deux types de méthodes sont distingués : les méthodes linéaires et les méthodes non linéaires.

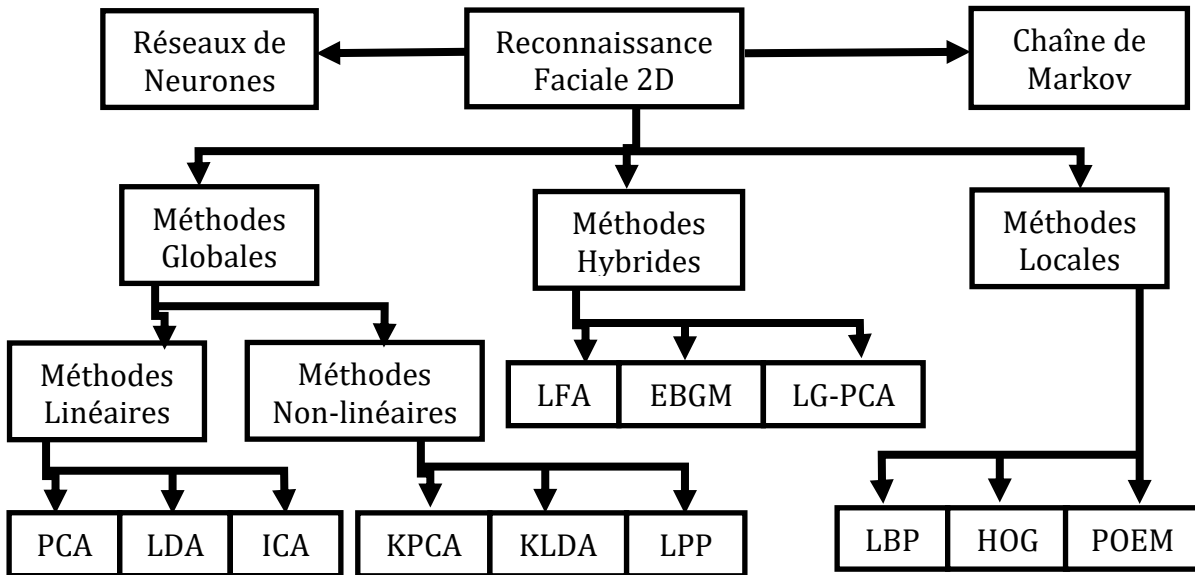


Figure I.7: Exemple de méthodes de reconnaissance de visages.

#### IV.3.1.1. Méthodes linéaires

Les méthodes linéaires sont des applications directes de l'algèbre linéaire dans le domaine d'analyse des données. Elles sont basées sur les résolutions des équations aux valeurs propres. L'analyse en composantes principales (*Principal Component Analysis, PCA*) et l'analyse discriminante linéaire (*Linear Discriminant Analysis, LDA*) sont les méthodes de référence de cette approche.

- PCA : c'est la méthode la plus ancienne et la mieux étudiée de l'état de l'art de réduction de la dimension. Elle consiste à déterminer la base de projection qui maximise la variance des composantes projetées. L'extension de cette méthode à la reconnaissance faciale est nommée Eigen-faces [10,32]. Plusieurs améliorations visant surmontées les limitations de

PCA ont été proposées, notamment Torre a proposé une variété qui maximise la variance totale tout en minimisant la variance dans les directions supposées contenir du bruit [33]. Moghaddam a proposé une variété qui intègre les probabilités de Bayes [34]. Tandis que, Wang a proposé de résoudre le problème de sélection optimale des composantes principales [35].

- LDA : c'est une méthode aussi répandue que PCA qui détermine le sous espace propre qui maximise le critère d'optimisation de Fisher [11]. Cela revient à maximiser les variations interclasses tout en minimisant les variations intra-classes. Tout comme les Eigen-faces, les Fisher-faces fut la première extension de LDA à la reconnaissance faciale proposée par Belhumeur et al[36-37].

Bien que les deux méthodes, PCA et LDA, reposent sur le même principe, réduction des endomorphismes, elles présentent quelques différences qui favorisent l'une à l'autre. En effet, PCA est favorisée pour le traitement des données intégralement réparties ; alors que LDA est favorisée pour le traitement des données regroupées en classes. En plus de PCA et LDA, nombreuses sont les méthodes linéaires proposées pour la reconnaissance de visages dont les principales sont : L'analyse en composantes indépendantes (*Independent Component Analysis, ICA*) [38-39], la factorisation de matrices non négatives (*Non-negative Matrix Factorization, NMF*) [40] et L'analyse discriminante bilinéaire (*Bi-linear Discriminant Analysis, BDA*) [41].

#### IV.3.1.2. Méthodes non linéaires

Généralement, l'aspect non linéaire des données traitées limitent les performances des méthodes de réduction linéaires. Afin de remédier à ce problème, des méthodes de réduction non linéaires sont proposées. Les plus utilisées sont celles basées sur les noyaux (*kernel based methods*) et celles basées sur l'apprentissage de variétés (*Manifold learning methods*).

- Méthodes basées sur les noyaux : elles consistent à projeter les données non linéaires dans un espace de Hilbert où leurs projections sont linéairement séparables. Ensuite appliquer les méthodes linéaires sur ces projections. Parmi les méthodes de cette classe nous citons : KPCA et KLDA [42- 43].
- Méthodes basées sur l'apprentissage de variétés : ce sont des méthodes qui considèrent les données comme des échantillons d'une variété Riemannien. Trois types sont distingués : les méthodes basées sur la conservation des distances comme le *Multi-dimensional Scaling*

(MDS) [44], l'*Isomap* [45] et les cartes de diffusion [46]. Celles fondées sur l'approximation locale par des espaces euclidiens dont les plus connues sont : *Local Linear Embedding (LLE)* [47] et *Local Preserving Projection (LPP)* [48]. Et les techniques utilisant la conservation globale des propriétés topologiques comme les auto-encodeurs, les réseaux de neurones et les arbres de décision.

### IV.3.2. Approche locale

Les méthodes de cette approche consistent à modéliser le visage similairement à notre perception visuelle, et le représente par un ensemble de caractéristiques invariantes. Ces méthodes sont regroupées en deux catégories, les méthodes basées sur les caractéristiques locales (*Local feature based method*) et les méthodes basées sur l'apparence locale (*Local appearance based method*).

#### IV.3.2.1. Méthodes basées sur les caractéristiques locales

Les méthodes basées sur les caractéristiques locales s'appuient principalement sur l'extraction des zones d'intérêts et l'élaboration des modèles du visage selon deux alternatives : la première utilise les distances entre les positions relatives aux points d'intérêts, tandis que la deuxième génère des graphes dont les sommets sont caractérisés par ces points d'intérêts.

- Méthodes géométriques : la première méthode de cette catégorie est proposée par Brunelli [49], elle représente chaque visage par un vecteur composé de 35 caractéristiques géométriques. De même, Heisele [50] a proposé de représenter chaque visage par un vecteur qui concatène les éléments des zones centrée sur dix points caractéristiques. Bien que le coût de stockage des méthodes géométriques soit très faible, elles souffrent des effets d'illumination, de pose, d'occultation et d'expression faciale.
- Méthodes basées sur les graphes : ces méthodes considèrent la reconnaissance comme un problème d'appariement de graphes. Elles associent à chaque visage un graphe topologique dont les nœuds représentent des caractéristiques locales et les arrêtes représentent les relations spéciales entre des caractéristiques. Manjunath et al[14] ont proposé de représenter les caractéristiques faciales à partir d'ondelettes de Gabor. Lades et al[51] ont proposé le modèle Architecture de Lien Dynamique (*Dynamic Link Architecture, DLA*) qui utilise un graphe topologique déformable au lieu d'un graphe topologique fixe.



#### IV.3.2.2. Méthodes basées sur l'apparence locale

Ces techniques définissent pour chaque visage un modèle globale par combinaison d'un ensemble de modèles locaux. Afin de définir chaque modèle local nous devons d'abord découper l'image en régions, ou patches, caractériser chaque région par un vecteur local, et concaténer tous les vecteurs locaux en un seul vecteur global représentant le visage. Parmi ces méthodes nous citons : les ondelettes de Gabor, le modèle binaire local (LBP), Histogramme de gradient orienté (HOG) et le modèle de magnitudes de bord orientées (POEM).

- Les ondelettes de Gabor : ce sont des filtres spectraux utilisés pour l'extraction de l'information utile suivant des orientations et des échelles spécifiques. Généralement, les filtres de Gabor utilisent huit orientations et cinq échelles pour représenter l'image par des coefficients complexes appelés composants de Gabor [52-53].
- Modèle binaire local (LBP) : depuis qu'il a été introduit par Ojala pour l'analyse de textures, LBP est rapidement étendu à la vision par ordinateurs. Il consiste à générer une représentation de l'image où chacun de ses pixels est recodé en fonction de son voisinage [54-55]. Motivé par les fonctionnalités de LBP, Ahonen propose une nouvelle description du visage appelée histogramme spatialement amélioré (eLBPH). Dans cette description les vecteurs locaux sont des histogrammes [16,56].
- Histogramme de gradient orienté (HOG) : proposé Dalal et Triggs le descripteur HOG décrit l'apparence locale d'un objet en divisant l'image en petites régions adjacentes, appelées cellules. Pour chaque cellule un histogramme des directions de gradient est déterminé par processus de vote [15-57].
- Modèle de magnitudes de bord orientées(POEM) : plus récemment, Vu a proposé d'extraire les caractéristiques locales du visage en tenant compte de l'information sur les orientations multi-échelles au niveau de ses contours locaux. Il a proposé d'appliquer le principe de LBP à des voisinages des orientations où chaque valeur de pixel est remplacée par l'amplitude de son gradient. Ils ont incorporé les informations de gradient des pixels voisins en calculant un histogramme local des orientations sur tous les pixels de la cellule. Enfin, chaque pixel est caractérisé par les valeurs de son niveau de gris et de ses orientations [58-59].

- Modèles de Markov cachés (MMC) : L'idée derrière l'utilisation des MMC réside dans le fait que la structure d'un visage peut être considérée comme une séquence de régions distinctes et ordonnées (front, yeux, nez, bouche, etc.), chaque région est assignée à un état du modèle MMC unidimensionnel. Dans leur travail, Samaria a découpé les images de visage en plusieurs bandes. Ils ont modélisé les transitions entre les bandes de manière probabiliste à l'aide de MMC, ainsi les mesures de la similarité entre visages sont ensuite évaluées de manière probabiliste par un décodage de leurs MMCs [60].

### IV.3.3.Approche hybride

Les méthodes hybrides résultent des combinaisons entre les méthodes des deux approches locale et globale. Elles exploitent la complémentarité entre ces méthodes en bénéficiant des leurs avantages. A titre d'exemples, on trouve :

- Pentland présente dans [61] l'approche dite des espaces propres modulaires (*Modular Eigenspaces*) Cette technique réalise une PCA et une classification sur des régions distinctes du visage (yeux, nez, bouche ou encore le visage entier).
- Proposée par Perlibakas [17], la méthode Log-Gabor-PCA consiste à appliquer la PCA pour réduire l'ensemble de vecteurs obtenus par application des filtres de Gabor autour de certains points caractéristiques du visage. Ces vecteurs contiennent à la fois la localisation et les amplitudes des énergies locales.
- Proposée par Wiskott et al [18], la méthode *Elastic Bunch Graph Matching* (EBGM) représente le visage par un Face Bunch Graph (FBG).A chaque nœud du graphe est associée des jets représentant les 40 coefficients du filtre de Gabor.

### IV.4.Base de données de visages

Au cours des années, de nombreuses bases de données ont été conçues pour répondre aux besoins spécifiques des problèmes de la reconnaissance faciale. Ces bases de données sont utilisées pour l'évaluation des performances des nouveaux algorithmes développés. Dans cette section, nous détaillons les bases de données utilisées pour l'évaluation de nos approches et nous listons les principales d'entre elles dans le Tableau I-2.

#### IV.4.1. Base de données ORL

Collectée par le laboratoire de, AT&T de l'université de Cambridge, la base de données ORL renferme 400 images, 10 vues pour chacune des 40 personnes (Figure I.8). Ces images ont la même taille de  $112 \times 92$  pixels, présentent des variations d'illumination, de poses, d'expressions faciales et d'occultations partielles [62].



Figure I.8 : Exemples des images extraites de la base ORL

#### IV.4.2. Base de données Yale

La base de données Yale est composée de 15 personnes différentes, 11 images pour chaque personne avec la résolution de  $320 \times 240$  pixels. Toutes les 165 images sont sous-échantillonnées pour avoir des images de tailles  $112 \times 92$  pixels. Les images de chaque personne (Figure I.9) présentent des variations dans les conditions d'éclairage, de poses, des expressions faciales et des occultations partielles [63].



Figure I.9 : Exemples des images extraites de la base Yale

#### IV.4.3. Base de données AR

La base de données AR a été enregistrée en 1998 au sein du laboratoire Computer Vision Center (CVC) en Espagne [64]. Elle présente 126 personnes pour un total de 3276 images en couleur . Les images sont présent en deux sessions à de deux semaines d'intervalle, 13 vues par

personne sont collectées dans chaque session (Figure I.10). Les 26 vues de chaque personne renferment des changements d'expression faciale, d'éclairage et des occultations partielles des yeux et de la partie basse du visage. Dans ce travail, nous avons utilisé un sous ensemble de la base AR formé par les premières 14 vues de 100 personnes. Les images de taille originale 768×576 sont sous-échantillonnées pour réduire la taille à 112×92 [65].



Figure I.10 : Exemples des images extraites de la base AR

Tableau I-2: Aperçu des principales bases de données de reconnaissance du visage.

Bases de données	Nombre de personnes	Nombre des images	Caractéristiques	Adresses
ORL	40	400	Variation de temps, illumination, expression, lunettes.	[62]
Yale	15	165	Expressions, lunettes, éclairage.	[63]
AR	126	4000	Pose frontale, expression, illumination, occlusions, lunettes, écharpes.	[64-65]
FERET	1199	14126	images en couleur, changements d'apparence à travers le temps, pose, expressions.	[66-67]
Extended Yale-B	38	2414	Poses multiples, conditions d'illumination par pose.	[68]
XM2VTS	295	1526	Changements de coiffure, poils du visage, lunettes.	[69]
Multi-PIE	337	750000	Haute résolution, Points de vue différents, illumination, expressions.	[70]
UMIST	20	564	Pose, sexe, race, niveaux de gris	[71]
MIFS	107	642	Maquillage tirées de vidéos aléatoires de YouTube.	[72]
VMU	51	204	Maquillage synthétique dans le jeu de données FRGC.	[72]
LFW	5749	13233	Images dégradées, illumination, poses, basse flou, race, résolution, vieillissement, occlusion, réels	[73-74]

## **V. Conclusion**

Dans ce chapitre, nous avons mis en relief les modalités biométriques utilisées pour l'identification des individus. Nous avons présenté leur architecture commune, ainsi que les principaux paramètres de mesure et d'évaluation de leurs performances. Ensuite, nous avons mis en évidence les avantages et les challenges liés à la biométrie faciale. Nous avons donné un aperçu global sur les méthodes de détection et de reconnaissance de visages. Une taxonomie de la reconnaissance faciale est présentée en classant les méthodes de l'état de l'art en trois approches (globale, locale et hybride). Finalement, quelques bases de données communément utilisées pour la validation des systèmes de reconnaissance des visages 2D sont exposées.

Dans le prochain chapitre, nous présentons les principes et les intérêts des méthodes utilisées dans nos contributions selon trois axes principaux: extraction des caractéristiques, réduction de la dimension et classification.

---

---

## **Chapitre II. Extraction des caractéristiques, réduction de la dimension et classification**

---

---

Un système de reconnaissance faciale est une association en cascade de trois phases, extraction des caractéristiques, réduction de la dimension et classification. Afin d'améliorer les performances du processus de reconnaissance, plusieurs méthodes assurant le bon fonctionnement de chacune de ces trois phases sont proposées. Dans ce chapitre, nous mettons la lumière sur quelques méthodes de la reconnaissance faciale 2D, en particulier celles utilisées dans nos contributions.

## I. Introduction

Pour un système de reconnaissance faciale, le principal objectif est d'atteindre un taux de reconnaissance élevé et un temps d'exécution minimal. Mais, nombreux sont les problèmes qui s'opposent à cet objectif, notamment, la taille des images, la structure des bases de données, la variabilité interclasse et la variabilité intra-classe. Afin de remédier à ces problèmes, plusieurs méthodes représentant l'exploration des données de la reconnaissance faciale 2D sont proposées. Dans ce chapitre, nous présentons en détail certaines méthodes de l'état de l'art de la reconnaissance faciale principalement celles utilisées dans la suite de cette thèse. Dans un premier temps, nous présentons l'aspect théorique et la particularité des méthodes de réduction de la dimension, notamment, ceux basés sur la PCA, la LDA, et la compressions spectrale. Dans un deuxième temps, nous présentons les algorithmes de classification, notamment, ceux basés sur la similarité et ceux basés sur l'apprentissage artificiel, particulièrement les SVMs et les MLPs. Enfin, nous présentons les méthodes d'extraction des caractéristiques, plus précisément celles basées sur le célèbre modèle binaire local (*Local Binary Pattern, LBP*) et ses variétés les plus récentes.

## II. Méthodes de réduction de la dimension

Dans la vision par ordinateur, les données brutes à traiter sont des images de grandes tailles, leur traitement nécessite un coût temporel assez élevé. Afin de réduire ce coût, ces images doivent être représentées par des vecteurs descripteurs de petite taille. Afin d'engendrer tels vecteurs descripteurs et sans pertes significatives de l'information pertinente, nous faisons appel à des méthodes de réduction de la dimension principalement celles basées sur la PCA, celles basées sur la LDA et celles basées sur l'analyse spectrale [75].

### II.1. Méthodes basées sur PCA

Comme il est déjà mentionné au premier chapitre de ce mémoire, PCA est une méthode d'analyse statistique qui permet de réduire la dimension d'un jeu de données tout en maximisant la variance des données projetées. Cela revient à déterminer le sous espace propre engendré par les vecteurs propres de la matrice de covariance. Étant donné un jeu de  $N$  échantillons  $x_i \in \mathbb{R}^D$  et un vecteur  $\vec{u} \in \mathbb{R}^D$  tel que :

$$\|\vec{u}\| = \vec{u}^t \vec{u} = 1 \quad (\text{II.1})$$

La variance de cette population est :

$$\sigma^2 = \frac{1}{N-1} \sum_i (x_i - \bar{x})^t (x_i - \bar{x}) \quad \text{avec,} \quad \bar{x} = \frac{1}{N} \sum_i x_i \quad (II.2)$$

On définit,  $v_i$ , la projection d'un échantillon  $x_i$  sur la direction  $\vec{u}$  par :  $v_i = \vec{u}^t x_i$ . Ainsi la variance projetée de la population sur la direction  $\vec{u}$  est :

$$\begin{aligned} \sigma_v^2 &= \frac{1}{N-1} \sum_i (v_i - \bar{v})^t (v_i - \bar{v}) = \frac{1}{N-1} \sum_i (\vec{u}^t (x_i - \bar{x}))^t (\vec{u}^t (x_i - \bar{x})) \\ &= \frac{1}{N-1} \sum_i (x_i - \bar{x})^t \vec{u} \vec{u}^t (x_i - \bar{x}) \\ &= \frac{1}{N-1} \sum_i ((x_i - \bar{x})^t \vec{u}) (\vec{u}^t (x_i - \bar{x}))^t \\ &= \frac{1}{N-1} \sum_i \vec{u}^t (x_i - \bar{x}) (x_i - \bar{x})^t \vec{u} \\ &= \frac{1}{N-1} \vec{u}^t [\sum_i (x_i - \bar{x}) (x_i - \bar{x})^t] \vec{u} \end{aligned}$$

$$\text{Soit,} \quad \sigma_v^2 = \vec{u}^t C \vec{u} \quad (II.3)$$

$$\text{Avec,} \quad C = \frac{1}{N-1} \sum_i (x_i - \bar{x}) (x_i - \bar{x})^t \quad \text{et} \quad \bar{v} = \frac{1}{N} \sum_i v_i \quad (II.4)$$

La matrice,  $C \in \mathbb{R}^{D \times D}$ , représente la matrice de covariance de l'ensemble de données d'apprentissage  $X = \{x_1, \dots, x_N\} \in \mathbb{R}^{D \times N}$ . La solution optimale du problème de maximisation de la variance projetée,  $\max_{\vec{u}} (\sigma_v^2 = \vec{u}^t C \vec{u})$ , peut être générée grâce au multiplicateur de Lagrange comme une équation aux valeurs propres de la matrice de covariance  $C$ . Par conséquent, le sous espace propre optimal est engendré par les vecteurs propres correspondant aux valeurs propres les plus élevées de  $C$ . Soit,

$$\text{Soit la fonction objectif :} \quad f(\vec{u}, \lambda) = \vec{u}^t C \vec{u} + \lambda(1 - \vec{u}^t \vec{u}) \quad (II.5)$$

$$\text{Dérivation partielle selon } \vec{u} : \quad \frac{\delta f(\vec{u}, \lambda)}{\delta \vec{u}} = 2C\vec{u} - 2\lambda\vec{u} = 0 \quad (II.6)$$

D'où l'équation aux valeurs propres :

$$C \vec{u} = \lambda \vec{u} \quad (II.7)$$

### II.1.1. Méthode des visages propres

Motivés par les travaux de Sirowich et Kirby [76], Turk et Pentland [10] ont proposé la méthode des Eigen-faces. Cette méthode consiste à réduire l'espace de projection des vecteurs caractéristiques des images de visages en appliquant l'algorithme PCA. Afin de définir la

---



matrice de covariance  $C$  régissant l'ensemble d'apprentissage de  $N$  images, PCA transforme chaque image,  $\gamma_i$ , de taille  $(m, n)$ , en un vecteur,  $x_i$ , de taille  $m \times n$ . Ensuite, utilise les vecteurs ainsi obtenus pour construire la matrice  $X = \{x_1, \dots, x_N\}$  de taille  $(m \times n, N)$ . Enfin, définit la matrice de covariance,  $C$ , de taille  $(m \times n, m \times n)$  par le produit  $XX^t$ . Cependant, les valeurs propres de  $C$  n'ont pas la même importance, sauf les «  $d$  » vecteurs propres correspondant aux «  $d$  » plus grandes valeurs propres qui contribuent à la construction de la base propre de projection. L'efficacité de PCA est justifiée par la réduction de la dimension des vecteurs caractéristiques de la valeur  $m \times n$  à la valeur  $d$ . Certes, la qualité de la réduction dépend d'un certain seuil de sélection des valeurs propres actives. Un seuil est dit optimal s'il conserve la majorité de l'information pertinente de l'ensemble d'apprentissage avec le minimum de valeurs propres actives. Deux cas sont à traiter : le cas d'un ensemble d'apprentissage de taille assez grande ( $N \gg m \times n$ ) et celui d'un ensemble d'apprentissage de taille assez petite ( $N \ll m \times n$ ).

- Cas des données de grande dimension: la réduction de la matrice de covariance de taille relativement grande est assez complexe. Pour un jeu de  $N$  images de taille  $n \times n$  ( $n \approx 100$ ), la taille de la matrice de covariance est  $n^2 \times n^2$ . En algèbre linéaire, pour toute matrice  $X$ , les deux produits  $XX^t$  et  $X^tX$  partagent les mêmes valeurs propres non nulles. Par conséquent, réduire  $C = XX^t$  de taille  $n^2 \times n^2$  revient à réduire la matrice  $M = X^tX$  de taille  $N \times N$  ( $N \ll n^2$ ). L'équation aux valeurs propres de  $M$  s'écrit:

$$M\vec{v}_i = \delta_i \vec{v}_i \Leftrightarrow X^t X \vec{v}_i = \delta_i \vec{v}_i \Leftrightarrow XX^t(X\vec{v}_i) = \delta_i(X\vec{v}_i) \quad (II.8)$$

D'où l'équation aux valeurs propres de  $C$ :

$$C\vec{u}_i = \delta_i \vec{u}_i \quad \text{Avec,} \quad \vec{u}_i = X\vec{v}_i \quad \text{et} \quad \delta_i = \lambda_i \quad (II.9)$$

Bien que PCA représente une méthode de référence pour la réduction de la dimension sans aucune connaissance a priori sur les images de visages, elle souffre principalement de trois limitations : La première limitation est liée aux variations intra-classes et interclasses, la deuxième dépend de l'aspect estimatif du seuil de sélection des éléments propres actifs, et la troisième est causée par la non linéarité des données traitées. Pour remédier aux limitations précédemment citées, plusieurs variétés de PCA ont été proposées dans [32-35]. Principalement, l'analyse en composantes principales à noyau (KPCA) et l'analyse en composantes principales à deux dimensions (2DPCA).

### II.1.2. Méthode PCA à noyau

Afin de traiter des données non linéaires, une variété de PCA non linéaire appelée Kernel-PCA (KPCA) a été proposée [42-43,77]. Elle consiste à projeter ces données dans un espace de Hilbert de grande dimension où leur répartition soit le plus linéaire possible. Cela revient à définir une fonction  $\varphi$  et un produit scalaire  $K(x, y) = \langle \varphi(x), \varphi(y) \rangle$  appelé noyau :

Noyau linéaire : 
$$K(x, y) = x \cdot y \quad (II.10)$$

Noyau polynomial : 
$$K(x, y) = (x \cdot y)^\alpha \quad (\alpha \in \mathbf{N}) \quad (II.11)$$

Noyau Gaussien : 
$$K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \quad (\sigma > 0) \quad (II.12)$$

Noyau sigmoïde : 
$$K(x, y) = \tanh(\alpha(x \cdot y) + \beta) \quad (\alpha > 0 \text{ et } \beta > 0) \quad (II.13)$$

Noyau RBF : 
$$K(x, y) = \exp(-\rho d(x, y)) \quad (\rho > 0) \quad (II.14)$$

$d(x, y)$  : une mesure de distance entre les vecteurs  $x$  et  $y$ .

La méthode KPCA est une généralisation de la méthode PCA qui se base sur la notion de noyau. Dans cette méthode, les données sont mappées de leur espace d'origine  $\mathbf{I}$  vers un espace  $\mathbf{H}$  ( $\dim(\mathbf{H}) \gg \dim(\mathbf{I})$ ) par une transformation non linéaire  $\varphi$ . La nouvelle matrice de covariance  $C_\varphi$  s'exprime dans  $\mathbf{H}$  par :

$$C_\varphi = \frac{1}{N} \sum_{j=1}^N \left( \varphi(x_j) - \varphi(\bar{x}) \right) \left( \varphi(x_j) - \varphi(\bar{x}) \right)^t \quad (II.15)$$

$$\text{avec, } \varphi(\bar{x}) = \frac{1}{N} \sum_{j=1}^N \varphi(x_j) = 0 \quad (II.16)$$

D'où, l'équation aux valeurs propres suivante:

$$C_\varphi V_k = \lambda_k V_k \quad \text{Avec: } k = 1, \dots, N \quad (II.17)$$

Les solutions de cette équation sont des combinaisons linéaires des  $\varphi(x_j)$ :

$$\exists \alpha_k; V_k = \sum_{i=1}^N \alpha_{ki} \varphi(x_i) \quad \text{Avec: } i = 1, \dots, N \quad (II.18)$$

D'où la forme matricielle de l'équation aux valeurs propres :

$$K^2 \alpha_k = \lambda_k N K \alpha_k \quad (II.19)$$

$$\text{Avec: } K_{ij} = K(x_i, x_j) = \varphi(x_i)^t \varphi(x_j) \quad (II.20)$$

### II.1.3. Méthode PCA à deux dimensions (2DPCA)

Proposée par Yang et al [78-79], 2DPCA est une méthode de réduction de la dimension basée sur la PCA qui représente chaque image par une matrice 2D au lieu de la transformée en un vecteur 1D. Telle représentation permet de générer des matrices de covariance de faible taille. En effet, pour un jeu de données d'apprentissage,  $X=\{x_1, \dots, x_N\} \in \mathbf{R}^{m \times n \times N}$ , formé de N images de visages, la taille de la matrice de covariance de 2DPCA prend la valeur n. Cette valeur est largement inférieure à  $n^2$  représentant la taille de la matrice de covariance de PCA.

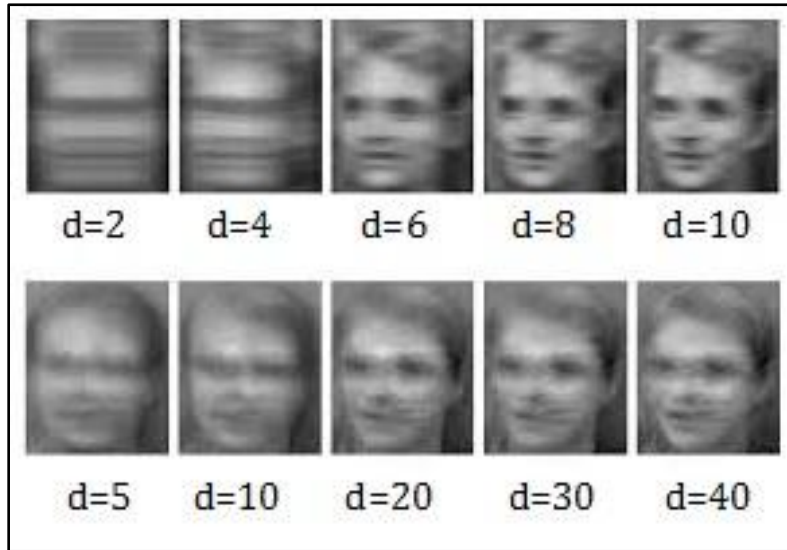


Figure II.1: Reconstruction des images par 2DPCA et PCA

En plus de ce crucial avantage, les auteurs ont rapportés que 2DPCA présente quelques avantages comparée aux méthodes: PCA, LDA, ICA et KPCA.

- Conservation de l'information temporelle et spatiale de l'image due à l'aspect 2D de la représentation de chaque image dans la matrice de covariance.
- Evaluation précise de la matrice de covariance et détermination efficace et rapide de ses éléments propres (taille de la matrice de covariance).
- Meilleure reconstruction des images d'origines en utilisant juste premières composantes principales qui emmagasine le maximum d'énergie (Figure II.1).

En effet, les images de cette figure montrent que même avec des faibles dimensions, l'apparence globale des visages obtenus par 2DPCA (images en haut) sont largement significative que celle des visages obtenus par PCA avec des dimensions très élevées (images

en bas). Enfin et suite à ses performances, 2DPCA est étendue à plusieurs variétés, y compris kernel-2DPCA [80], Improved-Modular-2DPCA [81].

## II.2. Méthodes basées sur LDA

Pour la majorité des problèmes de reconnaissance, les données d'apprentissage sont préalablement regroupées en classes, une classe par individu et plusieurs images par classe. Pour ce type de problème la méthode PCA permet de réduire la dimension de l'espace de projection des données, mais sans tenir compte de l'information intrinsèque des classes. Cependant, le principe de maximisation de la variance peut engendrer des variations intra-classe qui peuvent fausser la décision de classification. La solution la plus populaire à ce problème est la méthode LDA. En effet, LDA réalise la séparation entre les différentes classes par la détermination de l'espace de projection qui maximise les variations interclasses et minimise les variations intra-classes. Cela revient à maximiser le critère d'optimisation appelé critère de Fisher [11].

### II.2.1. Méthode des Fisher-faces

Par analogie avec les Eigen-faces, Belhumeur et al [36] ont étendu la méthode LDA pour la reconnaissance faciale en proposant la méthode des Fisher-faces. Tout comme les Eigen-faces, la méthode des Fisher-faces transforme toutes les images en vecteurs 1D. Ensuite, utilise tous ces vecteurs pour construire les matrices de dispersion intra-classes et interclasses. Pour une base d'apprentissage de  $N$  images de visages regroupées en  $c$  classes,  $p_i$  images par classe, l'algorithme des Fisher-faces requiert les étapes suivantes :

Transformer chaque image  $\gamma_i(m, n)$  en un vecteur  $x_{i(m \times n)}$

Le vecteur moyen de la base d'apprentissage :

$$\bar{x} = \frac{1}{N} \sum_i x_i \quad (II.21)$$

Le vecteur moyen pour chaque classe  $C_j$  de  $p_j$  images :

$$\bar{c}_j = \frac{1}{p_j} \sum_j x_j \quad (II.22)$$

La matrice de dispersion intra-classe :

$$S_w = \sum_{j=1}^c \sum_{x_i \in C_j} (x_i - \bar{c}_j) (x_i - \bar{c}_j)^t \quad (II.23)$$

La matrice de dispersion interclasse :

$$S_b = \sum_{j=1}^c p_j (\bar{c}_j - \bar{x})(\bar{c}_j - \bar{x})^t \quad (II.24)$$

La matrice de dispersion totale:

$$S_t = \sum_{i=1}^M (x_i - \bar{x})(x_i - \bar{x})^t \quad (II.25)$$

La base de projection optimale  $W$  doit minimiser,  $S_w$ , tout en maximisant,  $S_b$ , Cela revient à maximiser le critère d'optimisation de Fisher  $J(W)$  définis par :

$$J(W) = \frac{W^t S_b W}{W^t S_w W} \quad (II.26)$$

$$D'où l'équation: \quad S_b W = \lambda S_w W \quad (II.27)$$

Si la matrice  $S_w$  n'est pas singulière (inversible),  $W$  est une solution de l'équation aux valeurs propres suivantes :

$$S_w^{-1} S_b W = \lambda W \quad (II.28)$$

Dans ce cas, la réduction de l'espace de projection revient à déterminer la base de projection formée les « c-1 » vecteurs propres  $u_i \in \mathbf{R}^{m \times n}$  correspondant aux « c-1 » plus grandes valeurs propres de la matrice  $S_w^{-1} S_b$ . Cependant, la méthode LDA souffre principalement de deux problèmes. Le premier est lié à la non linéarité des données initiales et le deuxième est lié la singularité de la matrice  $S_w$ . Plusieurs extensions ont été proposées pour corriger ces deux problèmes dont les plus connues sont KLDA, LDA-directe et 2DLDA.

### II.2.2. Méthode LDA à noyau

La méthode KLDA est une extension de LDA basée sur les noyaux [42]. Elle consiste à projeter les données d'entrées, généralement non linéaires, dans un espace de Hilbert où les classes sont censées être linéairement séparables. D'où le critère de Fisher :

$$J(W) = \frac{V^t S_b^\varphi V}{V^t S_w^\varphi V} \quad avec \quad V = \sum_{i=1}^N \alpha_i \varphi(x_i) \quad (II.29)$$

$$S_w^\varphi = \sum_{j=1}^c \sum_{x_i \in C_j} (\varphi(\bar{c}_j) - \varphi(x_j))(\varphi(\bar{c}_j) - \varphi(x_j))^t \quad (II.30)$$

$$S_b^\varphi = \sum_{j=1}^c p_j (\varphi(\bar{c}_j) - \varphi(\bar{x}))(\varphi(\bar{c}_j) - \varphi(\bar{x}))^t \quad (II.31)$$

### II.2.3. Problème de singularité de LDA

Pour les applications de reconnaissance faciale, la dimension de la base d'apprentissage est généralement très inférieure à la dimension du vecteur représentant l'image ( $N \ll m \times n$ ). Par conséquent, la matrice  $S_w$  peut être non inversible et la maximisation de  $J(W)$  peut être possible. Ce problème est connu sous le nom du problème de la taille réduite des échantillons (*Small Sample Size, SSS*). Plusieurs solutions sont proposées à ce problème, Swets [82] a proposé de réduire les données initiales par la méthode PCA et ensuite appliquer la méthode des Fisher-faces sur les données projetées. Mais cette réduction peut supprimer l'information discriminante que contient l'espace nul de  $S_w$ . En plus, Yu et al ont proposé une méthode efficace, qui ne nécessite aucun prétraitement par PCA, appelée LDA-directe (DLDA). L'idée principale de DLDA est d'abandonner l'espace nul de  $S_b$ , qui ne contient aucune information utile, au lieu d'abandonner l'espace nul de  $S_w$ . Cela revient à maximiser le critère de Fisher en déterminant la matrice qui réduit à la fois  $S_b$  et  $S_w$  [83].

### II.2.4. Méthode LDA à deux dimensions (2DLDA)

Motivé par les travaux de Yang et al [78], Visani et al ont proposé d'étendre le principe de 2DPCA pour définir la nouvelle méthode nommée LDA à deux dimensions (2DLDA) [84-85]. Ainsi, 2DLDA utilise elle aussi une représentation matricielle au lieu de la représentation vectorielle de LDA. En effet, pour un jeu de données d'apprentissage,  $X = \{x_1, x_2, \dots, x_N\} \in \mathbf{R}^{m \times n \times N}$ , formé de  $N$  images de visages, les matrices de dispersion relatives à 2DLDA sont largement petite que celles relatives à LDA. Comparée à LDA, 2DLDA possède les avantages suivants :

- 2DLDA permet l'extraction rapide et efficace des vecteurs caractéristiques (tailles des matrices de dispersion).
- 2DLDA permet de conserver l'information temporelle et spatiale de l'image par représentation matricielle de chaque image.
- 2DLDA permet de corriger le problème de la singularité de la matrice de dispersion  $S_w$  (taille de matrice  $S_w < N$ ).

Tout comme LDA, la méthode 2DLDA a été étendue à plusieurs variétés, y compris kernel-2DLDA [86], (2D)2DLDA [87].

### II.3. Méthodes spectrales

En plus des méthodes de réduction de la dimension basées sur la réduction des endomorphismes présentées ci-dessus, d'autres méthodes peuvent assurer cette fonctionnalité en se basant sur la compression des images. En effet, la transformée en cosinus discrets (*Discrete Cosine Transform, DCT*) [88] et la transformée en ondelettes discrètes (*Discrete Wavelet Transform, DWT*) [89] sont deux méthodes d'analyse spectrales qui peuvent réduire la dimension des données initiales en utilisant un minimum de coefficients spectraux généralement situés dans les basses fréquences.

#### II.3.1. DCT à deux dimensions (2DDCT)

2DDCT est une transformée mathématique qui utilise des fonctions de base sinusoïdales non complexes pour décomposer une image  $\mathbf{J}$  en cosinus locaux suivant l'équation (II.32). Elle a effectivement la propriété d'emmagasiner le maximum d'énergie dans le minimum de coefficients spectraux de basses fréquences. Ces coefficients sont utilisés ultérieurement pour restaurer l'image  $\mathbf{J}$  à partir de sa transformée  $\mathbf{F}$  par l'opération inverse suivant l'équation (II.33). Par ailleurs, la décomposition d'une image par 2DDCT nécessite un grand nombre de fonctions de base sinusoïdales. Afin de remédier à ce problème, les chercheurs ont proposé une décomposition sous forme matricielle qui découpe l'image en blocs de même taille  $p \times q$  pixels, ensuite décomposer tous les blocs dans la même base de  $p \times q$  fonctions sinusoïdales. A titre d'exemple et pour des raisons de complexité la norme **JPEG** adopte la valeur  $p=q=8$  [90-91]. Cependant, seuls les tous premiers coefficients représentant les basses fréquences de chaque bloc sont utilisés pour construire les vecteurs descripteurs de taille assez réduite [92].

$$F(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} J(x, y) \beta(x, u) \beta(y, v) \quad (\text{II.32})$$

$$J(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \alpha(u)\alpha(v) F(u, v) \beta(x, u) \beta(y, v) \quad (\text{II.33})$$

$$\alpha(k) = \begin{cases} \sqrt{2/N} & \text{si } k = 0 \\ \sqrt{1/N} & \text{si } k > 0 \end{cases} \quad \text{et} \quad \beta(x, u) = \cos\left(\frac{(2x+1)u\pi}{2N}\right) \quad (\text{II.34})$$

#### II.3.2. DWT à deux dimensions (2DDWT)

2DDWT est une méthode d'analyse spectrale qui utilise des bancs de filtres pour décomposer l'image en quatre sous-bandes symbolisées par LL, HL, LH et HH [93]. La sous

bande LL est une approximation de l'image originale qui contient le maximum d'information spatiale, les trois autres sous-bandes reflètent les détails de l'image représentant les changements de gradient le long des trois directions, horizontale, verticale et diagonale. Par conséquent, l'approximative LL d'un niveau courant sera utilisée pour engendrer les quatre sous-bandes du niveau suivant pour transformer, après  $p$  décompositions successives, une image de taille  $(n, n)$  en sa dernière approximative de taille  $(n / 2^p, m / 2^p)$ .

### III. Classification

Pour les systèmes de reconnaissance faciale, la classification est une phase générique qui consiste à assigner un ou plusieurs individus à un corpus de classes. Cela revient à trouver une liaison fonctionnelle entre l'ensemble de vecteurs caractéristiques et l'ensemble de catégories (étiquettes, classes) selon des critères spécifiques. Les algorithmes de classification peuvent être regroupés en deux catégories à savoir, la classification par mesure de similarité et la classification par apprentissage artificiel.

#### III.1. Classification par mesure de similarité

La mesure de similarité entre deux vecteurs caractéristiques, d'l'arrière-plan et de test, est réalisée par l'une des distances suivantes :

- Distance City Block ( $L^1$ ) :

$$L^1(x, y) = \sum_i |x_i - y_i| \quad \text{Avec, } (x, y) \in R^n \times R^n \quad (II.35)$$

- Distance Euclidienne ( $L^2$ ) :

$$L^2(x, y) = (\sum_i |x_i - y_i|^2)^{\frac{1}{2}} \quad \text{Avec, } (x, y) \in R^n \times R^n \quad (II.36)$$

- Distance de Mahalanobis (Mah) :

$$Mah(x, y) = ((x - y)^t C^{-1} (x - y))^{\frac{1}{2}} \quad \text{avec, } C: \text{matrice de covariance} \quad (II.37)$$

- Distance Khi-carré ( $Chi^2$ ) :

$$Chi^2(h_1, h_2) = \sum_i \frac{|h_{1i} - h_{2i}|^2}{|h_{1i} + h_{2i}|} \quad (II.38)$$

- Intersection entre Histogrammes (HI) :

$$HI(h_1, h_2) = \sum_i \min(h_{1i}, h_{2i}) \quad (II.39)$$



## **III.2. Classification par apprentissage artificiel**

L'apprentissage artificiel est une technique de classification qui rassemble toutes les méthodes capables de générer un modèle réel à partir d'un ensemble de données d'apprentissage. Ce modèle doit être capable d'affecter chaque exemple de test à sa classe appropriée. D'une manière générale et suivant le mode d'apprentissage, les algorithmes de classification par apprentissage artificiel sont regroupés en algorithmes à apprentissage supervisé et algorithmes à apprentissage non supervisé [95]. Les réseaux de neurones artificiels (ANN), les machines à vecteurs de support (SVMs), les arbres de décision, les réseaux de Bayés et les K plus proches voisins sont parmi les principaux algorithmes de ce type de classification.

### **III.2.1. Réseaux de neurones artificiels**

Les réseaux de neurones artificiels sont des structures connexionnistes dont le développement a été motivé par l'ambition d'implémenter un système artificiel et intelligent assez comparable au cerveau humain [96]. L'application d'un réseau de neurones dans la classification doit tenir compte de la topologie du réseau (structure et fonction d'activation), de sa dynamique de propagation (en avant et/ou en arrière) et de son algorithme d'apprentissage (supervisé ou non supervisé). Plusieurs architectures de réseaux de neurones ont été proposées dont les plus connues sont : le perceptron multicouche (MLP) [97-98], le réseau de convolution (CNN) [99], le réseau de fonctions à base radial (RBF) [100] , la carte auto-organisatrice de Kohonen (SOM) [101], les réseaux récurrents [102], et probabilistes [103].

#### **III.2.1.1. Perceptron multicouche (Multi-Layer Perceptron, MLP)**

Le perceptron est connu comme premier réseau de neurones capable d'ajuster les poids des neurones de sa couche unique par apprentissage supervisé [104]. En effet, plusieurs limitations dues à l'architecture monocouche du perceptron sont évoquées par les recherches de Minsky [105]. Ces limitations sont surmontées par la conception du modèle perceptron multicouche (MLP) et l'algorithme de rétro-propagation de l'erreur [98,106]. C'est un graphe orienté, de l'entrée vers la sortie, qui organise ses neurones en couches : la couche d'entrée, des couches cachées et la couche de sortie. Partant de la couche d'entrée, chaque neurone de la couche courante est interconnecté avec tous les neurones de la couche suivante. Pour le réseau à trois

couches illustré dans la Figure II.2, la sortie  $z_i$  d'un neurone de la couche cachée est une combinaison linéaire de toutes les entrées  $x_j$ .

$$z_i = f\left(\sum_{j=0} w_{ij}^{(1)} x_j\right) \text{ où (1): neurones } i \text{ de la couche cachée} \quad (II.40)$$

En plus, la sortie  $y_k$  de la couche de sortie est une combinaison linéaire de toutes les  $z_i$ .

$$y_k = f\left(\sum_{i=0} w_{ki}^{(2)} z_i\right) \quad \text{où (2): neurones } k \text{ de la couche de sortie} \quad (II.41)$$

Enfin, la sortie  $y_k$  est exprimée en fonction des entrées  $x_j$  par l'équation :

$$y_k = f\left(\sum_{i=0} w_{ki}^{(2)} f\left(\sum_{j=0} w_{ij}^{(1)} x_j\right)\right) \quad (II.42)$$

Le processus d'apprentissage par rétro-propagation consiste à ajuster les poids de pondération de chaque neurone par la technique de descente de gradient proposée par Widrow-Hoff [104]. Cet algorithme s'exécute selon les quatre étapes suivantes:

E1 : Initialisation arbitraire des poids.

E2 : Propagation en avant de l'entrée vers la sortie et calcul des erreurs à la sortie.

E3 : Propagation en arrière de l'erreur à la sortie vers l'entrée.

E4 : Adaptation des poids au niveau des couches cachées et de la couche de sortie.

Les trois dernières étapes de ce processus s'effectuent itérativement jusqu'à ce que l'une des conditions d'arrêt soit réalisée.

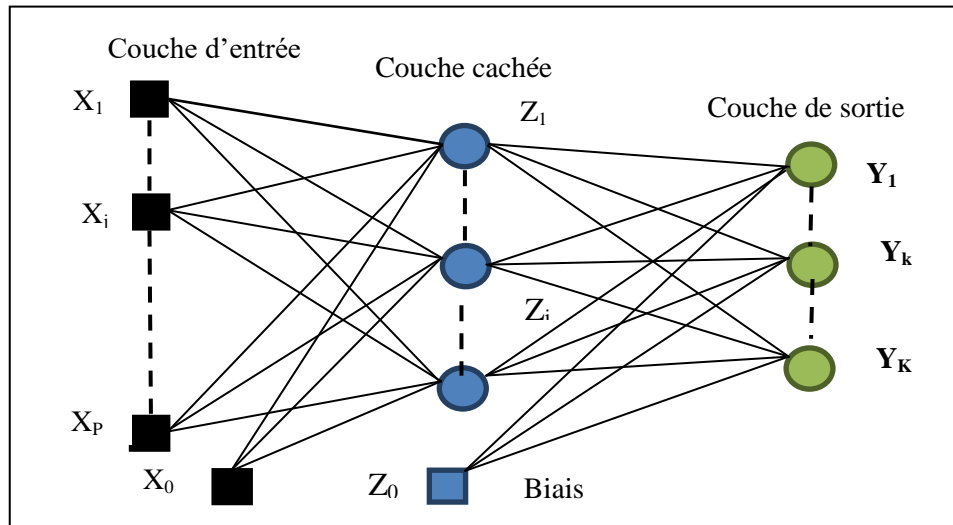


Figure II.2: Représentation du réseau MLP

Généralement, MLP présente deux défis principaux, le premier est lié à son architecture, nombre de couches cachées et nombre de neurones par couche, le deuxième est lié au rendement de la rétro-propagation. En effet, l'architecture du réseau peut être déterminée expérimentalement. Alors que l'amélioration de la rétro-propagation nécessite l'introduction de nouveaux algorithmes capables d'assurer la convergence rapide vers un minimum local [107-108]. Dans ce cadre, une nouvelle version de MLP nommée CS-MLP est mise en œuvre en améliorant la rétro-propagation par l'algorithme d'optimisation recherche des oiseaux.

### III.2.1.2. Optimisation du réseau MLP

CS-MLP est une version améliorée de MLP qui utilise un algorithme d'optimisation méta-heuristique appelé recherche des oiseaux (*Cuckoo search, CS*) [109-11]. Cet algorithme permet de surmonter les inconvénients de la rétro-propagation, principalement, le blocage des minima locaux et le faible taux de convergence. Le processus de la rétro-propagation améliorée s'exécute selon le schéma constitué des étapes suivantes :

E1 : l'algorithme CS est initialisé et transmet ses meilleurs poids à la rétro-propagation.

E2 : Lire les exemples de l'ensemble d'apprentissage.

Tant-que  $MSE < \text{critère d'arrêt}$  ( $MSE$  est l'erreur quadratique moyenne).

E3 : Initialiser tous les nids et passer leurs valeurs comme poids du réseau de neurones.

E4 : Exécuter la rétro-propagation avec les nouveaux poids et calculer l'erreur en arrière.

E5 : CS continue à calculer les meilleurs poids possibles à chaque époque jusqu'à la convergence du réseau de neurones.

Fin tant-que.

### III.2.2. Machine à vecteurs de support (SVM)

SVM est un classificateur linéaire à apprentissage supervisé, initialement introduit à la classification par Vapnik [112-113]. C'est un algorithme de classification binaire qui définit la frontière optimale séparant les deux classes par l'hyperplan qui maximise la marge entre les plus proches exemples d'apprentissage, de classes différentes, appelés vecteurs de support. De plus, SVM fait appel à des fonctions noyaux pour traiter les problèmes non linéaires. Ces fonctions projettent les données réelles d'apprentissage dans un espace de Hilbert de sorte que les données projetées sont linéairement séparables. Les principaux noyaux utilisés sont : le noyau linéaire, le noyau polynomial, le noyau gaussien et le noyau sigmoïde.

---

### III.2.2.1. Théorie des SVMs

Pour un problème de classification binaire, SVM mesure l'appartenance à l'une des deux classes via la fonction de décision définie par :

$$f(x) = \text{sign}(x^t w + b); \text{ où } w: \text{ vecteur des poids et } b: \text{ vecteur des biais} \quad (II.43)$$

Le problème de classification se réduit alors à un problème d'optimisation sous contrainte de l'ensemble d'apprentissage à l'extérieur de la marge :

$$\begin{cases} \min_x \frac{1}{2} \|w\|^2 \text{ sous contrainte:} \\ y_i(x_i^t w + b) \geq 1, \quad y_i \in \{-1, 1\} \forall i = 1, \dots, N \end{cases} \quad (II.44)$$

La solution de ce problème est réduite à la minimisation du lagrangien suivant :

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1} \alpha_i (y_i(x_i^t w + b) - 1) \quad (II.45)$$

où,  $\alpha_i$  sont des multiplicateurs de Lagrange

Ce qui mène à la maximisation du lagrangien dual par rapport à  $\alpha_i$  :

$$\begin{aligned} \max(L_D) &= \sum_{i=1} \alpha_i - \frac{1}{2} \sum_{i,j=1} \alpha_i \alpha_j y_i y_j (x_i^t y_j x_j) \\ &\text{avec } \alpha_i \geq 0 \text{ et } \sum_{i=1} \alpha_i y_i = 0 \end{aligned} \quad (II.46)$$

La fonction de décision devient donc :

$$f(x) = \text{sign}(\sum_{SVMs} y_i \alpha_i (x_i^t x) + b) \quad (II.47)$$

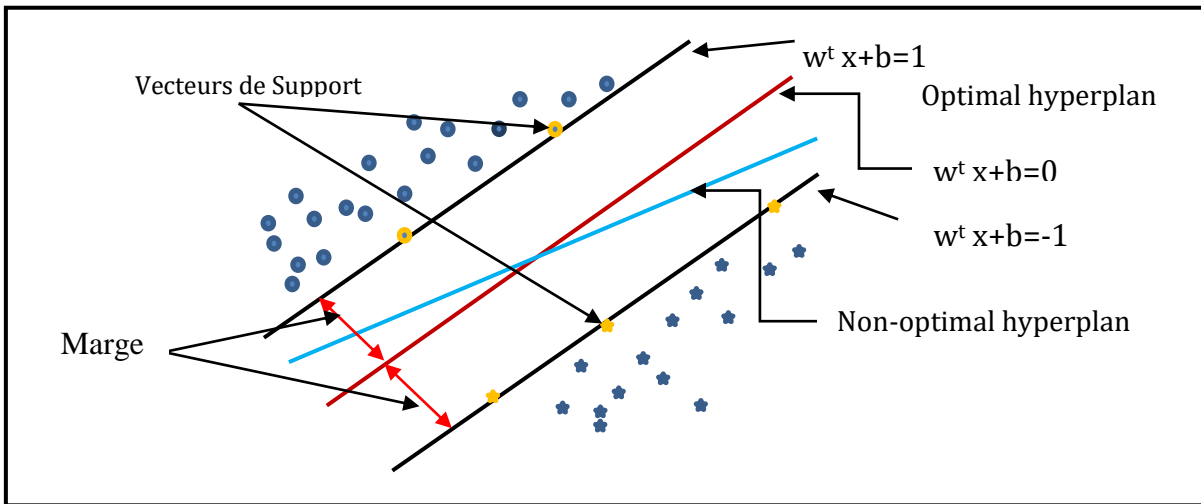


Figure II.3: Description de l'hyperplan optimal de SVM

Dans le cas d'un problème non-linéaire le lagrangien de la fonction à maximiser devient :

$$\max(L_D) = \sum_{i=1} \alpha_i - \frac{1}{2} \sum_{i,j=1} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (II.48)$$

La fonction de décision devient :

$$f(x) = \text{sign}(\sum_{i=1} y_i \alpha_i K(x_i, x) + b) \quad (II.49)$$

Généralement, les données traitées dans des applications de reconnaissance faciale sont regroupés en plusieurs classes (système multi-classes). Avant de procéder à la classification de données par SVM, ce type de problème doit être décomposé en problèmes binaires suivant l'un des trois processus suivants :

- Méthode dite un-contre-un : pour un problème de k classes, cette méthode consiste à induire  $k(k-1)/2$  sous-problèmes binaires. Pour chaque sous-problème seules deux classes du problème initial sont retenues.
- Méthode dite un-contre-tous : pour un problème de k classes, cette méthode consiste à induire k sous-problèmes binaires. Pour chaque sous-problème une des k classes est retenue, les k-1 classes restantes sont prises comme deuxième classe.
- Méthode moitié contre moitié : pour un problème de k classes, cette méthode consiste à induire  $2^{\lceil \log_2(k) \rceil} - 1$  sous-problèmes binaires par des dichotomies successives.

Plus récemment, le paquet LIBSVM est développé en tant que bibliothèque dédiée aux SVMs [114-115]. Cette librairie englobe un ensemble de fonctions implémentant des algorithmes de classification et de régression. Elle est disponible en langage C++ et en JAVA, et compatible avec diverses plateformes logicielles (Python, MATLAB, Weka et etc.).

### III.2.2.2. Optimisation Minimale Séquentielle (SMO)

Proposée par Platt, SMO est un algorithme rapide qui traite le problème de programmation quadratique des SVMs en le décomposant en plusieurs sous-problèmes à deux dimensions. Ainsi à chaque itération un sous problème est résolu en cherchant les valeurs optimales des deux multiplicateurs de Lagrange  $\alpha_i$  et en effectuant une mise à jour de la SVM pour donner le reflet des nouvelles valeurs optimisées [116-117]. En effet, l'algorithme SMO comporte trois éléments : une méthode analytique pour la résolution du problème quadrique ; deux heuristiques sont utilisées pour choisir les multiplicateurs candidats à l'optimisation ; et une méthode qui calcule le seuil b de l'équation (II.49).

### III.2.3. Les K-Plus Proches Voisins (KNN)

L'algorithme KNN est un classificateur supervisé simple et non paramétrique basé sur l'idée du vote majoritaire pour prédire la classe d'un nouvel exemple [118]. Par conséquent, la qualité de classification est améliorée par le choix de la valeur optimale de K et de la métrique adéquate pour la mesure de similarité.

### III.2.4. Les critères de mesure des performances des classificateurs

Afin d'évaluer la performance d'un classificateur, plusieurs paramètres sont utilisés, notamment, le rappel, la précision, F-mesure et l'exactitude [120]. Ces paramètres ainsi que d'autres notions d'apprentissage artificiel sont définis immédiatement.

- Validation croisée : c'est une technique qui divise, au hasard, l'ensemble des instances k échantillons de taille égale. Dans chacune des k itérations, un seul échantillon est utilisé pour le teste, les k-1 échantillons restants sont utilisés pour l'entraînement. L'avantage de cette méthode est que toutes les instances sont utilisées à la fois pour le test et pour l'arrière-plan.
- Matrice de confusion : c'est un outil servant à résumer la réussite des prédictions d'un classificateur. Chaque élément de cette matrice indique le nombre d'instances pour lesquelles la classe réelle est la ligne et la classe prédite est la colonne. Plus les valeurs des éléments diagonaux sont élevées plus les résultats de classification sont bons. Le Tableau II-1, représente la matrice de confusion à deux classes.

Tableau II-1: Matrice de confusion à deux classes

	Prédite +	Prédite -
Observée +	Vrais Positive ( <b>TP</b> )	Faux Négative ( <b>FP</b> )
Observée -	Faux Positive ( <b>FP</b> )	Vrais Négative ( <b>TP</b> )

- Précision : elle mesure la capacité du système à trouver des instances valides, et donne le pourcentage de réponses correctes parmi les résultats obtenus.

$$Précision = p = \frac{TP}{TP+FP} \quad (II.50)$$

- Rappel : il mesure la capacité du système à identifier tous les instances valides, et donne le pourcentage de réponses correctes renvoyées parmi toutes les instances pertinentes.

$$Rappel = r = \frac{TP}{TP+FN} \quad (II.51)$$

- F-mesure : elle représente la moyenne harmonique de la précision et du rappel qui mesure la capacité du système à donner toutes les solutions pertinentes.

$$F - mesure = 2pr / (p + r) \quad (II.52)$$

- Exactitude : elle désigne le pourcentage d'instances bien classées.

$$Exactitude = \frac{TP+TN}{TP+TN+FP+FN} \quad (II.53)$$

### III.2.5. Combinaison des classificateurs

Afin d'améliorer la qualité de la classification, les chercheurs ont proposé de combiner différents classificateurs au niveau des décisions (sorties) d'une manière parallèles et/ou séquentielles [121]. Dans nos travaux de thèse, nous proposons d'améliorer de la précision de classification à l'aide des méthodes parallèles, plus précisément les méthodes de vote. En effet, pour N classificateurs la décision d'un problème de classification à m classes, se fait suivant l'une des règles suivantes: vote majoritaire, moyenne des probabilités, produit des probabilités, maximum des probabilités et minimum des probabilités [122] :

- Vote majoritaire : l'individu  $X_i$  est Assigné à la classe  $C_j$  si :

$$\sum_{i=1}^N \Delta_{ji} = \max_{k=1,\dots,m} \sum_{i=1}^N \Delta_{ki} \quad \text{Où,}$$

$$\Delta_{ki} = \begin{cases} 1 & \text{si } P(c_k | x_i) = \max_{j=1,\dots,m} P(c_j | x_i) \\ 0 & \text{sinon} \end{cases} \quad (II.54)$$

- Moyenne des probabilités : l'individu  $X_i$  est Assigné à la classe  $C_j$  si :

$$\frac{1}{N} \sum_{i=1}^N P(c_j | x_i) = \max_{k=1,\dots,m} \frac{1}{N} \sum_{i=1}^N P(c_k | x_i) \quad (II.55)$$

- Produit des probabilités : l'individu  $X_i$  est Assigné à la classe  $C_j$  si :

$$\prod_{i=1}^N P(x_i | c_j) = \max_{k=1,\dots,m} P(c_k) \prod_{i=1}^N P(x_i | c_k) \quad (II.56)$$

- Maximum des probabilités : l'individu  $X_i$  est Assigné à la classe  $C_j$  si :

$$\max_{i=1,\dots,N} P(c_j | x_i) = \max_{k=1,\dots,m} \max_{i=1,\dots,N} P(c_k | x_i) \quad (II.57)$$

- Minimum des probabilités : l'individu  $X_i$  est Assigné à la classe  $C_j$  si :

$$\min_{i=1,\dots,N} P(c_j | x_i) = \max_{k=1,\dots,m} \min_{i=1,\dots,N} P(c_k | x_i) \quad (II.58)$$

## IV. Extraction des Caractéristiques

Dans le domaine de la reconnaissance faciale, l'extraction des traits remarquables du visage, telles que les yeux, le nez, la bouche et etc. reste une étape attrayante du processus de reconnaissance. Dans ce sens, plusieurs techniques ont été développées dont l'une des plus connues est le modèle binaire local (*Local Binary Pattern, LBP*). Cette section sera consacrée à ce célèbre descripteur et à ses variétés les plus récentes.

### IV.1.LBP d'origine et variétés de base

#### IV.1.1.LBP d'origine

LBP a été initialement proposé par Ojala en 1996 dans le but d'analyser les images de textures [54]. Il assigne à chaque pixel de l'image un code binaire obtenu par comparaison de sa valeur en niveau de gris avec celles de ses voisins suivant l'équation :

$$LBP_{P,R}(x_c, y_c) = \sum_{i=0}^{P-1} 2^i s(g_i, g_c) \text{ avec, } s(g_i, g_c) = \begin{cases} 1, & \text{si } g_i \geq g_c \\ 0, & \text{sinon} \end{cases} \quad (II.59)$$

$g_i$  : valeur du  $i^{\text{ème}}$  pixel voisin.  $g_c$  : valeur du pixel courant,  $S$  : fonction de décision,  $R$  : rayon du voisinage et  $P$  nombre de pixels à échantillonner.

#### IV.1.2.LBP uniforme et invariant par rotation

Dans leurs expériences, Ojala et al[55]ont constaté que seuls des motifs LBP spéciaux sont utiles lors de la description d'une image de textures. Par conséquent, Ils ont définis la variété LBP uniforme,  $LBP^{u2}$ . Dans cette variété sauf les motifs LBP dont le code binaire, vue comme une chaîne de bits circulaire, contient au plus deux transitions  $0 \leftrightarrow 1$ . Par exemple, 00011000 (2 transitions) est uniforme alors que 00010001 (4 transitions) ne l'est pas. Deux avantages de  $LBP^{u2}$  sont à signaler, la réduction du nombre des labels de l'histogramme et la discrimination des structures locales. En effet, pour un voisinage de  $P$  pixels on passe de  $2^P$  à  $P(P-1)+3$  labels distingués, et les structures locales décrites par ces motifs uniformes sont généralement des pics, des fins de lignes, des bordures, ou des coins. De même, les auteurs ont interprétés la rotation d'une image comme un déplacement des pixels sur le périmètre de leur voisinage circulaire, ils ont définis le LBP invariant par rotation, noté  $LBP^{ri}$ . La Figure II.4, illustre les 58 motifs LBP uniformes dont les neuf motifs encadrés en rouge représentent des invariants par rotation.



$$LBP_{P,R}^{u2} = \begin{cases} \sum_{i=0}^{P-1} 2^i s(g_i, g_c), & \text{si } U(LBP_{P,R}) \leq 2 \\ g_p & \text{sinon} \end{cases} \quad (\text{II.60})$$

$$U(LBP_{P,R}) = |s(g_{P-1}, g_c) - s(g_0, g_c)| + \sum_{i=1}^{P-1} |s(g_{P-1}, g_c) - s(g_i, g_c)| \quad (\text{II.61})$$

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i)\} \text{ avec, } i = 0 \dots P - 1 \quad (\text{II.62})$$

ROR(x, i) : effectue i décalages successifs, en bits, vers la droite des P bits du modèle x.

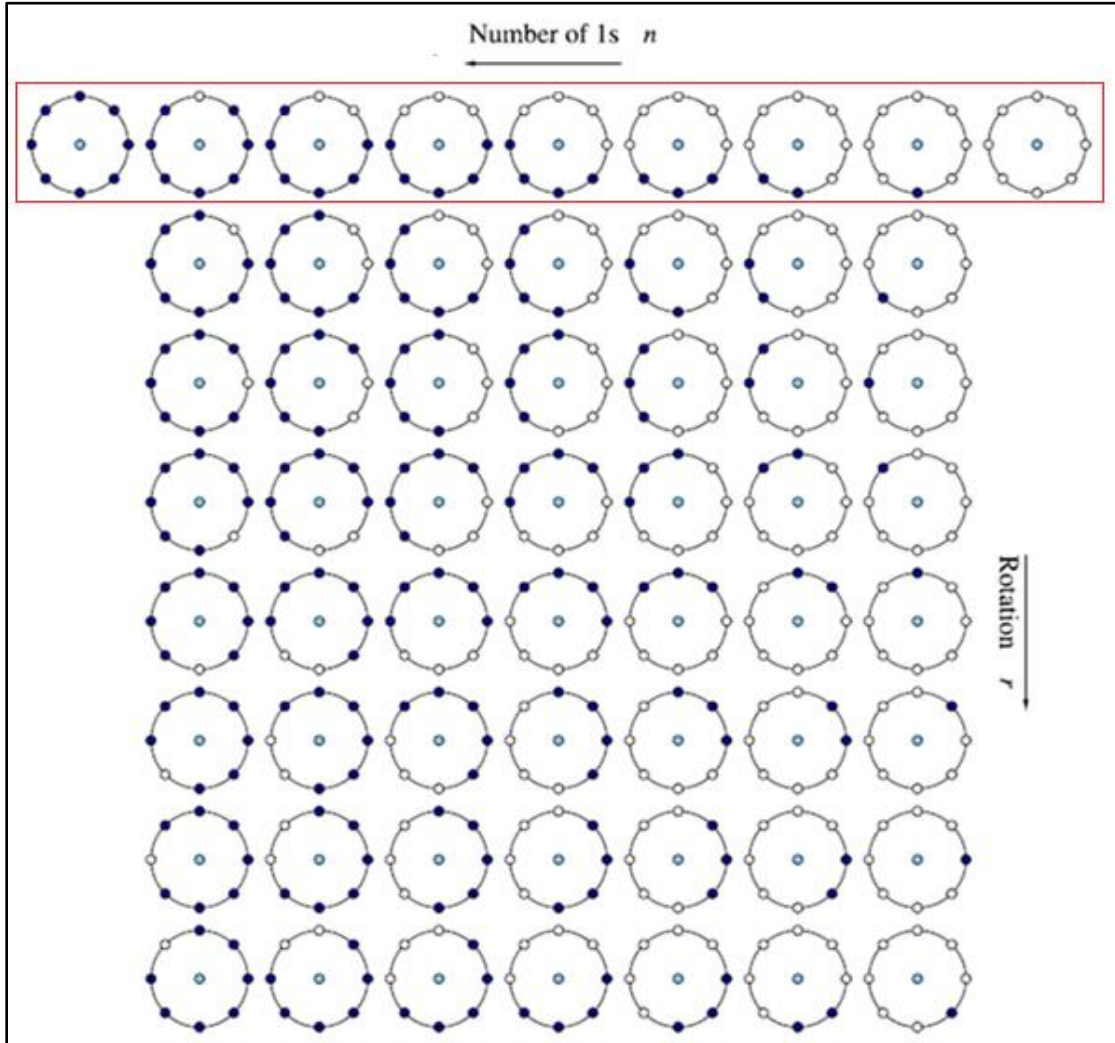


Figure II.4: Les 58 modèles uniformes de LBP<sub>8,1</sub>

## IV.2. Nouvelles variétés de LBP

Afin d'améliorer les performances du descripteur LBP, plusieurs extensions de celui-ci ont été proposées. Ces extensions portent principalement sur les paramètres caractérisant le LBP

d'origine notamment le mode de codage, la topologie du voisinage et l'aspect de la fonction de décision. Ainsi nous citons :

- Local Ternary Pattern (LTP) : pour surmonter la sensibilité de LBP aux bruits, Tan et Triggs [122] ont proposé la variété LTP. Cette variété utilise une fonction de décision à trois niveaux (+1, 0, 1) pour générer 38 motifs distingués.

$$LTP_{P,R} = \sum_{i=0}^{P-1} 3^i s(g_i, g_c) \text{ Avec, } s(g_i, g_c, t) = \begin{cases} 1, & \text{si } g_i - g_c \geq t \\ 0, & \text{sinon} \\ -1, & \text{si } g_i - g_c \leq -t \end{cases} \quad (II.63)$$

- Local Gradient Pattern (LGP) : Jun et Kim ont proposé leur nouvelle variété nommée LGP en remplaçant les huit pixels de bords par leurs gradients [123]. Cette variété a montré son efficacité vis-à-vis les différents changements d'illumination. Soit,

$$LGP_{P,R} = \sum_{i=0}^{P-1} 2^i s(G_i, G_c) \text{ avec,} \quad (II.64)$$

$$G_i = |g_i - g_c|; \quad G_c = \sum_{i=0}^{P-1} G_i \text{ et } s(G_i, G_c) = \begin{cases} 1, & \text{si } G_i \geq G_c \\ 0, & \text{sinon} \end{cases} \quad (II.65)$$

- Fuzzy-LBP (FLBP): Afin de remédier au problème de choix du seuil de segmentation, la variété FLBP remplace l'indicateur du signe de LBP par les deux fonctions d'appartenance floues suivantes [124].

$$s_{1,d}(x) = \begin{cases} 0, & \text{si } x < -d \\ 0.5 + 0.5 \frac{x}{d}, & \text{sinon} \\ 1, & \text{si } x > d \end{cases} \quad (II.66)$$

$$s_{0,d}(x) = 1 - s_{1,d}(x) \quad (II.67)$$

Avec,  $d$  paramètre de contrôle de la quantité de *fuzzification* de la fonction floue.

Plusieurs autres variétés de LBP ainsi que leurs domaines d'application (Tableau II-2) ont été proposées dans plusieurs travaux de recherches de l'état de l'art, principalement celles évoquées dans les deux livres [125-126].

Tableau II-2: Aperçu des principales variétés de LBP.

Variétés	symboles	Caractéristiques	Réfs
Improved-LBP	ILBP	Engendre $2^9$ labels en comparant les neuf pixels voisins à leur moyenne.	[127]
Median-LBP	MBP	Engendre $2^9$ labels en comparant les neuf pixels voisins à leur médiane.	[128]
Center Symmetry-LBP	CS-LBP	Permet de mieux décrire les régions d'intérêt sans utilisé le pixel du centre du voisinage.	[129]
Centralized-LBP	CBP	Mémé principe que CS-LBP à la différence de prendre en compte tous les pixels du voisinage.	[130]
Local directional pattern	LDP	Remplace les pixels de bords du voisinage par les résultants de convolution des huit masques directionnels de Kirsch.	[131]
Fuzzy-LBP	FLBP	Remplace l'indicateur du signe par deux fonctions d'appartenance floues.	[124]
Elliptical -LBP	ELBP	Capture mieux les caractéristiques micro-faciales du visage en remplaçant le voisinage circulaire par un voisinage elliptique.	[132]
Completed-LBP	CLBP	Décrit les régions locales en tenant compte de la composante centrale (CLBP_C), de la composante magnitude (CLBP_M) et de la composante signe (CLBP_S).	[133]
Local Phase Quantization	LPQ	Basé sur la quantification de la Phase de transformation de Fourier dans les voisinages locaux.	[134]
Moment-based - LBP	Mb-LBP	Utilise l'invariance des moments vis-à-vis les transformations géométriques et applique le principe de LBP aux momentgram.	[135]
Multi- block LBP	MBLBP	Multi-block LBP remplace les descripteurs de Haar dans Viola et Jones par un codage LBP entre zones rectangulaires.	[136]
Pyramidal-LBP	PLBP	Analyse spatiale par multi-résolution pyramidale.	[137]
Local Ternary Patterns	LTP	Utilise un mode de codage sur 3 bits pour corriger les effets des bruits	[122]
1DLBP	1DLBP	Utilise un voisinage 1D pour traiter les signaux sonore	[138]
3DLBP	3DLBP	Application de LBP aux images 3D	[139]

## **V. Conclusion**

Dans ce chapitre, nous avons étudié les méthodes utilisées dans les trois phases d'un système de reconnaissance faciale. Ainsi, nous avons mis l'accent sur les méthodes d'extraction des caractéristiques, les méthodes de réduction de la dimension et les algorithmes de classification. Concernant les méthodes de réduction de la dimension, nous avons focalisé les méthodes basées sur la réduction des endomorphismes (PCA, LDA, 2DPCA et 2DLDA) et celles basées sur la décomposition spectrale (2DDCT et 2DDWT). Ainsi, nous avons présenté leurs formulations, leurs caractéristiques et leurs avantages. Pour les algorithmes de classification, nous avons distingué ceux basés sur la mesure de la similarité et ceux basés sur l'apprentissage artificiel. Nous avons ciblé les MLPs et les SVMs. Ainsi, nous avons présenté leurs formulations et avantages. Finalement, nous avons présenté le célèbre descripteur, LBP ainsi que ses variétés les plus utilisées.

Le chapitre suivant sera consacré à la présentation et l'étude de notre première contribution qui s'articule principalement autour de notre descripteur LBPP.

---

---

## **Chapitre III. Modèle Probabiliste Binaire Local (Local Binary Probabilistic Pattern, LBPP) : Application à la Biométrie Faciale**

---

---

Dans le contexte image, les zones frontières qui séparent l'objet de son arrière plan, ainsi que les contours de cet objet sont des zones de grand intérêt pour le processus d'extraction des caractéristiques. Dans notre courante contribution, nous concevons un nouveau descripteur nommé modèle probabiliste binaire local (LBPP). Ce dernier localise les zones de fortes variations du gradient en se basant sur l'intervalle de confiance déterminé en fonction du degré de normalité de la distribution qui régit chaque voisinage de l'image. L'intégration de ce descripteur à la reconnaissance faciale a montré la robustesse de la description locale du visage via l'évaluation par intervalle de confiance.

## **I. Introduction**

Les processus aléatoires sont des outils précieux qui contribuent à la compréhension de plusieurs phénomènes naturels, dont la description nécessite la décomposition en entités élémentaires régie par des distributions empiriques. Ces distributions sont généralement approximées par l'une des lois de probabilités usuelles, particulièrement la loi normale. Dans ce contexte, nous utilisons la distribution qui régit les pixels de chaque voisinage de l'image pour définir notre nouveau modèle d'extraction de caractéristiques locales appelé modèle probabiliste binaire local (Local Binary Probabilistic Pattern, LBPP). Ce descripteur est une variété probabiliste du célèbre descripteur LBP qui s'appuie sur des résultats importants de la théorie des probabilités, y compris la loi des grands nombres et le théorème central limite. A fin d'engendrer une description faciale robuste et fiable, LBPP étudie la convergence en loi de la distribution qui gouverne chaque voisinage, détermine son intervalle de confiance et recode son pixel centrale sur la base de cet intervalle de confiance. Cette vision permet de pallier certains défauts de LBP d'origine, en particulier les pertes d'informations associées à sa nature binaire et déterministe. Pour mettre en évidence les performances de LBPP nous avons mis en place un système de reconnaissance faciale où ce dernier est hybridé avec des méthodes de réduction de la dimension notamment 2DDCT, 2DPCA et 2DLDA, ainsi que des algorithmes de classification par apprentissage y compris KNNs, MLPs et SVMs. Les résultats expérimentaux obtenus à l'aide des bases de données ORL, Yale et AR démontrent l'efficacité et la robustesse de nos différentes approches.

## **II. Test de normalité et moments statistiques**

Parmi plusieurs distributions observées dans la nature, la distribution des niveaux de gris dans une zone locale quasi-homogène de l'image, spécialement une image de visage, suit approximativement une loi normale. Cette estimation fait appel à des paramètres statistiques appelés moments statistiques. En effet, la loi normale est la loi de probabilité la plus pratique et la plus qualifiée pour estimer le comportement de plusieurs phénomènes naturels. Elle est complètement définie par la donnée de la densité de probabilité  $f$  et/ou la fonction de répartition  $F$ , dépendant principalement de la moyenne  $\mu$  et de l'écart type  $\sigma$ .

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right] \quad (\text{III.1})$$

$$F(x) = \int_{-\infty}^x f(t) dt \quad (\text{III.2})$$

La plus remarquable caractéristique de la loi normale réside dans la définition des intervalles relatifs aux pourcentages des valeurs observées. Ces intervalles sont appelés intervalles de confiance. Parmi les principaux intervalles de confiance sont ceux illustrés dans la Figure III.1 et qui valent :

- 68.27% des valeurs observées sont situées à  $+\sigma$  /  $-\sigma$  de la moyenne
- 95.45% des valeurs observées sont situées à  $+2\sigma$  /  $-2\sigma$  de la moyenne
- 99.73% des valeurs observées sont situées à  $+3\sigma$  /  $-3\sigma$  de la moyenne
- 99.99% des valeurs observées sont situées à  $+4\sigma$  /  $-4\sigma$  de la moyenne

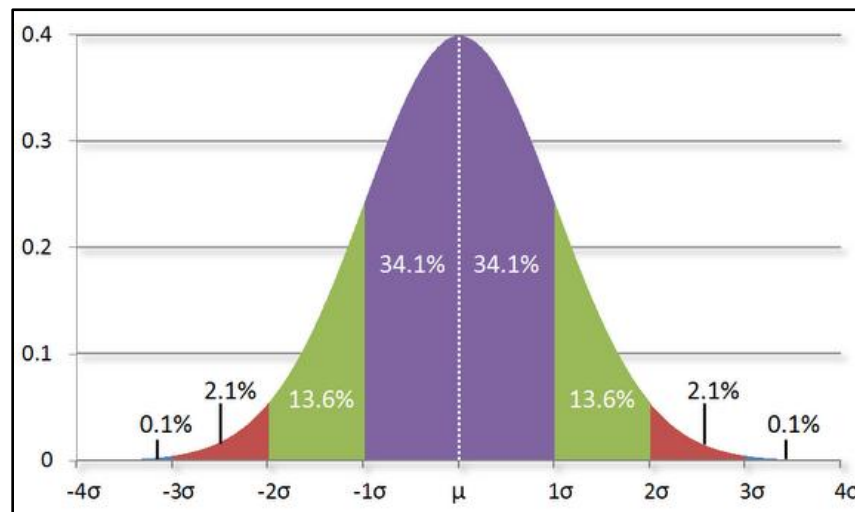


Figure III.1 : Différents intervalles de confiance de la loi normale

## II.1. Moments Statistiques

### II.1.1. Tendances Centrale et Dispersion

Pour une population de  $n$  individus  $x = \{x_1, x_2, \dots, x_n\}$ , les moments d'ordre un et deux mesurent respectivement la tendance centrale et la dispersion d'une distribution. En effet, la tendance centrale est définie comme étant la valeur spécifique qui résume les  $n$  individus de

---

cette population. Elle est généralement mesurée à l'aide du mode ( $M_{od}$ ), de la moyenne ( $\mu$ ) ou de la médiane ( $M_{ed}$ ).

$$M_{od} = x \text{ tel que } \max_i (P(X = x_i)) \quad \forall i = 1 \dots n \quad (\text{III.3})$$

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad \forall i = 1 \dots n \quad (\text{III.4})$$

$$M_{ed} = x_{\left(E\left(\frac{n}{2}\right)\right)} \quad \forall i = 1 \dots n \quad (\text{III.5})$$

Où,  $E(p)$  partie entière de  $p$  et  $x_{(i)}$  la population ordonnée

En plus, l'écart-type est un coefficient qui mesure la dispersion d'une distribution comme étant la moyenne quadratique des écarts des individus par rapport à leur moyenne. En outre, un autre coefficient de dispersion qui peut être, parfois, préféré à l'écart type ( $\sigma$ ) est le coefficient de variation ( $\delta$ ). C'est une valeur sans unité rapportant l'écart-type à la moyenne.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{(x_i - \mu)^2}{n}} \quad \forall i = 1 \dots n \quad (\text{III.6})$$

$$\delta = \frac{\sigma}{\mu} \quad (\text{III.7})$$

### II.1.2. Asymétrie et Aplatissement

En plus des moments d'ordre un et deux, plusieurs moments d'ordre supérieurs peuvent contribuer à l'étude du comportement d'une population. En particulier, le moment d'ordre trois (*Skewness*,  $S_k$ ) et le moment d'ordre quatre (*Kurtosis*,  $K_u$ ) qui mesurent respectivement la symétrie et la courbure d'une distribution.

$$S_k = \frac{1}{n\sigma^3} \sum_{i=1}^n (x_i - \mu)^3 \quad \forall i = 1 \dots n \quad (\text{III.8})$$

$$K_u = \frac{1}{n\sigma^4} \sum_{i=1}^n (x_i - \mu)^4 \quad \forall i = 1 \dots n \quad (\text{III.9})$$

Pour une loi normale, qui est parfaitement symétrique, le coefficient d'asymétrie est nul et celui d'aplatissement est égal à 3. Cette propriété est utilisée pour juger grossièrement la normalité d'une distribution empirique. En effet, la tendance asymptotique d'une distribution empirique  $D(\mu, \sigma)$  vers une loi normale  $N(\mu, \sigma)$  augmente avec des valeurs de *skewness* très proche de 0 et des valeurs de *kurtosis* très proche de 3. En effet, la Figure III.2 représente respectivement les courbes correspondantes aux valeurs de *skewness* (à gauche) et de *kurtosis*

---



(à droite). Pour les courbes *skewness*, nous remarquons que : pour une loi normale  $S_k=0$ , si la distribution est positivement asymétrique  $S_k>0$  et  $S_k<0$  si cette distribution est négativement asymétrique. En conséquence, le nombre des niveaux de gris de part et d'autre de la moyenne ainsi que leurs fréquences d'apparition est significatif soit à droite soit à gauche suivant la valeur de  $S_k$ . Tandis que, les courbes *kurtosis*, montrent que lorsque  $K_u=3$ , la distribution est une loi normale. En revanche,  $K_u<3$  indique que la distribution est plus aplatie que la loi normale (platykurtique), alors que  $K_u>3$  indique que la distribution est effilée (leptokurtique). En fait, le nombre des niveaux de gris de part et d'autre de la moyenne ayant une même fréquence d'apparition augmente en diminuant  $K_u$  et vice-versa.

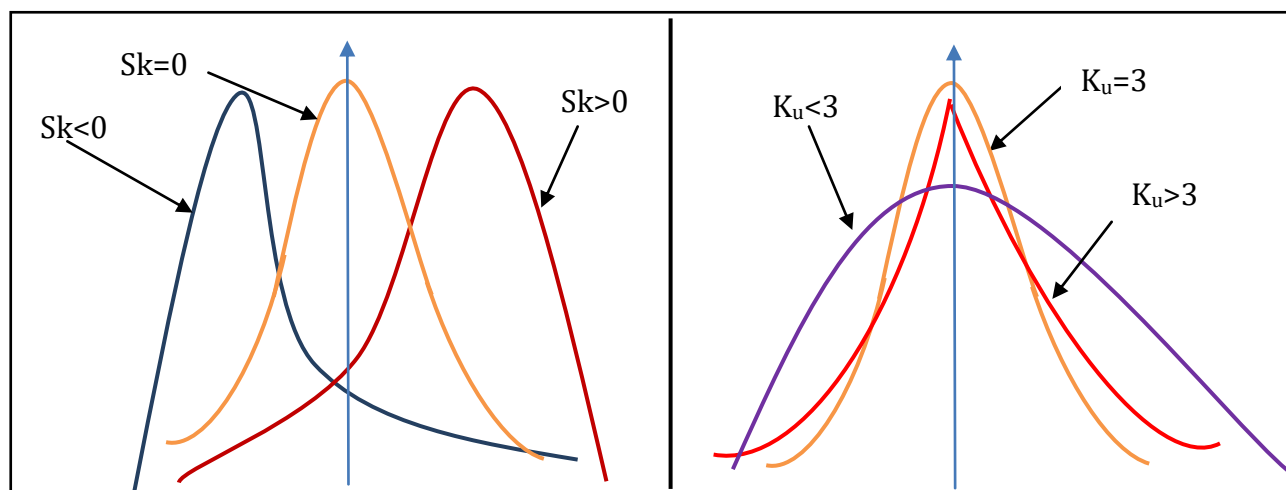


Figure III.2 : Skewness et Kurtosis de quelques distributions

## II.2. Test de Normalité de Jarque–Bera (JB-test)

En analyse statistique, l'hypothèse selon laquelle une distribution empirique est asymptotiquement normale représente un outil efficace pour l'étude de plusieurs populations naturelles. Afin de valider cette hypothèse plusieurs tests de normalité ont été proposés, spécialement le test de Jarque–Bera (JB-test) [140-142]. En effet, ce test évalue la normalité approximative d'une population, de  $n$  individus, en regroupant les deux coefficients  $S_k$  et  $K_u$  suivant la statistique définie par :

$$JB - test = \frac{n}{6} \left( S_k + \frac{(K_u - 3)^2}{4} \right) \quad (III.10)$$

En outre, plus la valeur de cette statistique est élevée, plus l'écart entre la distribution empirique et la loi normale est manifeste. Le Tableau III-1 montre que la valeur critique  $T_{JB}$  dépend de la taille  $n$  et de l'ordre du risque  $\alpha$  de la population. La condition de normalité asymptotique est définie par :

$$JB - test > T_{JB}, \text{ où } T_{JB} \text{ est une valeur critique} \quad (III.11)$$

Tableau III-1: Valeurs critiques  $T_{JB}$  pour différentes tailles  $n$  et différents risque  $\alpha$ .

$\alpha$	n=10	n=20	n=25	n=50	n=100	n=200	n=500	n $\rightarrow$ $\infty$
0.01	5.738	9.458		12.331	12.296	11.750	10.601	9.219
0.02	4.274	6.583		8.721	9.089	8.788	8.349	7.824
0.05	2.535	3.768	4.15	5.004	5.448	5.728	5.825	5.991
0.10	1.618	2.335		3.192	3.643	4.081	4.324	4.605
0.20	1.125	1.556		2.122	2.474	2.748	2.985	3.219

### III. Modèle Probabiliste Binaire Local (LBPP)

#### III.1. Présentation de LBPP

L'idée principale du descripteur LBPP se base sur l'hypothèse suivante : « Dans une image, en particulier de visage, chaque voisinage local de l'image est régi par une distribution empirique qui peut converger asymptotiquement vers une loi normale ». Cette convergence est évidente avec les différentes conditions de normalité. En effet, tout voisinage remplissant ces conditions de normalité est qualifié de zone quasi-homogène. En se basant sur cette proposition, nous définissons notre nouvelle variété LBPP du descripteur LBP. Cette variété évalue chaque pixel de l'image en fonction de l'intervalle de confiance déterminé suivant la normalité ou la non-normalité de la distribution empirique qui caractérise son voisinage. Cette vision permet de passer d'un mode de seuillage déterministe de LBP à un mode de seuillage probabiliste pour LBPP. Ainsi, ceci permet de séparer les zones *quasi-homogènes* présentant de faibles fluctuations et les zones *de pics* présentant de fortes variations du gradient. Afin de confirmer cette assertion nous analysons les histogrammes de différentes zones d'une image LBPP de visage. La Figure III.3 illustre deux histogrammes, le premier correspond à une zone de pic prise de la région de la bouche et le deuxième représente une zone presque homogène

du front à proximité de l'œil droite. En effet, l'histogramme de la Figure III.3-(b) montre que presque tous les pixels de la zone quasi-homogène choisie sont proches du niveau de gris 255 (666 valeurs du mode parmi 750). En plus, ils sont repartis autour de leur moyenne avec un faible écart-type  $\sigma=19.27$  et un faible coefficient de variation  $\delta=0.077$ . Par contre, l'histogramme de la Figure III.3-(a) représentant la zone de pic choisie montre que les pixels sont étalés sur toute la plage  $0 : 255]$  (59 valeurs du mode parmi 750). En plus, ses pixels sont distribués avec un écart-type élevé  $\sigma=71.93$  et un coefficient de variation élevé  $\delta=0.41$ . Ceci affirme l'hypothèse des zones quasi-homogène déjà annoncée.

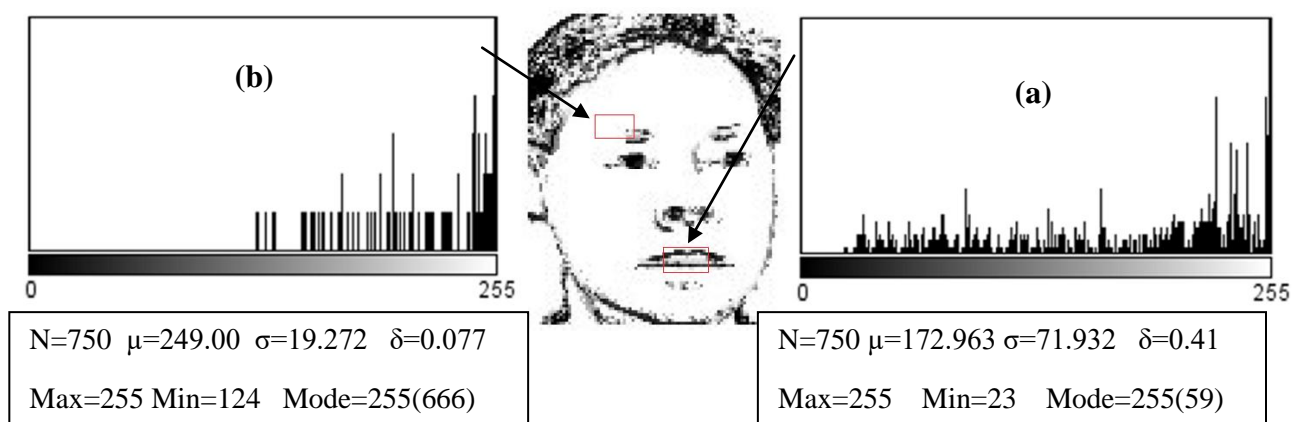


Figure III.3 : Histogrammes des zones de l'image LBPP

Partant de cette hypothèse, chaque pixel de l'image est recodé à l'aide d'un intervalle de confiance déterminé à partir de la distribution qui régit son voisinage selon la procédure détaillée dans la section suivante.

### III.2. Évaluation de l'intervalle de confiance

Pour coder un pixel courant avec notre descripteur LBPP, nous devons estimer l'intervalle de confiance  $[\alpha_1, \alpha_2]$  qui caractérise son voisinage  $5 \times 5$  pixels. Cela revient à mesurer le degré de normalité de la distribution empirique qui régit ce voisinage. Plusieurs tests sont proposés pour estimer la normalité d'une distribution empirique ; notamment ceux basés sur les moments statistiques. Dans notre cas, nous proposons de mesurer l'écart d'une distribution empirique  $D(\mu, \sigma)$  à la loi normale  $N(\mu, \sigma)$  en apposant des conditions sur le coefficient d'asymétrie  $S_k$  et le coefficient de variation  $\delta$  et/ou la valeur critique de Jarque–Bera ( $T_{JB}$ ). Dans ce contexte, chaque image est considérée comme étant une réalisation d'un champ

aléatoire où chaque pixel est vu comme une variable aléatoire discrète. Par conséquent, chaque voisinage remplissant les conditions de l'équation (III.12) sera considéré comme zone quasi-homogène et son intervalle de confiance sera défini par l'équation (III.13). Les voisinages qui ne remplissent pas ces conditions sont qualifiés de zone de pic et leur intervalle de confiance sera défini par l'équation (III.14).

$$Si \{(S_k = 0 \text{ ou } \delta < \beta) \text{ ou bien } JB > TJB\} \quad (III.12)$$

$$Alors, [\alpha_1, \alpha_2] = [\mu - k \cdot \sigma, \mu + k\sigma] \text{ où } k = 1,2,3,4 \quad (III.13)$$

$$Sinon, [\alpha_1, \alpha_2] = [\mu - K, \mu + K]; \text{ où } K = \begin{cases} \sigma & \text{si } \beta < \delta < 0.2 \\ 0.2 \mu & \text{ailleurs} \end{cases} \quad (III.14)$$

Avec,  $0.1 < \beta < 0.2$  est une valeur expérimentale et  $3.768 < T_{JB} < 5.004$  est une valeur tabulée

Compte tenu de ce qui précède, nous concluons que l'intervalle de confiance d'un voisinage est principalement influencé par la forme de la courbe de la distribution qui le régit. Ainsi, pour un voisinage quasi-homogène, cette courbe est très proche d'une gaussienne et ses pixels sont dispersés autour de leur moyenne dans un intervalle de taille  $8\sigma$  avec un degré de confiance de 99.99%. Plus cette courbe devient aplatie ou dissymétrique plus la taille de l'intervalle de confiance tend vers  $2\sigma$ . Par contre, si les éléments du voisinage sont assez dispersés, la taille de l'intervalle de confiance prend une valeur limite de  $0,4\mu$ .

### III.3. Formulation de LBPP

En se basant sur les intervalles de confiance définis par les équations (III.12),(III.13) et (III.14) le descripteur LBPP évalue la valeur du pixel courant en utilisant le processus suivant : « chacun des huit pixels voisins ayant une valeur située dans l'intervalle de confiance  $[\alpha_1, \alpha_2]$  se code par '1'. Et les pixels qui restent se codent par "0" ». Le mot de code binaire ainsi obtenu est ensuite convertit en niveau de gris à l'aide de l'équation suivante :

$$LBPP_{P,R,P',R'} = \sum_{n=0}^{P-1} s(i_n) 2^n \quad (III.15)$$

$$\text{où, } s(x) = \begin{cases} 1, & \text{si } \alpha_1(\mu, \sigma, k) \leq x \leq \alpha_2(\mu, \sigma, k) \\ 0, & \text{ailleurs} \end{cases} \quad (k \in [-4,4]) \quad (III.16)$$

R et P sont respectivement le rayon de voisinage et le nombre de pixels voisins qui contribuent au calcul du code binaire. R' et P' sont respectivement le rayon de voisinage et le nombre de pixels voisins qui contribuent au calcul des moments statistiques ( $\mu$ ,  $\sigma$ ,  $\delta$ ,  $S_k$ ,  $K_u$  et  $T_{JB}$ ).

Comme exemple d'un tel processus de codage la Figure III.4 illustre le résultat obtenu pour un voisinage défini par les valeurs suivantes : R=1, P=8, R'=2 et P'=25.

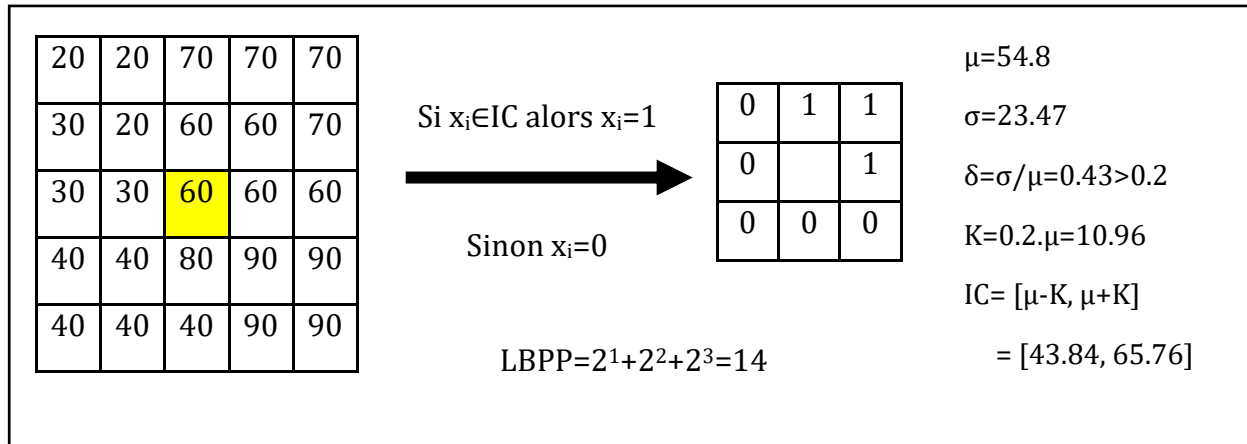


Figure III.4 : Description d'un pixel par la méthode LBPP

### III.4. Les avantages de LBPP

Contrairement à la majorité des variétés de LBP qui sont généralement des modèles déterministes, l'approche LBPP considère les zones de visage comme des distributions empiriques de pixels. En outre, la mesure de l'écart de ces distributions par rapport à la loi normale permet de définir une nouvelle description probabiliste de l'image. Dans cette description, le concept d'intervalle de confiance permet de localiser les zones de fort gradient qui sont caractérisées par des changements brutaux de l'intensité. Par ailleurs, cette description du visage est robuste aux changements d'éclairage et aux bruits. En effet, la Figure III.5-(a) montre que les codes générés par LBP et LBPP sont les mêmes lorsque le changement d'éclairage affecte le fond et le premier plan en même temps. Par contre, les exemples de Figure III.5-b et Figure III.5-c montrent que seul le code LBPP qui demeure invariant avec les changements d'éclairage affectant séparément le fond et/ou le premier plan. De plus, LBPP(k=4) présente d'autres avantages : il est robuste vis-à-vis le changement de contraste, le balancement des couleurs, le changement de la brillance, le bruit gaussien et le bruit de flou.

En outre, la Figure III.6 représente les résultats de l'application du descripteur LBPP(k=4) sur différentes images d'entrées. Ces dernières représentent une image originale d'un visage et les images obtenues par des opérations de changement de couleur et d'éclairage, ainsi que des additions des bruits gaussien et flou. Cette figure montre que LBPP(k=4) décrit d'une manière assez comparable l'image d'origine et toutes les images résultantes de l'application de différents changements précédentes. En effet, il est visiblement claire que dans toutes les images obtenues par LBPP(k=4), les zones de pics représentant les principaux composants du visage sont facilement extraites car elles ont été fortement différenciées des zones quasi-homogènes.

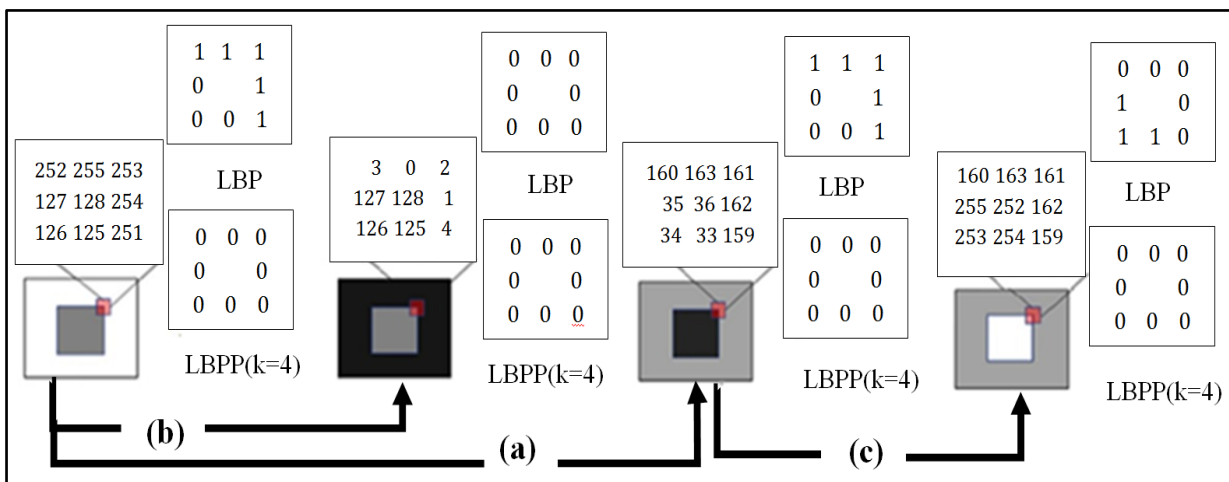


Figure III.5 : Comparaison entre LBP et LBPP pour des changements d'éclairage

En outre, la Figure III.6 représente les résultats de l'application du descripteur LBPP(k=4) sur différentes images d'entrée. Ces dernières représentent une image originale d'un visage et les images obtenues par des opérations de changement de couleur et d'éclairage, ainsi que des additions des bruits gaussien et flou. Cette figure montre que LBPP(k=4) décrit d'une manière assez comparable l'image d'origine et toutes les images résultantes de l'application de différents changements précédentes. En effet, il est visiblement claire que dans toutes les images obtenues par LBPP(k=4), les zones de pics représentant les principaux composants du visage sont facilement extraites car elles ont été fortement différenciées des zones quasi-homogènes.

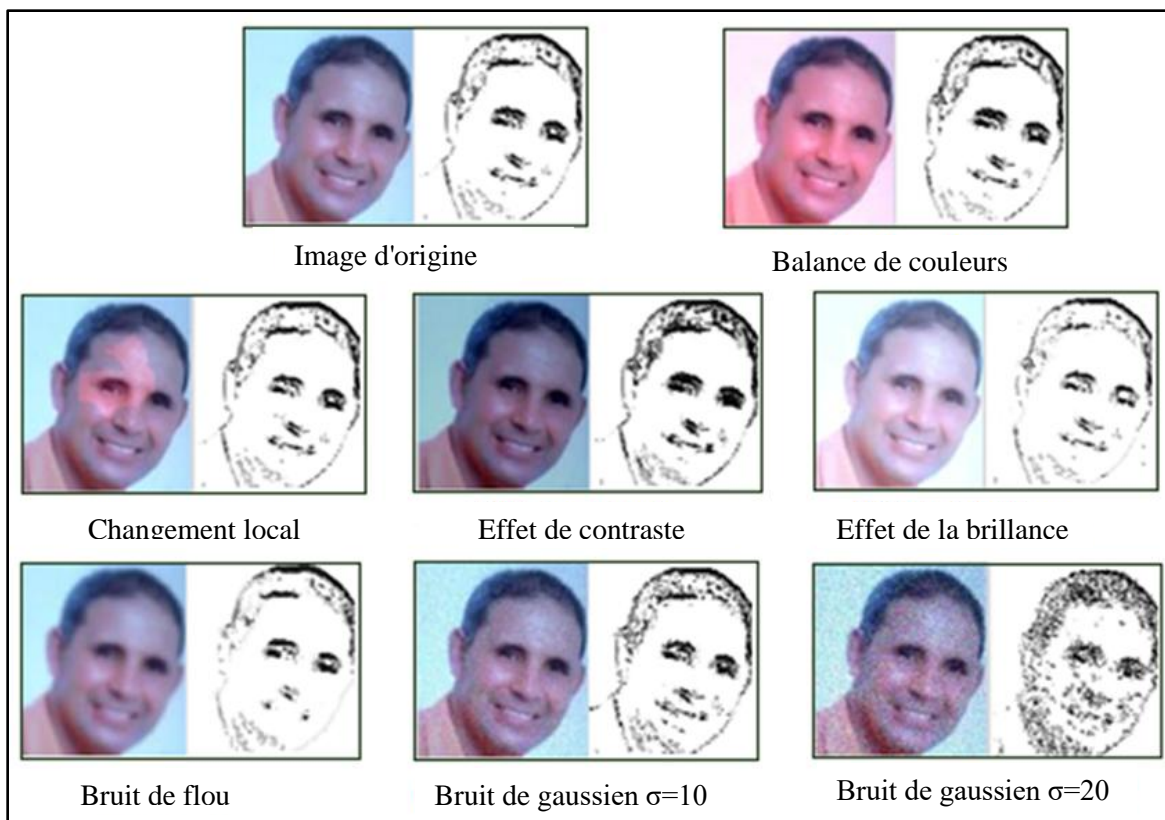


Figure III.6 : Résultats LBPP(k=4) avec changements d'éclairage et bruits

## IV. Reconnaissance faciale par LBPP

### IV.1. Système de reconnaissance faciale proposé

Dans cette sous-section, nous proposons un système de reconnaissance faciale (Figure III.7) qui utilise la nouvelle description de visage obtenue par LBPP. Ce système fonctionne en trois phases : initialement, chaque visage détecté et normalisé sera traité par LBPP afin de localiser ses principaux composants. Ensuite, les méthodes de réduction de la dimension y compris PCA, LDA, 2DPCA, 2DLDA et 2DDCT engendrent un ensemble de vecteurs caractéristiques de faibles dimensions. Cet ensemble sera sujet de classification à l'aide des algorithmes : KNN, LIBSVM, SMO et MLP (présentés au chapitre 2).

#### IV.1.1. Prétraitement

Dans cette phase deux opérations sont successivement réalisées : le recadrage géométrique et l'égalisation d'histogramme. La première élimine les composantes non pertinentes de

l'arrière-plan pour ne pas entraver la performance globale du système, la deuxième égalise l'équiprobabilité de tous les niveaux de gris de l'image.

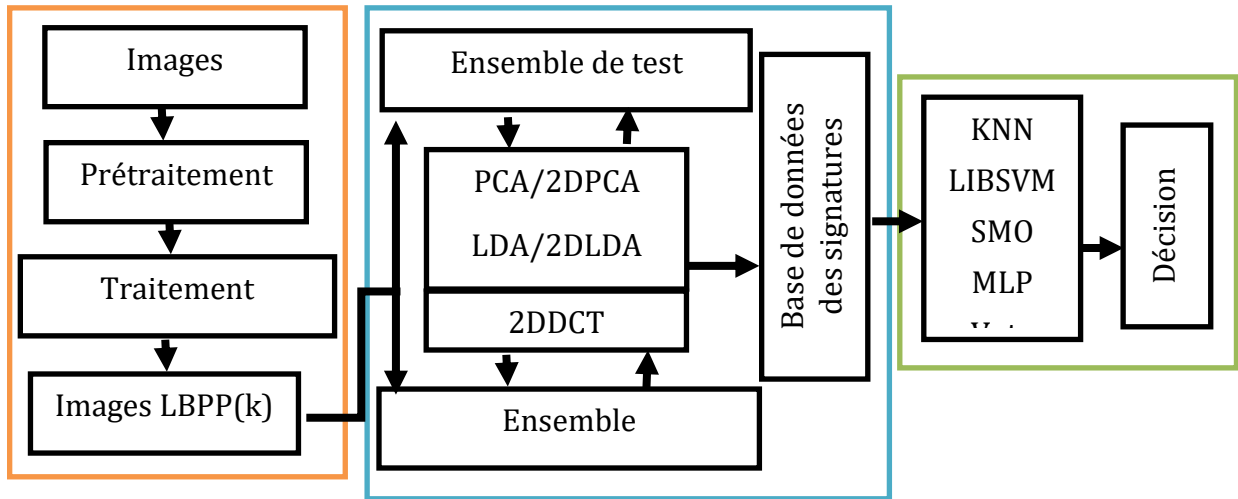


Figure III.7 : Diagramme blocs du système de reconnaissance

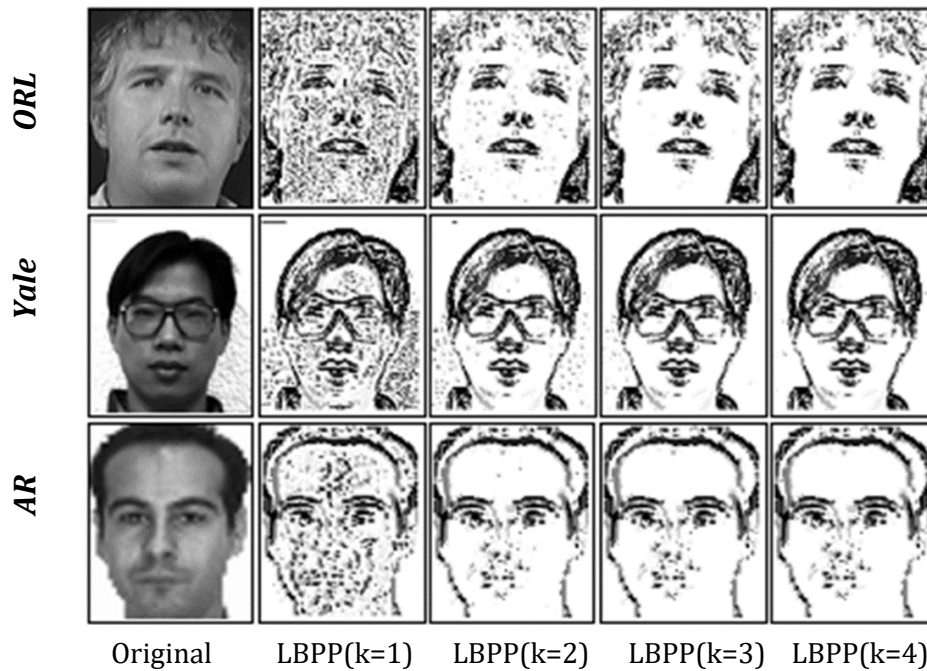


Figure III.8 : Images LBPP(k = 1,2,3,4) des bases de données ORL, Yale et AR

#### IV.1.2. Traitement par LBPP

Généralement, les yeux, le nez et la bouche sont les éléments les plus descriptifs de la topologie du visage. Par conséquent, la sélection de ces composants est une étape décisive dans le processus de reconnaissance. L'idée principale de LBPP consiste à attribuer des codes



similaires aux zones quasi-homogènes et des codes différents aux zones de pics. Ces dernières seront de plus en plus localisées en augmentant la taille de l'intervalle de confiance. Cette propriété est mise en évidence à l'aide des images issues des trois bases de données ORL, Yale et AR. La Figure III.8 montre les images résultantes de l'application de LBPP avec des intervalles de confiance déterminées par  $k=1, 2, 3$  ou  $4$ . Dans ces images les principaux composants des visages sont évidemment distingués tout en allant de  $k=1$  vers  $k=4$ .

#### IV.1.3. Réduction de la dimension

Il est évident que la grande dimension des images de visages affecte la complexité de calcul et la performance globale des systèmes de reconnaissance. Pour pallier ce problème, nous utilisons des méthodes de réduction de la dimension par projection et des méthodes de réduction de la dimension par compression.

- Méthodes de réduction par projections : elles engendrent des représentations de faible dimension en projetant les images d'origines, de haute dimension, sur des bases de projection propres. Ces bases de projection sont formées par les vecteurs propres d'une matrice spécifique. Dans cette catégorie, nous utilisons principalement les méthodes suivantes : PCA, LDA, 2DPCA, 2DLDA.
- Méthode de réduction par compression (2DDCT) : elle représente chaque image par un vecteur caractéristique de taille  $n_b \times b$ . où,  $n_b$  représente le nombre de blocs non chevauchants ;  $b$  est le nombre des premiers coefficients de la transformée 2DDCT (*K-First Neighboring coefficients of Direct Current, KFNC-DC*). La Figure III.9 montre un exemple de construction d'un vecteur caractéristique 2DDCT. Ainsi, l'image LBPP( $k=4$ ) est subdivisée en  $n_b$  blocs de  $p \times p$  pixels, chaque bloc est transformé en  $p \times p$  coefficients spectraux. Ensuite, seuls les  $b$  premiers coefficients de chaque bloc sont concaténés par la technique du zigzag pour construire son vecteur local. Enfin, les vecteurs locaux de tous les blocs sont concaténés en un seul vecteur global formant le vecteur descripteur de l'image. En outre, la taille du vecteur descripteur d'une image dépend des valeurs des deux paramètres  $n_b$  et  $b$ . Ces valeurs dépendent elles aussi du degré d'homogénéité de l'image. En effet, plus l'image est localement homogène plus la taille des blocs est grande et le nombre de coefficients significatifs est faible et vice-versa. Cette propriété est en parfaite harmonie avec le principe de LBPP. Pour affirmer cette assertion, nous avons appliqué

2DDCT à deux zones locales d'une image LBPP(k=4). La première zone est quasi-homogène (front), l'autre de pic (œil). Les résultats illustrés dans la Figure III.10 montrent que : pour la zone quasi-homogène seule la composante DC (b=1) est non nulle. Alors que pour la zone de pic tous les coefficients sont non nuls, mais seuls les b premiers sont significatifs ( $1 < b < 10$ ).

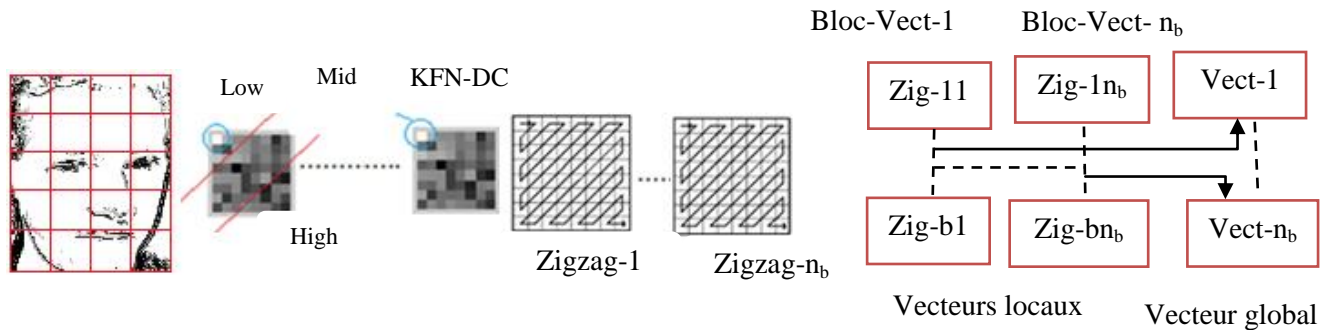


Figure III.9 : Différentes étapes de la méthode 2DDCT

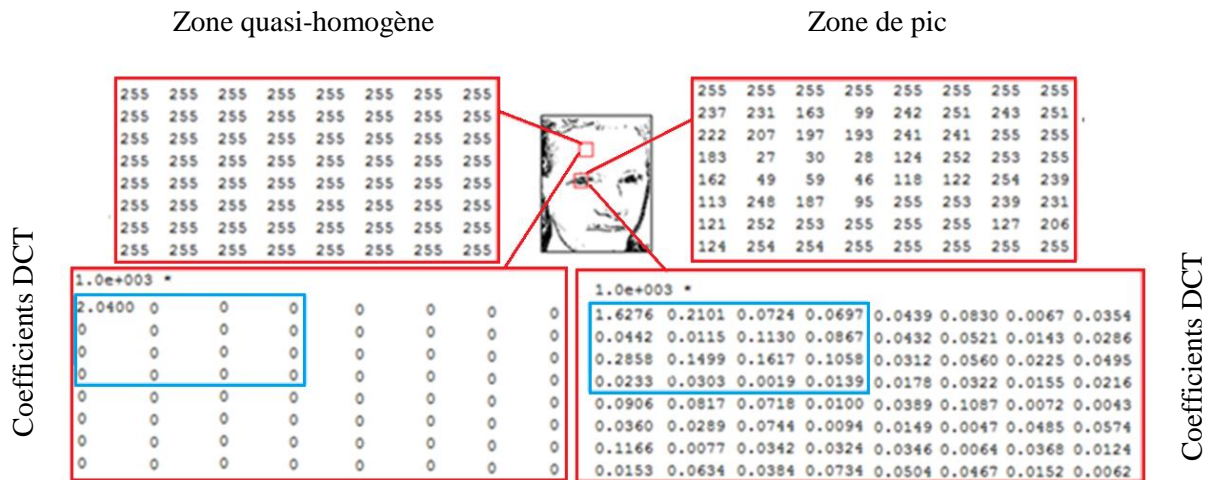


Figure III.10 : Description de différentes zones de visage par 2DDCT

#### IV.1.4. Classification

Dans cette phase, pour chaque base de données les vecteurs réduits précédemment engendrés représentent l'ensemble des données d'entrée pour les algorithmes de classification. Généralement, les performances de ces algorithmes sont évaluées à l'aide des paramètres suivants : le taux de reconnaissance, le temps d'exécution, le taux d'exactitude, la précision, le rappel et la F-mesure.

## IV.2. Résultats expérimentaux

Dans cette sous-section, nous évaluons notre système de reconnaissance de visages à travers une série d'expériences effectuées sur les trois bases de données ORL, Yale, et AR. Dans chaque phase de notre système nous testons plusieurs méthodes afin de déterminer la meilleure d'entre elles. Le long du processus global de reconnaissance, nous déterminons la combinaison qui donne la meilleure performance en jouant sur les valeurs des paramètres critiques de chaque méthode. En effet, l'intervalle de confiance, la dimension de la base de projection, la taille des blocs et le nombre des KFNC-DC éléments sont des variables caractérisant respectivement le descripteur LBPP et les méthodes de réduction de la dimension. Deux phases expérimentales sont mises en place. La première consiste à réaliser des expériences pour évaluer les paramètres de performance où chaque expérience est répétée 100 fois. Dans chaque répétition, les images de chaque base de données appartenant à  $n_c$  classes sont aléatoirement subdivisées en un ensemble d'entraînement et un ensemble de test. Le premier comporte  $n_{tr}$  images par classe ; le second comporte  $n_{ts}$  images. Ensuite, les valeurs optimales des paramètres critiques de chaque méthode seront choisies de manière à répondre au compromis entre le taux de reconnaissance et le temps d'exécution. Ce temps d'exécution mesure la durée entre l'entrée des images et la prise de décision de toute la base de données envisagée. Finalement et dans la deuxième phase expérimentale, les données sont analysées par des algorithmes de classification artificiels en mode de validation croisée. Enfin, les performances des classificateurs (KNN, SVM et MLP) sont évaluées suivant le taux d'exactitude, la précision, le rappel et la F-mesure.

### IV.2.1. Valeurs optimales des paramètres

Pour déterminer les valeurs optimales des paramètres critiques des méthodes LBPP, PCA, LDA, 2DPCA, 2DLDA et 2DDCT, nous appliquons les méthodes de réduction de la dimension sur les images obtenues par LBPP. Ensuite, nous comparons pour chaque expérience : les taux de reconnaissance atteints et les temps d'exécution réalisés.

#### IV.2.1.1. Dimension optimale de la base de projection

Afin de déterminer la dimension optimale de la base de projection de chacune des méthodes de réduction par projection, évoquées précédemment, nous réalisons des expériences où les

bases de données formées par les images obtenues par LBPP(k=4) sont utilisées. Pour chacune des méthodes (PCA, 2DPCA et 2DLDA), nous comparons les taux de reconnaissance et les temps d'exécution réalisés par les différentes valeurs de la dimension des bases de projection (dim). Alors que suite à l'objectif principal de la méthode LDA qui vise à réduire la dimension de base de projection tout en préservant les classes cette dimension est effectivement fixée à la valeur  $n_c-1$ .

#### a. Résultats de la base ORL

ORL est subdivisée en deux ensembles : un pour l'arrière-plan avec  $n_{tr}=5$  et l'autre pour les tests avec  $n_{ts}=5$ . Le Tableau III-2 résume les résultats obtenus avec PCA, 2DPCA et 2DLDA pour des valeurs de la dimension. Ces résultats montrent que 2DPCA et 2DLDA surpassent PCA et LDA (91.00% avec dim=39) en termes de taux de reconnaissance et temps d'exécution. En effet, pour la même dimension (dim=8 << 20), 2DPCA et 2DLDA atteignent le même taux de reconnaissance de 96.50%. Mais leurs temps d'exécution réalisés diffèrent, soit une durée de 7.59s pour 2DPCA et 9.48s pour 2DLDA.

Tableau III-2 : Dimension optimale de la base ORL

(dim)	Taux de reconnaissance (%)					Temps d'exécution (s)				
	4	8	12	16	20	4	8	12	16	20
LBPP+PCA	57.00	78.50	85.50	86.50	87.50	10.92	10.78	10.82	10.84	10.93
LBPP+2DPCA	95.00	96.50	95.00	95.00	94.00	6.34	7.59	9.17	10.44	11.95
LBPP+2DLDA	95.50	96.50	96.50	96.00	95.50	7.99	9.48	11.19	12.16	13.80

#### b. Résultats de la base Yale

Yale est subdivisée en deux ensembles avec  $n_{tr}=5$  et  $n_{ts}=6$ . Le Tableau III-3 récapitule les résultats obtenus avec PCA, 2DPCA et 2DLDA pour des valeurs de la dimension. Ces résultats montrent que 2DPCA et 2DLDA surpassent PCA et LDA (93.33% avec dim=14) en termes de taux de reconnaissance et temps d'exécution. En effet, pour la même dimension (dim=4 << 14), 2DPCA et 2DLDA atteignent respectivement les taux 95.56% et 96.67%. Leurs temps d'exécution réalisés sont respectivement 1.95s et 2.58s.

Tableau III-3 : Dimension optimale de la base Yale

(dim)	Taux de reconnaissance (%)					Temps d'exécution (s)				
	4	8	12	14	15	4	8	10	12	14
LBPP+PCA	66.67	85.56	91.11	91.11	91.11	2.83	2.84	2.84	2.87	2.89
LBPP+2DPCA	95.56	95.56	94.44	94.44	92.22	1.95	2.18	2.44	2.50	2.59
LBPP+2DLDA	96.67	94.44	94.44	93.33	93.33	2.58	2.88	3.18	3.42	3.41

### c. Résultats de la base AR

AR est subdivisée en deux ensembles avec  $n_{tr}=7$  et  $n_{ts}=7$ . Le Tableau III-4 récapitule les résultats obtenus avec PCA, 2DPCA et 2DLDA pour des valeurs de la dimension. Ces résultats montrent que 2DPCA et 2DLDA surpassent PCA et LDA (95.43% avec dim=99) en termes de taux de reconnaissance et temps d'exécution. En effet, pour la même dimension (dim=8<<40), 2DPCA et 2DLDA atteignent respectivement les taux 96.43% et 96.70% avec un même temps d'exécution de 7.82s.

Tableau III-4 : Dimension optimale de la base AR

(dim)	Taux de reconnaissance (%)					Temps d'exécution (s)				
	4	8	12	20	40	4	8	12	20	40
LBPP+PCA	52.65	81.43	89.29	93.57	95.14	5.17	5.74	6.38	8.17	12.59
LBPP+2DPCA	95.71	96.43	96.14	96.57	94.29	6.08	7.82	9.90	14.82	25.38
LBPP+2DLDA	96.40	96.70	96.70	96.80	96.08	5.80	7.82	9.40	10.11	14.92

Compte tenu de tous les résultats précédemment obtenus avec les trois bases de données ORL, Yale et AR, nous adoptons les deux méthodes 2DPCA et 2DLDA avec leurs dimensions optimales qui valent respectivement dim=8 pour ORL et AR, dim=4 pour Yale pour des expériences ultérieures de ce manuscrit. Outre la dimension de la base de projection la taille optimale de l'intervalle de confiance, représentant l'objectif des expériences suivantes, doit nécessairement être déterminée.

#### IV.2.1.2. Intervalle de confiance optimal

D'après les images de la Figure III.8 précédemment illustrée, le descripteur LBPP(k=4) a bien assuré la localisation des principaux composants du visage y compris le nez, les yeux et la bouche. Ceci favorise la caractérisation de l'intervalle de confiance optimal par la valeur

$k=4$ . Afin d'affirmer cette caractérisation, nous comparons les taux de reconnaissance enregistrés par les méthodes de réduction de la dimension pour chaque valeur de  $k$ ; avec,  $k \in \{1;2;3;4;\dots\}$ .

#### a. Résultats de la base ORL

ORL est subdivisée en deux ensembles avec  $n_{tr}=5$  et  $n_{ts}=5$ . Le Tableau III-5 résume les résultats obtenus avec PCA, LDA, 2DPCA, 2DLDA et 2DDCT pour des valeurs de  $k$ . Ces résultats montrent que l'intervalle de confiance optimal correspond bien à la valeur  $k=4$ . Cet intervalle donne de bons résultats avec presque toutes les méthodes de réduction de la dimension. En effet, les taux enregistrés sont respectivement 96.5% avec 2DPCA, 96.5% avec 2DLDA et 94.87% avec 2DDCT.

Tableau III-5 : Intervalle de confiance optimal de la base ORL

	PCA	LDA	2DPCA	2DLDA	2DDCT
LBPP $k=1$	88.00±1.29	88.50±1.22	95.00±1.11	95.50±1.09	94.72±1.12
LBPP $k=2$	88.50±1.29	91.00±1.22	96.00±1.11	96.00±1.09	94.81±1.12
LBPP $k=3$	87.50±1.29	91.00±1.22	96.50±1.09	96.50±1.08	94.85±1.12
LBPP $k>3$	87.50±1.29	91.00±1.22	96.50±1.09	96.50±1.08	94.87±1.10

#### b. Résultats de la base Yale

Yale est subdivisée en deux ensembles avec  $n_{tr}=5$  et  $n_{ts}=6$ . Le Tableau III-6 récapitule les résultats obtenus avec PCA, LDA, 2DPCA, 2DLDA et 2DDCT pour des valeurs de  $k$ . Ces résultats montrent que l'intervalle de confiance optimal correspond aussi à la valeur  $k=4$ . Ce même intervalle donne de bons résultats avec presque toutes les méthodes de réduction de la dimension. En effet, les taux enregistrés sont respectivement 95.56% avec 2DPCA, 96.67% avec 2DLDA et 99.23% avec 2DDCT.

Tableau III-6 : Intervalle de confiance optimal de la base Yale

	PCA	LDA	2DPCA	2DLDA	2DDCT
LBPP $k=1$	94.44±1.05	93.33±1.03	94.44±1.06	95.56±1.03	98.83±1.09
LBPP $k=2$	91.11±1.05	93.33±1.03	95.56±1.07	96.67±1.06	99.13±1.08
LBPP $k=3$	91.11±1.05	93.33±1.03	95.56±1.09	96.67±1.08	99.23±1.01
LBPP $k>3$	91.11±1.05	93.33±1.03	95.56±1.05	96.67±1.08	99.23±0.90

**c. Résultats de la base AR**

Quant à AR, elle est subdivisée en deux ensembles avec  $n_{tr}=7$  et  $n_{ts}=7$ . Ainsi, le Tableau III-7 récapitule les résultats obtenus avec 2DPCA, 2DLDA et 2DDCT pour des valeurs de  $k \in \{1; 2; 3; 4\}$ . Ces résultats montrent que l'intervalle de confiance optimal correspond toujours à  $k=4$ . En effet, les taux enregistrés sont respectivement 96.42% pour 2DPCA, 96.75% pour 2DLDA et 96.51% pour 2DDCT.

Tableau III-7 : Intervalle de confiance optimal de la base AR

	LBPP k=1	LBPP k=2	LBPP k=3	LBPP k>3
2DPCA	95.72±1.63	96.01±1.04	96.14±1.35	96.42±1.35
2DLDA	96.57±1.35	96.68±1.16	96.74±1.20	96.75±1.20
2DDCT	96.50±1.14	96.45±1.16	96.50±1.17	96.51±1.20

Compte tenu des résultats obtenus par les trois expériences précédentes, la valeur 4 du paramètre  $k$  correspond bien à la taille optimale de l'intervalle de confiance. Par conséquent, l'hypothèse de normalité des zones quasi-homogènes est largement confirmée. Par conséquent, les images obtenues par le descripteur LBPP( $k=4$ ) seront utilisées dans la suite de cette section. En effet, la détermination des paramètres de 2DDCT caractérisant la taille du vecteur descripteur de chaque image représente l'objectif principal des expériences immédiates.

**IV.2.1.3. Paramètres de la méthode 2DDCT**

Présentement, nous passons à la construction des vecteurs descripteurs assez discriminants et de faible taille à l'aide de la méthode 2DDCT. Pour y parvenir, nous décomposons chaque image LBPP( $k=4$ ) en une série de coefficients spectraux. Deux opérations sont envisagées. La première consiste à découper chaque image en  $n_b$  blocs non chevauchants et déterminer pour chaque bloc les  $p \times p$  coefficients spectraux. La deuxième, consiste à concaténer les  $b$  premiers coefficients de chaque bloc en utilisant la technique de zigzag. Les  $n_b$  vecteurs locaux ainsi engendrés sont concaténés dans un seul vecteur global de taille  $n_b \times b$  représentant image.

**a. Résultats de la base ORL**

ORL, est subdivisée en deux ensembles avec  $n_{tr}=5$  et  $n_{ts}=5$ . Les taux de reconnaissance illustrés dans le Tableau III-8 correspondent aux différents triplets  $(p, n_b, b)$ . Les paramètres de chaque triplet caractérisent respectivement la taille et le nombre des blocs ainsi que le nombre

*Chapitre III. Modèle Probabiliste Binaire Local (Local Binary Probabilistic Pattern, LBPP) :  
Application à la Biométrie Faciale*

des coefficients actifs au sein de chaque bloc. Les résultats obtenus montrent que le triplet ( $p=14, n_b=48, b=3$ ) atteint le taux de reconnaissance le plus élevé soit, 95.57%. Par conséquent, chaque image est représentée par un vecteur de  $48 \times 3 = 144$  coefficients au lieu de  $112 \times 92 = 10304$  pixels.

Tableau III-8 : Paramètres optimaux de 2DDCT pour ORL

p×p	p=8	10×10	12×12	13×13	14×14	15×15	16×16
b=3	94.08±1.19	94.19±1.21	94.23±1.12	95.30±1.18	95.57±0.95	94.88±1.05	95.22±1.12
b=6	93.37±1.51	93.27±1.33	93.73±1.10	94.07±1.22	94.85±1.08	93.98±0.99	94.62±0.89
b=10	91.58±1.57	92.56±1.40	92.92±1.40	93.03±1.39	93.80±1.28	93.61±1.07	94.29±1.13

**b. Résultats de la base Yale**

Yale, est subdivisée en deux ensembles avec  $n_{tr}=5$  et  $n_{ts}=6$ . Le Tableau III-9 illustre les taux de reconnaissance correspondent aux différents triplets ( $p, n_b, b$ ). Les résultats obtenus montrent que le triplet ( $p=13, n_b=63, b=6$ ) atteint le taux de reconnaissance le plus élevé soit, 99.70%. Par conséquent, chaque image est représentée par un vecteur de  $63 \times 6 = 378$  coefficients. Par contre, le triplet ( $p=14, n_b=48, b=3$ ) atteint un taux de 99.36%, et un vecteur de 144 coefficients. Ce dernier triplet réalise une réduction de la taille du vecteur descripteur de 61.90% avec une faible perte de taux de reconnaissance de 0.34%. Enfin, chaque image sera représentée par un vecteur de 144 coefficients au lieu de 378.

Tableau III-9 : Paramètres optimaux de 2DDCT pour Yale

p×p	8×8	10×10	12×12	13×13	14×14	15×15	16×16
b=3	99.53±0.79	99.63±0.61	99.45±0.82	99.50±0.63	99.36±0.74	99.47±0.69	99.52±0.79
b=6	99.26±1.01	99.33±0.96	99.67±0.63	99.70±0.70	99.28±0.65	99.67±0.90	99.58±0.62
b=10	98.77±1.22	98.96±1.20	99.25±0.94	99.51±0.85	99.32±0.89	99.48±0.76	99.47±0.76

**c. Résultats de la base AR**

Quant à AR, elle est subdivisée en deux ensembles avec  $n_{tr}=7$  et  $n_{ts}=7$ . Les taux de reconnaissance illustrés dans le Tableau III-10 correspondent aux différents triplets ( $p, n_b, b$ ). Les résultats obtenus montrent que le triplet ( $p=9, n_b=120, b=3$ ) atteint le taux de



reconnaissance le plus élevé soit, 96.84%. Par conséquent, chaque image est représentée par un vecteur de  $120 \times 3 = 360$  coefficients. Par contre, le triplet ( $p=14, n_b=48, b=3$ ) atteint un taux de 96.50 %, et un vecteur de 144 coefficients. Ce dernier triplet réalise une réduction de la taille du vecteur descripteur de 60.00% avec une faible perte de taux de reconnaissance de 0.35%. Enfin, chaque image sera représentée par un vecteur de 144 coefficients au lieu de 360.

Tableau III-10 : Paramètres optimaux de 2DDCT pour AR

p×p	8×8	9×9	12×12	13×13	14×14	15×15	16×16
b=3	96.44±1.07	96.84±0.94	96.60±1.02	96.62±1.12	96.50±1.04	96.44±1.10	96.34±0.95
b=6	95.94±1.12	96.40±0.97	96.50±0.94	96.81±1.25	96.60±1.22	96.74±1.02	96.67±0.99
b=10	95.24±1.24	95.70±1.08	96.15±1.07	96.74±0.98	96.34±1.19	96.42±1.30	96.63±1.05

En se basant sur les résultats des trois expériences précédentes, un compromis entre la taille du vecteur descripteur et le taux de reconnaissance détermine que les vecteurs descripteurs optimaux sont de 144 coefficients. Enfin, à partir du Tableau III-11 qui résume les tailles optimales des vecteurs descripteurs des méthodes de réduction de la dimension utilisées, nous concluons que 2DDCT est la plus qualifiés pour la phase d'extraction de caractéristiques de notre système. Les vecteurs résultants de cette méthode seront utilisés subséquentment dans la phase de classification pour évaluer les performances des algorithmes de classification par apprentissage artificiel.

Tableau III-11 : Taille optimale des vecteurs descripteurs

	PCA	LDA	2DPCA	2DLDA	2DDCT(14,3)
ORL	20×200	39×200	8×92=736	8×92=736	56×3=144
Yale	14×75	14×75	4×92=198	4×92=198	56×3=144
AR	20×700	99×700	8×92=736	8×92=736	56×3=144

#### IV.2.2. Evaluation des classificateurs

Dans cette section, les performances des classificateurs KNN, LIBSVM, SMO et MLP sont jugées par les valeurs des paramètres suivants : la précision, le rappel, la F-mesure et l'exactitude [119-121]. Ces valeurs sont déduites des matrices de confusion engendrées par des expériences réalisées en mode de validations croisées de *k-folders* ( $k_f = 4, \dots, 10$ ). L'analyse de

ces matrices de confusion montre que les instances males classées différent d'un classificateur à un autre. Raison pour laquelle, nous avons proposé leur fusion par les règles de vote suivantes: vote majoritaire ( $V_{maj}$ ), maximum de probabilités ( $V_{max}$ ), minimum de probabilités ( $V_{min}$ ), produit des probabilités ( $V_{prod}$ ) et moyenne des probabilités ( $V_{avg}$ ).

#### a. Résultats de la base ORL

Le Tableau III-12 résume les valeurs des paramètres d'évaluation des classificateurs (KNN, LIBSVM, SMO et MLP). Chaque classificateur est appliqué seul. En effet, les meilleurs résultats sont réalisés par les SVMs (LIBSVM et SMO). A titre d'exemple, la F-mesure atteint 98.58% pour LIBSVM et 98.23% pour SMO. En plus, le Tableau III-13 récapitule les valeurs des paramètres d'évaluation de la fusion des quatre classificateurs à l'aide des cinq règles de la stratégie du vote. Les résultats obtenus montrent que ce mode de fusion améliore la qualité de classification, plus précisément la règle ( $V_{avg}$ ) qui atteint une F-mesure de 98.74% et une exactitude de 98.70%.

Tableau III-12 : Comparaison de différents classificateurs pour ORL

Classificateurs	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
LIBSVM	98.67±0.47	98.64±0.51	98.58±0.51	98.61±0.52
SMO	98.34±0.65	98.27±0.75	98.23±0.72	98.25±0.74
MLP	97.59±0.49	97.44±0.55	97.37±0.55	97.39±0.56
KNN	97.00±0.49	96.77±0.63	96.66±0.61	96.72±0.59

Tableau III-13 : Comparaison des différents modes de vote sur ORL

Classificateurs	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
$V_{maj}$	98.46±0.60	98.54±0.53	98.48±0.58	98.44±0.61
$V_{max}$	98.60±0.51	98.67±0.46	98.63±0.52	98.58±0.51
$V_{min}$	98.57±0.55	98.65±0.48	98.58±0.56	98.55±0.54
$V_{prod}$	98.60±0.52	98.67±0.47	98.62±0.52	98.58±0.51
$V_{avg}$	98.71±0.66	98.78±0.61	98.74±0.67	98.70±0.68

### b. Résultats de la base Yale

Le Tableau III.14 résume les valeurs des paramètres d'évaluation de chaque classificateur appliqué seul. Effectivement, les meilleurs résultats sont réalisés par les SVMs (LIBSVM et SMO). A titre d'exemple, F-mesure atteint 99.4% pour LIBSVM et SMO atteint 99.40%. En plus, la fusion des quatre classificateurs (Tableau III-15) améliore la qualité de classification, plus précisément la règle de votes ( $V_{\text{maj}}$ ) qui atteint une F-mesure de 99.74% et une exactitude de 99.74%.

Tableau III-14 : Comparaison de différents classificateurs pour Yale

Classificateurs	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
LIBSVM	99.4±0.23	99.4±0.23	99.4±0.23	99.39±0.22
SMO	99.4±0.21	99.4±0.21	99.4±0.21	99.39±0.18
MLP	99.23±0.27	99.23±0.27	99.23±0.27	99.23±0.24
KNN	99.08±0.54	99.08±0.54	99.08±0.54	99.08±0.53

Tableau III-15 : Comparaison des différents modes de vote sur Yale

Classificateurs	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
$V_{\text{maj}}$	99.74±0.23	99.74±0.23	99.74±0.23	99.74±0.23
$V_{\text{max}}$	99.10±0.29	99.10±0.29	99.10±0.29	99.10±0.29
$V_{\text{min}}$	99.12±0.27	99.12±0.27	99.12±0.27	99.12±0.27
$V_{\text{prod}}$	99.23±0.21	99.23±0.21	99.23±0.21	99.23±0.21
$V_{\text{avg}}$	99.48±0.32	99.48±0.32	99.48±0.32	99.48±0.32

### c. Résultats de la base AR

Pour cette base, le Tableau III-16 récapitule les valeurs des paramètres d'évaluation de chaque classificateur appliqué seul. En effet, les meilleurs résultats sont obtenus par les SVMs et le MLP. En effet, la F-mesure atteint respectivement 99.4%, 99.40% et 98.3% pour LIBSVM, SMO et MLP. En plus, la fusion des quatre classificateurs (Tableau III-17) améliore la qualité de classification, plus précisément la règle de votes ( $V_{\text{maj}}$ ) qui atteint une F-mesure de 98.68% et une exactitude de 98.70%.

Tableau III-16 : Comparaison de différents classificateurs sur AR

Classificateurs	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
LIBSVM	98.58±0.12	98.50±0.13	98.48±0.14	98.58±0.12
SMO	98.35±0.25	98.23±0.29	98.23±0.29	98.35±0.25
MLP	98.38±0.13	98.3±0.13	98.3±0.13	98.28±0.13
KNN	97.68±0.32	97.51±0.34	97.51±0.34	97.68±0.32

Tableau III-17 : Comparaison de différents modes de vote sur AR

Classificateurs	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
V <sub>maj</sub>	98.80±0.18	98.68±0.18	98.68±0.19	98.70±0.12
V <sub>max</sub>	98.70±0.21	98.60±0.28	98.60±0.28	98.68±0.18
V <sub>min</sub>	98.68±0.19	98.60±0.26	98.60±0.26	98.60±0.23
V <sub>prod</sub>	98.68±0.19	98.60±0.25	98.60±0.25	98.60±0.24
V <sub>avg</sub>	98.74±0.19	98.64±0.19	98.64±0.19	98.66±0.21

### IV.3. Comparaison avec les méthodes de l'état de l'art

Finalement, nous comparons notre approche basée principalement sur notre descripteur LBPP avec quelques méthodes de l'état de l'art. En fait, les résultats illustrés au Tableau III-18 représentent les taux de reconnaissance relatifs aux diverses méthodes de reconnaissance faciale 2D. Ces taux montrent que nos approches sont très compétitives par rapport aux méthodes de l'état de l'art. En effet, pour ORL, ces approches ont atteint des taux plus élevés que ceux rapportés dans [88, 143-147], ils valent effectivement: 98.61% pour LIBSVM, 98.25% pour SMO et 97.39% pour MLP. De même, pour Yale, elles ont atteint des taux plus dominants que ceux rapportés dans [88, 143-147], ils valent effectivement: 99.39% pour LIBSVM, 99.39% pour SMO et 99.23% pour MLP. Enfin, pour AR, elles ont atteint des taux plus supérieurs que ceux rapportés dans [144-147]. ils valent effectivement: 98.58% pour LIBSVM, 98.35% pour SMO et 98.28% pour MLP.

Tableau III-18 : Comparaison avec des méthodes de l'état de l'art

Méthodes	HDCT	GWT+ DCT	PCA+ DCT	2DPCA	PCA+ SVM	LBP	ULBP/ LPQ	Notre approche		
								LBPP+2DDCT+		
Réfs	[88]	[143]	[143]	[144]	[145]	[146- 147]	[146- 147]	LIBSVM	SMO	MLP
ORL	88.43	96.50	95.75	96.00	97.00	91.00	91.50	<b>98.61</b>	<b>98.25</b>	<b>97.39</b>
Yale	68.00	96.00	95.50	84.24	99.39	96.00	98.00	<b>99.39</b>	<b>99.39</b>	<b>99.23</b>
AR	****	****	****	96.10	92.67	98.00	92.70	<b>98.58</b>	<b>98.35</b>	<b>98.28</b>

L'objectif principal de cette section est de déterminer la meilleure combinaison des méthodes au niveau de chaque étage du système de reconnaissance faciale. En effet, l'extraction des caractéristiques par LBPP(k=4), la réduction de la dimension par 2DDCT et la classification par LIBSVM ont significativement amélioré le taux de reconnaissance et le temps d'exécution de notre système. Ainsi les paramètres k=4, dim=8, p=14 et b=3 sont les valeurs optimales caractérisant respectivement l'intervalle de confiance et les différents paramètres de la réduction de la dimension.

## V. Conclusion

Dans ce chapitre, nous avons proposé la méthode d'extraction des caractéristiques locales nommée modèle probabiliste binaire local (LBPP). Dans un premier temps, nous avons mis en évidence ses fonctionnalités et ses avantages. Dans le deuxième, nous l'avons utilisé dans le processus de reconnaissance d'un système de reconnaissance faciale. Nous avons montré que la performance de la description faciale par LBPP augmente avec la taille de l'intervalle de confiance. De plus, nous avons intégré LBPP dans la phase d'extraction des caractéristiques de notre système de reconnaissance faciale algorithmes de classification. Deux phases d'expérimentation utilisant trois bases de données ORL, Yale et AR ont été réalisées. Dans la première phase, nous avons combiné LBPP avec les méthodes de réduction de la dimension (PCA, LDA, 2DPCA, 2DLDA et 2DDCT). Les résultats obtenus ont atteints des taux de reconnaissance élevés et des temps d'exécution réduits. Dans la deuxième phase, nous avons amélioré les performances de la classification en intégrant les algorithmes de la *machine-*

*Learning* notamment KNN, MLP, SMO, LIBSVM et leur fusion via la stratégie du vote. Les résultats de cette intégration ont été très remarquables. Il est important de noter que la comparaison de l'approche proposée avec les méthodes récentes de l'état de l'art a montré sa compétitivité et sa prépondérance.

Vue l'intérêt que présente le gradient dans la description de l'image ainsi que son pouvoir de détecter les zones de pics, nous proposons dans notre prochain chapitre notre descripteur LGPP qui bénéficie des avantages des deux concepts gradient et intervalle de confiance.

---

---

## **Chapitre IV. Modèle Probabiliste Gradient Local (Local Gradient Probabilistic Pattern, LGPP) : Application à la Biométrie Faciale**

---

---

En analyse vectorielle, la transformation d'une fonction scalaire en champ vectoriel et inversement utilise des opérateurs dont le gradient est le plus fondamental. Ce dernier caractérise les variations orientées d'une quantité scalaire en dépendance de ses propriétés intrinsèques et locales. Dans ce chapitre, nous visons améliorer la localisation des zones de fortes variations d'intensité en utilisant les deux concepts gradient et intervalle de confiance. Dans ce sens, nous avons proposé un descripteur local basé sur ces deux concepts et nommé modèle probabiliste gradient local (LGPP).

## **I. Introduction**

La principale contribution des méthodes d'extraction de caractéristiques au système de reconnaissance faciale est la génération d'une description intrinsèque discriminante à partir d'une image de visage. Pour y parvenir, nous proposons dans ce chapitre notre deuxième descripteur nommé modèle probabiliste gradient local (Local Gradient Probabilistic Pattern, LGPP). Ce descripteur étudie la normalité de la distribution qui régit les valeurs des gradients de chaque voisinage de l'image et définit l'intervalle de confiance qui régit ces gradients. Il est, en fait, une généralisation probabiliste du descripteur LGP [123]. En fait, nous commençons par la formulation et la mise en évidence des avantages du descripteur LGPP. Ensuite nous mettons en place un système de reconnaissance faciale qui utilise en plus de LGPP les méthodes de réduction de la dimension (2DDWT, 2DPCA et 2DLDA), ainsi que les algorithmes de classification (KNN, SVMs et MLPs). L'évaluation de ce système de reconnaissance faciale est réalisée à l'aide d'une multitude d'expériences utilisant les bases de données ORL, Yale et AR. Les résultats fournis par ces expériences ont montré que l'intégration du descripteur LGPP à la phase d'extraction des caractéristiques du système de reconnaissance a amélioré la performance globale de ce système de reconnaissance faciale.

## **II. Modèle Probabiliste Gradient Local (LGPP)**

### **II.1. Présentation de LGPP**

LGPP est extension probabiliste de la variété LGP proposée par Bongjin [123] et qui repose sur les deux concepts intervalle de confiance et gradient. Il consiste en deux étapes principales. Dans la première, nous définissons un nouveau voisinage formé par les valeurs des gradients, la deuxième étape utilise ce voisinage pour évaluer l'intervalle de confiance qui régit ces gradients. Certes, à l'instar de LBPP, le principal objectif de LGPP est de séparer les zones quasi-homogènes des zones de pics. Afin d'affirmer grossièrement que cet objectif est atteint, nous étudions les histogrammes de différentes zones d'une image de visage obtenue par LGPP. En effet, l'histogramme de la Figure IV.1-b montre que presque tous les pixels de la zone quasi-homogène choisie sont proches du niveau de gris 255 (670 valeurs du mode parmi 750). En plus, ils sont repartis autour de leur moyenne avec un écart-type ( $\sigma=9.43 < 19.27$ ) et un



coefficient de variation ( $\delta=0.037 < 0.077$ ). Par compte, les pixels de la zone de pic choisie, sur de la Figure IV.1-a, sont étalés sur toute la plage d'histogramme [0:255] avec un écart-type ( $\sigma=60.03 < 71.93$ ) et un coefficient de variation ( $\delta=0.304 < 0.4$ ). En comparaison avec les résultats obtenus par LBPP et qui sont illustrés dans la Figure III.3, nous concluons que LGPP a permis d'améliorer l'apparence globale du visage. En effet, le nombre des pixels quasi-homogènes a augmenté et la dispersion des pixels de l'image a diminué.

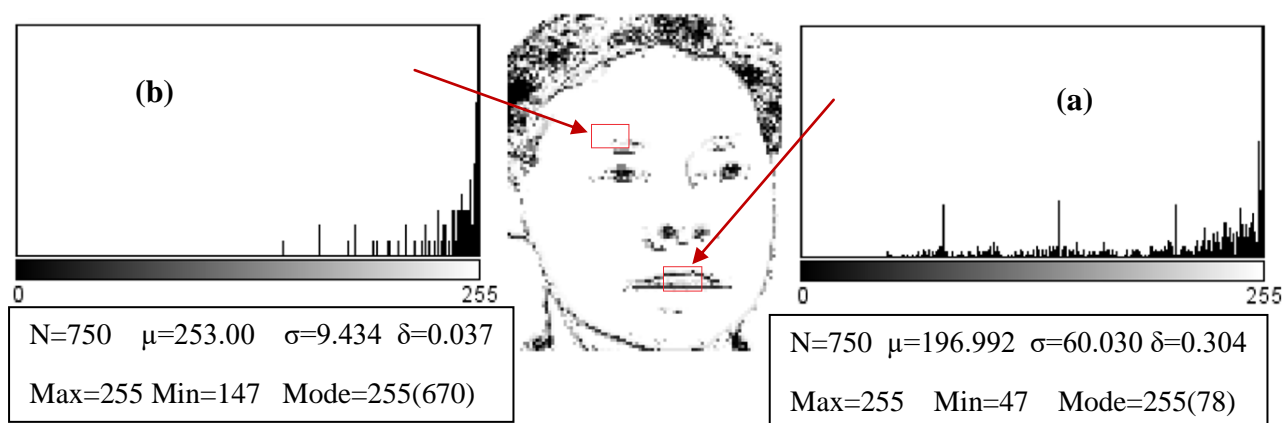


Figure IV.1 : Histogrammes des zones de l'image LGPP

Similairement à LBPP, le recodage des pixels par LGPP nécessite la définition de l'intervalle de confiance régissant chaque voisinage des gradients. La formulation de cet intervalle de confiance sera traitée dans la section immédiate.

## II.2. Évaluation de l'intervalle de confiance

Partant de l'intervalle de confiance,  $[\alpha_1, \alpha_2]$ , qui caractérise un voisinage de pixels,  $\mathbf{I}$ , nous pouvons déduire l'intervalle de confiance,  $[\gamma_1, \gamma_2]$ , qui caractérise le voisinage  $\mathbf{G}$  qui correspond à  $\mathbf{I}$  et qui est formé par les valeurs des gradients. Pour y parvenir, nous devons tout d'abord définir les éléments de  $\mathbf{G}$  tout en se basant sur ceux de  $\mathbf{I}$ . En effet, pour chacun des  $P-1$  pixels de l'intervalle  $I = \{i_1, \dots, i_{P-1}, i_c\}$ , nous définissons la valeur de gradient  $g_n$  de l'intervalle  $G = \{g_1, \dots, g_{P-1}, g_c\}$  tel que :

$$g_n = |i_n - i_c| \quad \forall n = 0, \dots, P - 1 \quad (\text{IV.1})$$

La valeur de gradient du pixel central est donnée par :

$$g_c = g_m = \frac{1}{p} \sum_{n=0}^{p-1} g_n \quad \forall n = 0, \dots, p-1 \quad (IV.2)$$

Soit  $[\alpha_1, \alpha_2]$ , l'intervalle de confiance régissant les valeurs en niveau de gris :

$$[\alpha_1, \alpha_2] = [\mu - K, \mu + K]; \text{ où } K = \begin{cases} k\sigma & \text{si } \{S_k = 0 \text{ ou } \delta < \beta\}; k = 1, \dots, 4 \\ \sigma & \text{si } \beta < \delta < 0.2 \\ 0.2\mu & \text{others} \end{cases} \quad (IV.3)$$

Nous définissons l'intervalle de confiance  $[\gamma_1, \gamma_2]$  régissant les valeurs des gradients en suivant les étapes :

$$\begin{aligned} \forall i_n \in I, \quad i_n \in [\alpha_1, \alpha_2] &\Rightarrow \mu - K \leq i_n \leq \mu + K \\ &\Rightarrow |i_n - \mu| \leq K \\ &\Rightarrow |(i_n - i_c) - (i_c - \mu)| \leq K \\ &\Rightarrow ||i_n - i_c| - |\mu - i_c|| \leq K \\ &\Rightarrow -K + |\mu - i_c| \leq |i_n - i_c| \leq K + |\mu - i_c| \end{aligned}$$

$$\text{Soit, } \max(0, -K + g_m) \leq g_n \leq K + g_m \text{ avec } g_m = |\mu - i_c|. \quad (IV.4)$$

$$D'où, \quad [\gamma_1, \gamma_2] = [\max(0, -K + g_m), K + g_m] \quad (IV.5)$$

$$\text{avec, } K = \begin{cases} k\sigma & \text{si } \{S_k = 0 \text{ ou } \delta < \beta\}; k = 1, \dots, 4 \\ \sigma & \text{si } \beta < \delta < 0.2 \\ 0.2\mu & \text{others} \end{cases} \quad (IV.6)$$

En se basant sur ce qui précède, nous associons à chaque distribution des pixels régi par l'intervalle de confiance  $[\alpha_1, \alpha_2]$ , une distribution des gradients régi par l'intervalle de confiance  $[\gamma_1, \gamma_2]$ . En effet, si le premier intervalle nécessite l'hypothèse d'approximation à une loi normale, le deuxième nécessite l'hypothèse d'approximation à une loi normale tronquée.

### II.3. Formulation de LGPP

Conformément aux équations (IV.5) et (IV.6), qui définissent l'intervalle de confiance du voisinage des gradients, le descripteur LGPP encode chaque pixel courant en utilisant le processus suivant (Figure IV.2) : chacun des huit pixels voisins ayant une valeur de gradient

---

située dans l'intervalle de confiance  $[\gamma_1, \gamma_2]$  se code par '1', les autres par '0'. Le code binaire ainsi obtenu est converti en niveau de gris par :

$$LGPP_{P,R,P',R'} = \sum_{n=0}^{P-1} S(g_n)2^n \quad (IV.7)$$

$$\text{où, } S(x) = \begin{cases} 1, & \text{si } \gamma_1(\mu, \sigma, k) \leq x \leq \gamma_2(\mu, \sigma, k) \\ 0, & \text{ailleurs } (k \in [-4,4]) \end{cases} \quad (IV.8)$$

R et P, sont respectivement le rayon et le nombre de pixels contribuant au calcul du code binaire. R' et P' sont respectivement le rayon et le nombre de pixels contribuant au calcul des moments statistiques ( $\mu, \sigma, \delta$  et  $S_k$ ).

Comme exemple d'un tel codage la Figure IV.2 illustre le résultat obtenu pour un voisinage dont  $R=1, P=8, R'=2$  et  $P'=25$ .

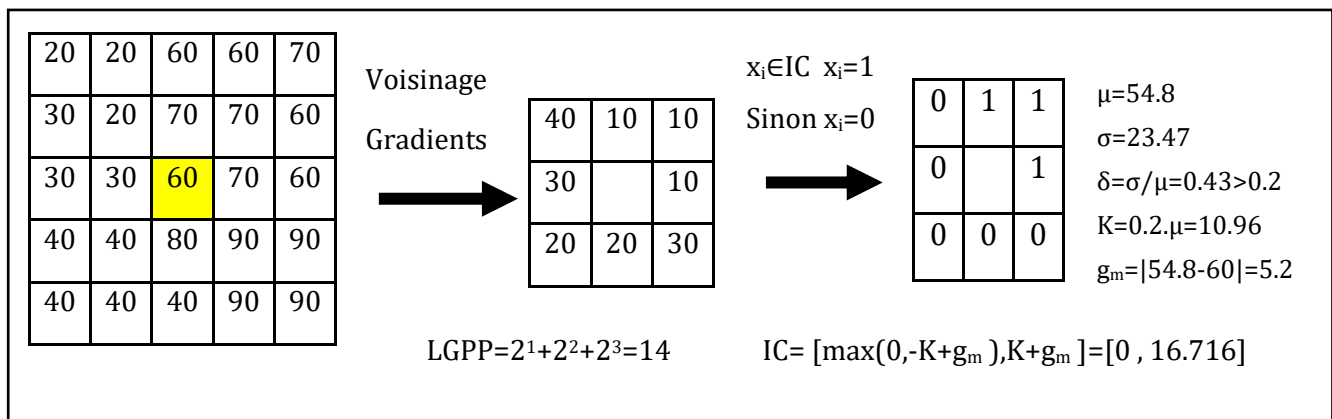


Figure IV.2 : Description d'un pixel par la méthode LGPP

## II.4. Avantages de LGPP

Tout comme le descripteur LBPP du chapitre précédent, le descripteur LGPP présente lui aussi plusieurs avantages. En effet, la caractéristique la plus importante de ce descripteur est sa capacité de générer une bonne description de la surface faciale. Dans cette description, les principaux composants du visage sont de mieux en mieux localisés tout en augmentant la taille de l'intervalle de confiance, Voir, Figure IV.3. Cet intervalle génère une représentation permanente et interchangeable pour toutes les valeurs de  $k \geq 4$ . Le deuxième avantage de LGPP

réside dans son comportement vis-à-vis des changements d'illumination, généralement liés à l'état d'éclairage de la scène et/ou à la qualité des capteurs.

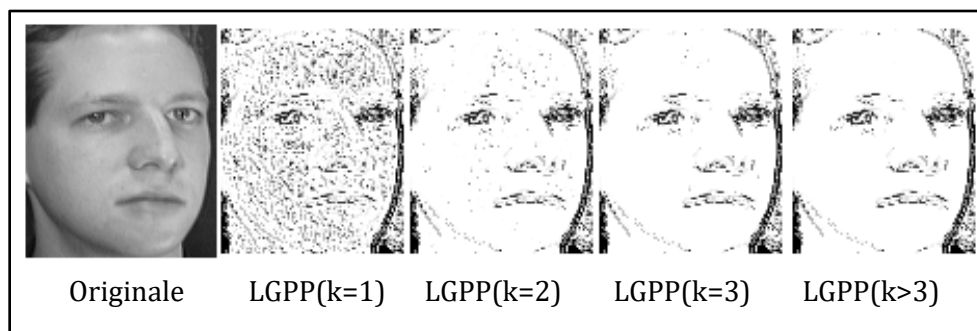


Figure IV.3 : Images obtenues par LGPP pour différents intervalles de confiance.

En outre, ces changements peuvent être simulés à l'addition ou la soustraction de l'intensité lumineuse, à l'accentuation ou l'affaiblissement du contraste, et à l'ajustement de la netteté par des corrections gamma. Ainsi, dans les différentes images de la Figure IV.4 les zones quasi-homogènes et les zones de pics sont toujours bien distinguées. En effet, cette figure illustre les images résultantes de l'application de LGPP(k=4) à une image originale et aux images obtenues par des changements d'éclairage, de contraste et de netteté.

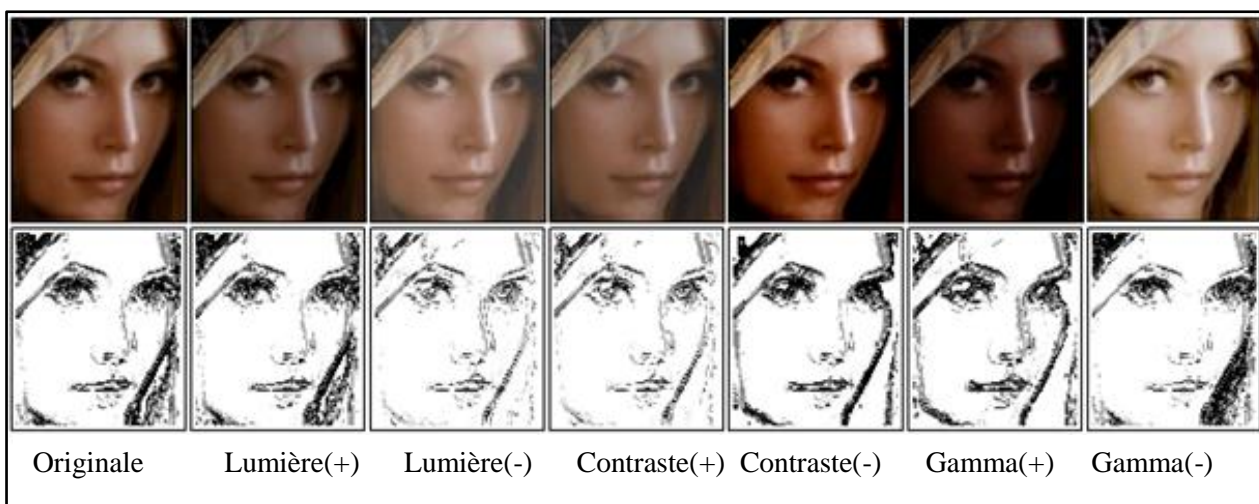


Figure IV.4 : Comportements de LGPP vis-à-vis différents d'éclairages.

De même, le descripteur LGPP est robuste aux changements locaux de lumière, aux changements de la structure de visage et aux bruits. Partant des images illustrées dans la Figure IV.5 et qui représentent les résultats de l'application de LGPP(k=4) à l'image originale

---

et aux images obtenues par des changements de couleurs sur des zones du front et des joues, des déformations structurelles par ondulation verticale et horizontale ou création de reliefs et des additions du bruit flou ou Gaussien. Nous remarquons que dans toutes ces images les éléments principaux du visage restent distingués.

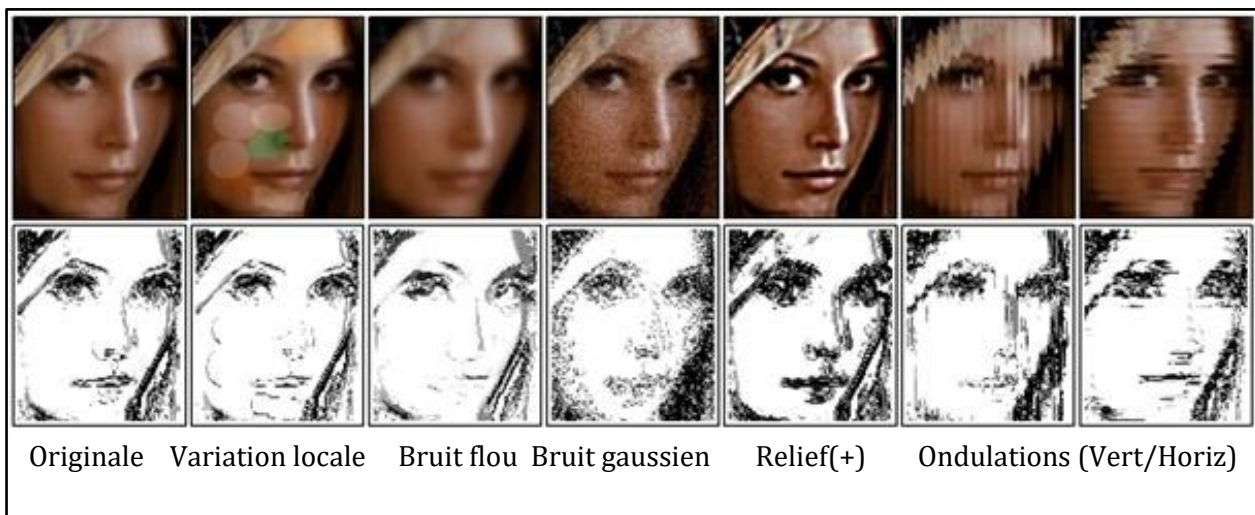


Figure IV.5 : Comportements de LGPP vis-à-vis différents changements et bruits.

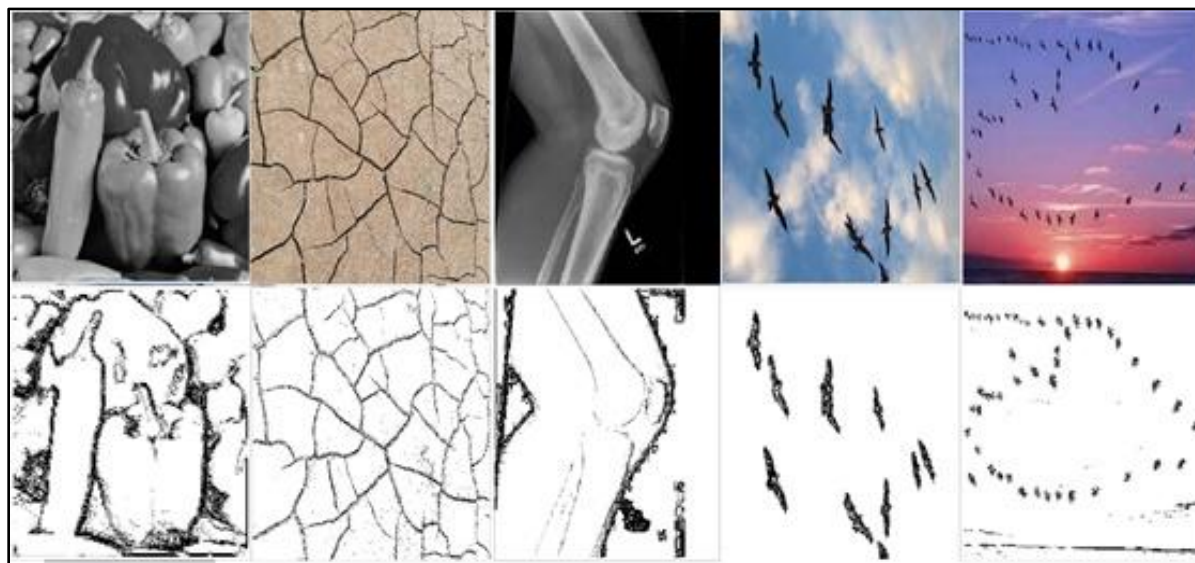


Figure IV-6 : Domaines d'application de LGPP.

De plus, LGPP(k=4) donne des résultats prometteurs avec des images de différents domaines de vision par ordinateur et de traitement d'image. Pour affirmer cette déclaration nous avons appliqué LGPP(k=4) à des images de différents contextes. Les résultats sont

---

effectivement illustrés dans la Figure IV.6. D'une part, LGPP(k=4) permet de bien localiser les zones de fortes variations du gradient et peut être utilisé dans des applications où l'analyse des textures et l'extraction des contours est un traitement indispensable (analyse des textes, détection de fissures et etc.). D'autre part, LGPP(k=4) permet de bien éliminer les composants de l'arrière-plan et peut être utilisé dans des applications où la soustraction de l'arrière-plan est un traitement essentiel (analyse de mouvement, suivi des objets et etc.).

### III. Reconnaissance de visage par LGPP

L'objectif de cette section est de mettre en évidence l'effet de la description faciale par LGPP à l'implémentation du système de reconnaissance faciale. Dans ce système de reconnaissance, nous combinons le descripteur LGPP, les méthodes de réduction de dimension et les algorithmes de classification afin d'atteindre des meilleures performances en termes de taux de reconnaissance et de temps d'exécution.

#### III.1. Système de reconnaissance faciale proposé

Le système de reconnaissance proposé dans la Figure IV.7 comporte trois phases principales. Dans la première phase, chaque visage détecté et normalisé sera traité par le descripteur LGPP. L'image résultante est décomposée, au deuxième niveau, par 2DDWT en sept subbandes, une approximative et six détails.

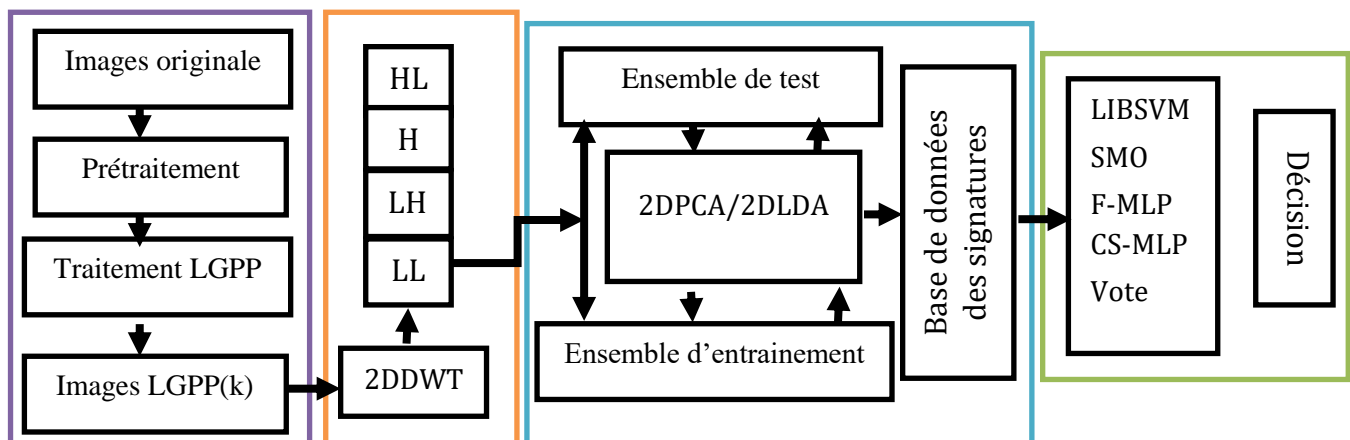


Figure IV-7 : Diagramme blocs du système de reconnaissance proposé

L'ensemble des approximations, de taille réduite, sont traitées par les méthodes 2DPCA et/ou 2DLDA pour générer des vecteurs descripteurs de très faibles dimensions. Enfin et dans la troisième phase, l'ensemble formé par ces vecteurs descripteurs sera classifié à l'aide des algorithmes SVMs et/ou MLPs. Le fonctionnement des trois phases du système proposé sera détaillé dans les sous-sections suivantes. Ces vecteurs descripteurs sera classifié à l'aide des algorithmes SVMs et/ou MLPs. Le fonctionnement des trois phases du système proposé sera détaillé dans les sous-sections suivantes.

### **III.1.1. Prétraitement**

Dans cette phase deux opérations sont successivement réalisées : le recadrage géométrique et l'égalisation d'histogramme. La première opération exclue les composantes non pertinentes de l'arrière-plan et la deuxième assure l'équiprobabilité de tous les niveaux de gris.

### **III.1.2. Traitement par LGPP**

Dans cette phase, les images normalisées sont traitées par le descripteur LGPP qui localise les principaux composants qui marquent la topologie de la surface faciale. Cette propriété est mise en évidence par application de LGPP( $k=4$ ) à des images de trois individus, sept images pour chacun, extraites des trois bases de données ORL, Yale et AR. Dans toutes les images de la Figure IV.8, les traits marquants du visage sont effectivement bien localisés et affirment l'hypothèse des zones quasi-homogènes de LGPP. En effet, les contours des lunettes des images ORL sont accentués, les arrière plans des images Yale sont éliminés et les effets d'éclairage des images AR sont annulés.

### **III.1.3. Réduction de dimension et vecteurs caractéristiques**

Dans cette phase les trois méthodes de réduction de la dimension : 2DDWT, 2DPCA et 2DLDA sont utilisées pour engendrer un ensemble de vecteurs descripteurs de dimension réduite. Initialement, 2DDWT décompose l'image LGPP en sept sub-bandes (Figure IV.9): l'approximative (LL2) et les six détails (LH1, HL1, HH1, LH2, HL2 et HH2). En effet, dans le premier niveau de décomposition l'image LGPP est décomposée en l'approximative (LL1) et les trois détails (LH1, HL1 et HH1). L'approximative (LL1) est elle aussi décomposée au deuxième niveau en une approximative (LL2) et trois détails (LH2, HL2 et HH2). Enfin, l'ensemble des approximations (LL2) sont sujet de réduction de la dimension par les deux

---

méthodes 2DPCA et/ou 2DLDA. Le choix de se limiter à ces deux méthodes est justifié par la forte ressemblance entre les deux descriptions LGPP et LBPP du visage. Ainsi que par la supériorité de ces deux méthodes vis-à-vis leurs antécédentes PCA et LDA. Cette supériorité a été largement affirmée dans le chapitre précédent.

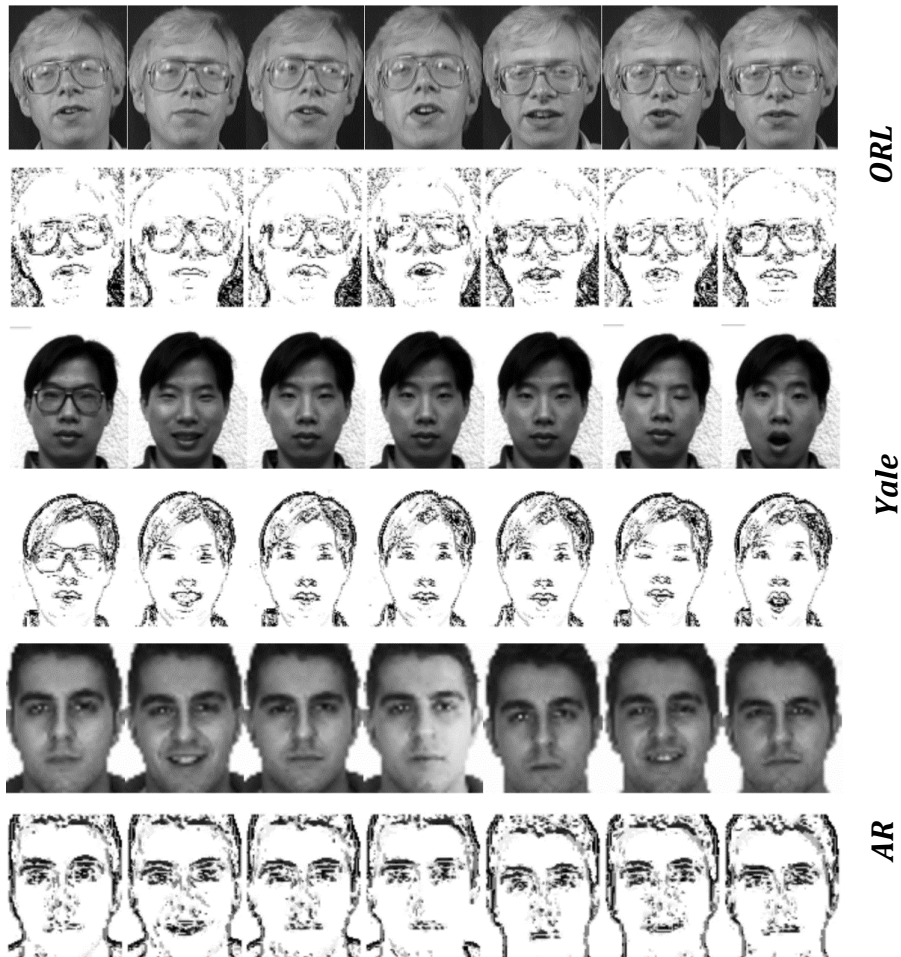


Figure IV.8 : Images LGPP(k= 4) pour ORL, Yale et AR.

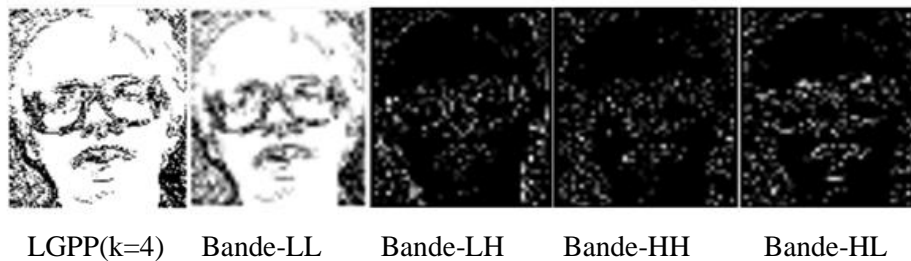


Figure IV.9 : Sub-bandes de décomposition au premier niveau de 2DDWT.



### III.1.4. Classification

Dans cette phase, Les vecteurs descripteurs, de taille réduite, précédemment engendrés pour chaque base de données représentent l'ensemble des données d'entrée pour les algorithmes de classification suivants : LIBSVM, SMO, F-MLP et CS-MLP. Les performances de ces algorithmes sont évaluées à l'aide des paramètres de classification suivants : le taux d'exactitude, la précision, le rappel, la F-mesure, le ROC-area et la courbe ROC.

## III.2. Les résultats expérimentaux

Dans cette sous-section, nous évaluons notre système de reconnaissance de visages à travers des expériences réalisées sur les trois bases de données ORL, Yale et AR. Dans chaque phase de notre système nous testons plusieurs méthodes afin de déterminer la meilleure d'entre elles. Le long du processus global de reconnaissance, nous déterminons la combinaison qui donne la meilleure performance en jouant sur les valeurs des paramètres critiques de chaque méthode. En effet, l'intervalle de confiance, la dimension de la base de projection, et le niveau de décomposition spectrale sont des variables caractérisant respectivement le descripteur LGPP et les méthodes de réduction de la dimension. Deux phases expérimentales sont mises en place. La première consiste à réaliser des expériences pour évaluer les paramètres de performance où chaque expérience est répétée 100 fois. Dans chaque répétition, les images de chaque base de données appartenant à  $n_c$  classes sont aléatoirement subdivisées en un ensemble d'entraînement et un ensemble de test. Le premier comporte  $n_{tr}$  images par classe ; le second comporte  $n_{ts}$  images. Ensuite, les valeurs optimales des paramètres critiques de chaque méthode seront choisies de manière à répondre au compromis entre le taux de reconnaissance et le temps d'exécution. Finalement et dans la deuxième phase expérimentale, les données sont analysées par des algorithmes de classification artificiels en mode de validation croisée. Enfin, les performances des classificateurs (LIBSVM, SMO, F-MLP et CS-MLP) sont évaluées suivant le taux d'exactitude, la précision, le rappel la F-mesure, le ROC-area et la courbe ROC.

### III.2.1. Valeurs optimales des paramètres

Afin de profiter au maximum des aptitudes des méthodes utilisées, nous déterminons les valeurs optimales de leurs paramètres critiques. Nous comparons donc les taux de reconnaissance et les temps d'exécution réalisés pour chaque expérience.

### III.2.1.1. Dimension optimale de la base projection

Dans le chapitre précédent nous avons montré que les méthodes de réduction de la dimension 2DPCA et 2DLDA sont plus efficace que leurs antécédentes PCA et LDA. Dans ce chapitre, nous nous limitons à tester cette propriété sur la base de données ORL formée de deux ensembles d'images LGPP(k=4): le premier d'arrière-plan ( $n_{tr}=5$ ), le second de test ( $n_{ts}=5$ ). Les résultats illustrés dans le Tableau IV-1 montrent que 2DPCA et 2DLDA surpassent largement PCA et LDA en termes de taux de reconnaissance et temps d'exécution. En effet, pour la même dimension ( $dim=8 \ll 20$ ), 2DPCA et 2DLDA atteignent un taux de reconnaissance de 96.00%. Leurs temps d'exécution sont respectivement 7.09s pour 2DPCA et 9.22s pour 2DLDA. Par conséquent, seules 2DPCA et 2DLDA seront utilisées ultérieurement et avec les mêmes dimensions du chapitre précédent { $dim(ORL)=8$ ;  $dim(Yale)=4$ ;  $dim(AR)=8$  }.Après avoir fixé les dimensions optimales des deux méthodes 2DPCA et 2DLDA, nous passons à la mise en évidence des possibilités offertes par les ondelettes, spécialement celle de Haar.

Tableau IV-1 : Dimension optimale pour ORL

(dim)	Taux de reconnaissance (%)					Temps d'exécution (s)				
	6	8	12	16	20	6	8	12	16	20
LGPP+PCA	71.00	81.50	88.00	89.00	89.50	10.17	10.45	10.39	10.61	11.73
LGPP+2DPCA	94.00	96.00	95.00	94.50	94.00	6.31	7.09	8.44	9.86	12.73
LGPP+2DLDA	94.50	96.00	95.00	95.00	94.00	8.11	9.22	10.59	11.45	13.34

### III.2.1.2. Effet de la décomposition par 2DDWT

Dans cette sous-section, la méthode 2DDWT décompose chaque image LGPP(k=4) en sept sub-bandes, notamment l'approximative (LL2) et les six détails (LH1, HL1, HH1, LH2, HL2 et HH2). En effet, c'est l'approximative qui emmagasine le maximum d'informations issues de LGPP(k=4). Par conséquent, chaque image LGPP(k=4) de taille 112×92 sera représentée par une image réduite 28×23,soit une réduction de 75% .Cette réduction se traduira au niveau du système de reconnaissance par une réduction du temps d'exécution et une amélioration du taux de reconnaissance. Le Tableau IV-3résume les résultats obtenus par des tests réalisés sur les bases de données ORL, Yale et AR avec et sans utilisation de la méthode 2DDWT.

- Subdivisée en deux ensembles, d'entraînement avec  $n_{tr}=5$  et du test avec  $n_{ts}=5$ , ORL enregistre de bons résultats. En effet, 2DPCA atteint un taux de 97.00% et réduit de 34% le temps d'exécution. De même, 2DLDA atteint un taux de 96.50% et réduit de 40% le temps d'exécution.
- Yale est elle aussi subdivisée en deux ensembles  $n_{tr}=5$  et  $n_{ts}=6$ . Les résultats du Tableau IV-2 montrent que 2DPCA atteint un taux de 98.89% en réduisant le temps d'exécution de 29%. De même, 2DLDA atteint un taux de 98.89% en réduisant le temps d'exécution de 34%.
- Quant à AR, elle est subdivisée en deux ensembles,  $n_{tr}=n_{ts}=7$ . Les résultats du Tableau IV-2 montrent que 2DPCA atteint un taux de 96.71% en réduisant le temps d'exécution de 32%. De même, 2DLDA atteint un taux de 96.43% en réduisant le temps d'exécution de 32%.

Tableau IV-2 : Effet de 2DDWT sur l'efficacité du système de reconnaissance

BD	ORL ( $n_{tr}=5$ et $n_{ts}=5$ )		Yale ( $n_{tr}=5$ et $n_{ts}=6$ )		AR ( $n_{tr}=7$ et $n_{ts}=7$ )	
	Taux(%)	Temps(s)	Taux(%)	Temps(s)	Taux(%)	Temps(s)
2DPCA	96.00	1.60	96.00	0.54	95.86	10.1
2DLDA	95.50	2.02	97.78	0.71	96.57	11.9
2DDWT+2DPCA	97.00	1.05	98.89	0.38	96.71	6.84
2DDWT+2DLDA	96.50	1.21	98.89	0.47	96.43	8.08

Compte tenu des résultats précédents, Il est évident que la méthode 2DDWT contribue à l'amélioration de la performance globale de notre système de reconnaissance en termes de taux de reconnaissance et temps d'exécution. Dans la prochaine sous-section, nous menons des tests qui confirment l'hypothèse des zones quasi-homogènes en étudiant l'effet du paramètre  $k$  sur la performance du système proposé.

### III.2.1.3. Intervalle de confiance optimal

Dans la conception de notre descripteur LGPP, nous avons montré que l'apparence du visage dépend principalement de la taille de l'intervalle de confiance en chaque point de l'image, et que les principales composantes du visage deviennent de plus en plus mieux localisées en allant de  $k=1$  vers  $k=4$ . Afin d'affirmer que la valeur  $k=4$  définit bien l'intervalle

de confiance optimal, nous comparons les taux de reconnaissance enregistrés pour chaque valeur de  $k \in \{1, 2, 3, 4\}$ . Les résultats obtenus sont illustrés par le Tableau IV-3.

- ORL est subdivisée en deux ensembles avec  $n_{tr} = n_{ts} = 5$ . Les résultats du Tableau IV-3 montrent que l'intervalle de confiance optimal correspond bien à  $k=4$ . En effet, les taux enregistrés avec cet intervalle sont respectivement 97.00% pour 2DPCA et 96.50% pour 2DLDA.
- Yale est subdivisée en deux ensembles avec  $n_{tr} = 5$  et  $n_{ts} = 6$ . Les résultats du Tableau IV-3 montrent que  $k=4$  donne de bons résultats avec 2DPCA et 2DLDA. En effet, 2DPCA et 2DLDA réalisent toutes les deux un taux de 98.89%.
- Quant à AR, elle est subdivisée en deux ensembles avec  $n_{tr} = n_{ts} = 7$ . Les résultats du Tableau IV-3 montrent que, même pour cette base de données,  $k=4$  donnait de bons résultats. En effet, les taux enregistrés sont respectivement 96.71% avec 2DPCA et 96.43% avec 2DLDA.

Tableau IV-3 : Intervalle de confiance optimal

Bases	ORL ( $n_{tr}=5$ et $n_{ts}=5$ )		Yale ( $n_{tr}=5$ et $n_{ts}=6$ )		AR ( $n_{tr}=7$ et $n_{ts}=7$ )	
Méthodes	2DDWT+		2DDWT+		2DDWT+	
	2DPCA	2DLDA	2DPCA	2DLDA	2DPCA	2DLDA
LGPP $k=1$	96.00%	95.50%	97.78%	97.78%	96.14%	96.29%
LGPP $k=2$	96.50%	96.00%	98.67%	98.67%	96.57%	96.29%
LGPP $k=3$	97.00%	96.50%	98.89%	98.89%	96.71%	96.43%
LGPP $k>3$	97.00%	96.50%	98.89%	98.89%	96.71%	96.43%

Compte tenu de tous les résultats expérimentaux précédents, la valeur  $k=4$  du paramètre caractérisant la taille de l'intervalle de confiance de LGPP correspond bien à la valeur optimale. De plus, cette valeur représente une limite au-delà de laquelle l'apparence globale du visage est immuable. Par conséquent, les images obtenues par LGPP ( $k = 4$ ) seront utilisées dans la suite de ce chapitre. À l'instant, nous testons l'impact de plusieurs classificateurs dans le but de déterminer le plus approprié pour notre système de reconnaissance.

### III.2.2. Evaluation par classification artificielle

Dans cette section, nous utilisons les valeurs optimales des paramètres des méthodes précédemment étudiées, notamment LGPP, 2DPCA, 2DLDA et 2DDWT. Nous représentons chaque image LGPP(k=4) de taille 112×92 par un vecteur descripteur de 8×28=224 instances pour les bases ORL et AR. Pour Yale ce vecteur est de 4×28=112 instances. L'ensemble de ces vecteurs descripteurs est utilisé en mode de validations croisées de *k-folders* (kf =4, 6, 8, 10) pour évaluer les performances des classificateurs (LIBSVM, SMO, F-MLP et CS-MLP). D'une part, les Tableau IV-4 et Tableau IV-5 représentent les résultats obtenus par les deux algorithmes de type SVMs, à savoir, LIBSVM et SMO. D'une autre part, les Tableau IV-6 et Tableau IV-7 représentent les résultats obtenus par les deux algorithmes de type MLPs, à savoir, F-MLP et CS-MLP. En effet, à partir des résultats illustrés dans le Tableau IV-4, nous pouvons conclure que LIBSVM atteint une performance remarquable avec le noyau polynomial  $K(x_i, x_j) = (x_i x_j + 1)^3$ . Effectivement, LIBSVM+2DPCA atteint respectivement une F-mesure de 98.41% pour ORL, 99.67% pour Yale et 98.09% pour AR, et une exactitude de 98.40% pour ORL, 99.69% pour Yale et 98.09% pour AR. De même LIBSVM+2DLDA atteint respectivement une F-mesure de 98.25% pour ORL, 99.67% pour Yale, et 98.46% pour AR ; et une exactitude de 98.25% pour ORL, 99.69% pour Yale et 98.45% pour AR.

Tableau IV-4 : Résultats du classificateur LIBSVM

Méthodes		BD	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)	ROC-area (%)
LGPP+2DDWT	2DPCA	ORL	98.62±0.32	98.50±0.37	98.41±0.37	98.40±0.35	99.66±0.11
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.23±0.23	98.10±0.28	98.10±0.28	98.09±0.26	99.00±0.14
	2DLDA	ORL	98.46±0.32	98.35±0.37	98.25±0.37	98.25±0.35	99.66±0.11
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.56±0.14	98.46±0.14	98.46±0.14	98.45±0.16	99.20±0.11

En plus, les résultats illustrés au Tableau IV-5, montrent que la performance de l'algorithme SMO avec le noyau polynomial  $K(x_i, x_j) = (x_i x_j + 1)^2$  est remarquable. En effet, SMO+2DPCA atteint respectivement une F-mesure de 97.96% pour ORL, 99.67% pour Yale

et 98.10% pour AR, et une exactitude de 98.00% pour ORL, 99.69% pour Yale et 97.81% pour AR. De même, SMO+2DLDA atteint respectivement une F-mesure de 98.14% pour ORL, 99.67% pour Yale, et 98.06% pour AR, et une exactitude de 98.06% pour ORL, 99.69% pour Yale et 98.06% pour AR.

Tableau IV-5 : Résultats du classificateur SMO

Méthodes		BD	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)	ROC-area (%)
LGPP+2DDWT	2DPCA	ORL	98.16±0.32	98.05±0.37	97.96±0.37	98.00±0.35	99.46±0.11
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.30±0.14	98.10±0.14	98.10±0.14	97.81±0.30	99.06±0.1
	2DLDA	ORL	98.20±0.32	98.16±0.37	98.14±0.37	98.06±0.35	99.40±0.11
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.26±0.18	98.06±0.18	98.06±0.18	98.06±0.20	99.03±0.1

En ce qui concerne l'algorithme F-MLP, les résultats illustrés dans le Tableau IV-6 démontrent la performance remarquable de celui-ci. Effectivement, F-MLP+2DPCA atteint respectivement une F-mesure de 97.77% pour ORL, 99.67% pour Yale et 98.56% pour AR, et une exactitude de 97.75% pour ORL, 99.69% pour Yale et 98.52% pour AR. De même, F-MLP+2DLDA atteint respectivement une F-mesure de 98.20% pour ORL, 99.67% pour Yale, et 98.10% pour AR, et une exactitude de 98.02% pour ORL, 99.69% pour Yale et 98.14% pour AR.

Tableau IV-6 : Résultats des classificateurs F-MLP

Méthodes		BD	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)	ROC-area (%)
LGPP+2DDWT	2DPCA	ORL	97.95±0.34	97.77±0.37	97.77±0.37	97.75±0.35	99.70±0.09
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.63±0.37	98.53±0.37	98.53±0.37	98.52±0.36	99.67±0.28
	2DLDA	ORL	98.26±0.36	98.20±0.40	98.20±0.40	98.02±0.39	99.70±0.09
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.30±0.28	98.16±0.32	98.16±0.32	98.14±0.30	99.83±0.23

En plus, les résultats illustrés au Tableau IV-7, montrent que la performance de l'algorithme CS-MLP est remarquable. En effet, CS-MLP+2DPCA atteint respectivement une F-mesure de 97.85% pour ORL, 99.65% pour Yale et 97.65% pour AR, et une exactitude de 97.75% pour ORL, 99.69% pour Yale et 98.52% pour AR. De même, CS-MLP+2DLDA atteint respectivement une F-mesure de 98.18% pour ORL, 99.74% pour Yale, et 98.35% pour AR, et une exactitude de 98.02% pour ORL, 99.69% pour Yale et 98.14% pour AR.

Tableau IV-7 : Résultats des classificateurs CS-MLP

Méthodes		BD	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)	ROC-area (%)
LGPP+2DDWT	2DPCA	ORL	97.95±0.34	97.77±0.37	97.77±0.37	97.75±0.35	99.70±0.09
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.63±0.37	98.56±0.32	98.56±0.32	98.54±0.33	99.90±0.24
	2DLDA	ORL	98.26±0.36	98.20±0.40	98.20±0.40	98.02±0.39	99.70±0.09
		Yale	99.67±0.28	99.67±0.28	99.67±0.28	99.69±0.29	99.63±0.34
		AR	98.30±0.29	98.16±0.33	98.10±0.28	98.14±0.30	99.83±0.23

Outre les différents paramètres d'évaluation de classificateurs déduits de la matrice de confusion, notamment la F-mesure et l'exactitude, la courbe ROC [11,119] représente un outil graphique d'évaluation qui trace le taux des vrais positifs en fonction du taux des faux positifs. En effet, plus cette courbe tend vers le point (0,1), plus la surface sous la courbe ROC ou ROC-area tend vers 1 et plus le classificateur est évalué efficace. Dans cette étude, nous avons tracé les courbes ROCs relatives aux différents triplets (base de données, méthode de réduction de la dimension, classificateur). Les graphiques des 24 triplets sont illustrés dans les Figure IV.10 et Figure IV.11. En effet, la courbe ROC de chacun des 24 triplets tend à se marier avec l'axe des ordonnés, sa surface sous ROC est élevée (ROC-area>0.99) et sa qualité de classification est pertinente.

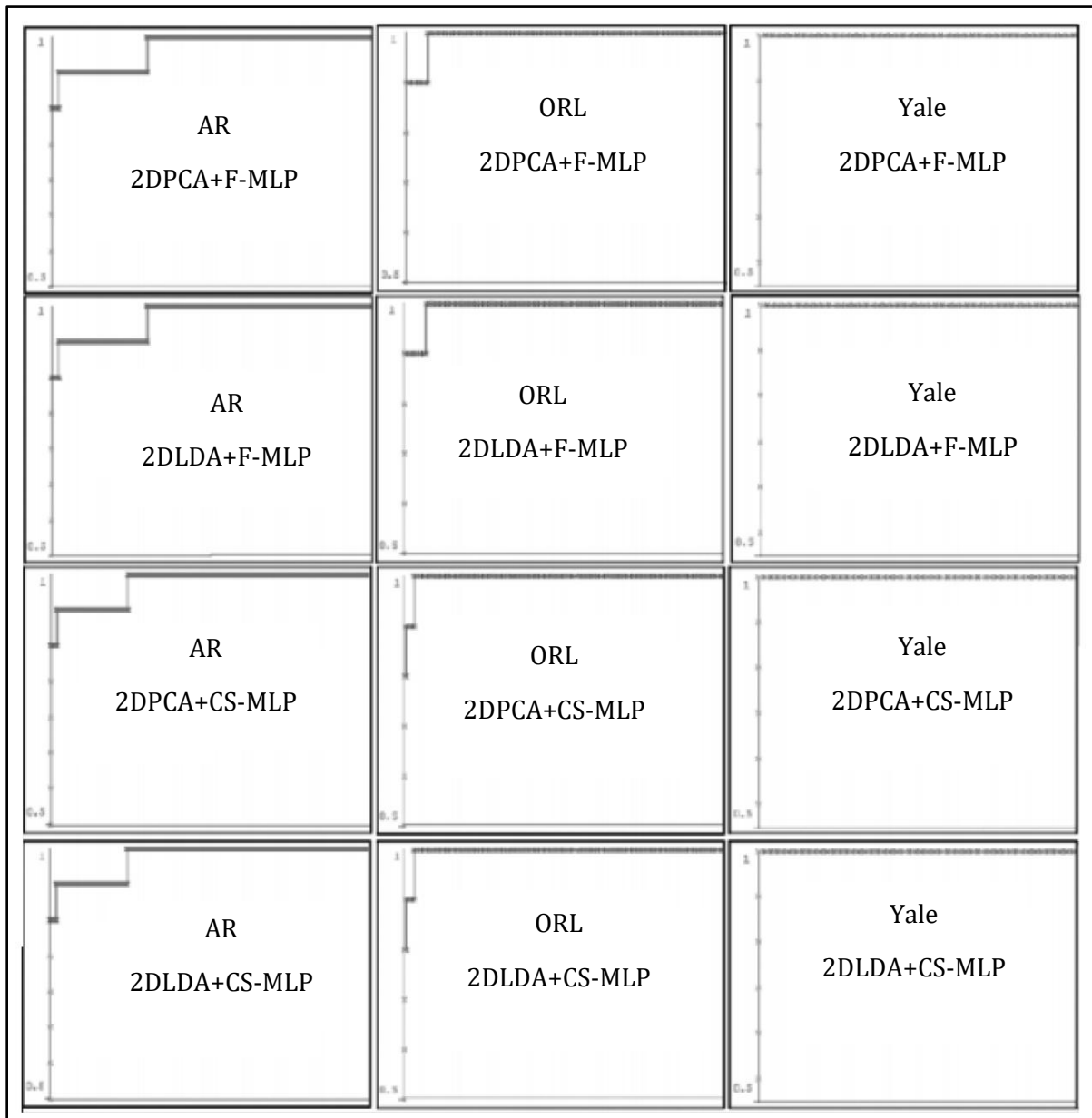


Figure IV.10 : Les courbes ROC relatives aux MLPs



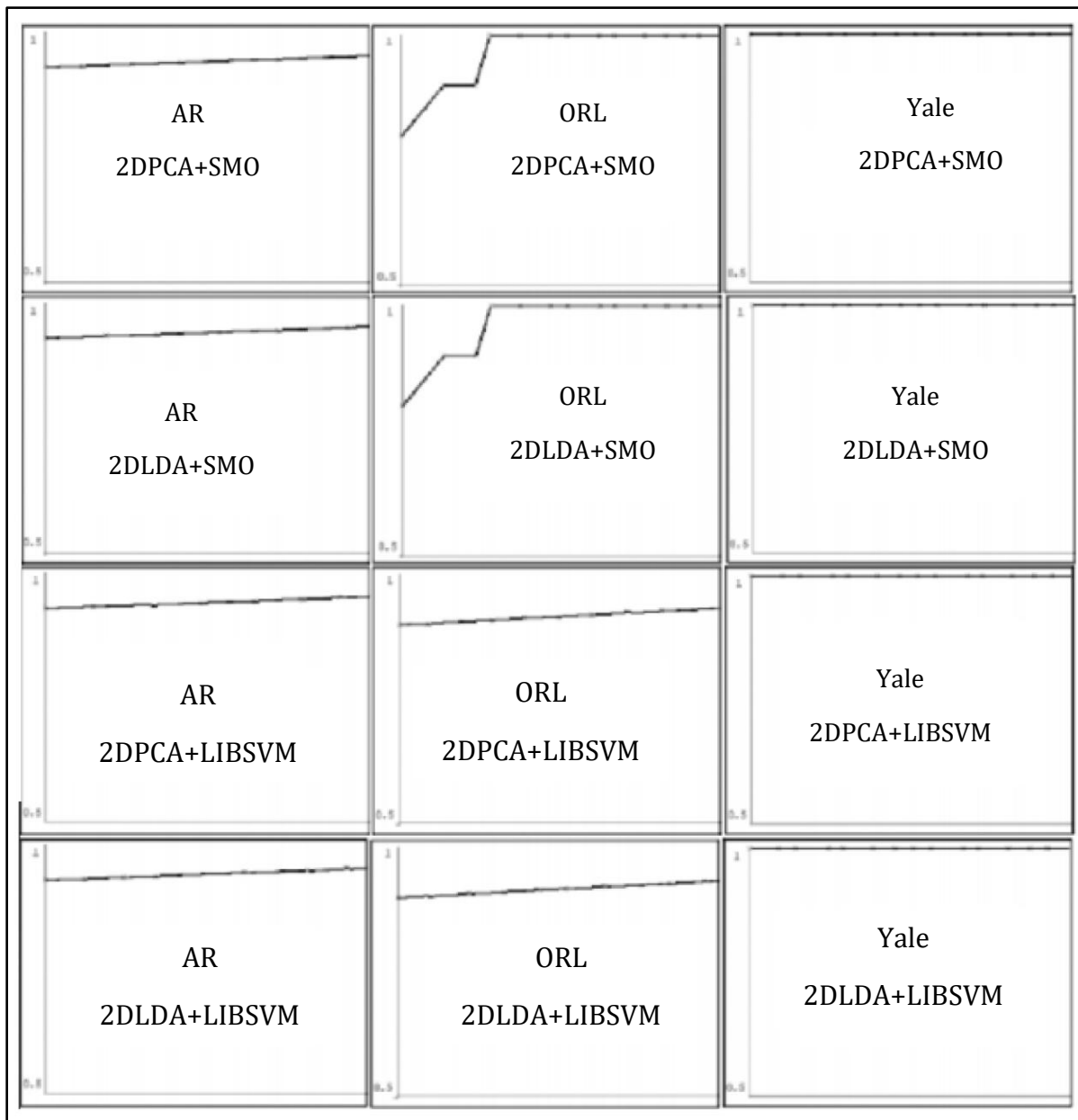


Figure IV.11 : Les courbes ROC relatives aux SVMs

### III.2.3. Combinaison des classificateurs

Durant ces dernières années, la combinaison des classificateurs est devenue un sujet de recherche assez motivant qui vise à améliorer les performances de la classification artificielle des données. Dans cette section, nous utilisons les règles de décision de la stratégie de vote, y compris le vote majoritaire (MV) et la moyenne des probabilités (AP) [120-121] pour fusionner les deux algorithmes SVMs. En effet, tous les tests de cette section sont réalisés en

mode de validations croisées de *k-folders* ( $k_f = 10$ ). En effet, les résultats illustrés dans le Tableau IV-8 montrent que la règle de décision AP a amélioré la performance des deux classificateurs en termes de F-Mesure et d'exactitude. En utilisant la méthode 2DPCA, ces deux paramètres atteignent respectivement : (98.50%, 98.50%) pour ORL, (99.91%, 99.91%) pour Yale et (98.30%, 98.28%) pour AR. De plus, en utilisant 2DLDA, ils atteignent respectivement : (99.20%, 99.25%) pour ORL ; (99.91%, 99.91%) pour Yale et (98.50%, 98.58%) pour AR.

Tableau IV-8 : Résultats de la fusion des SVMs par règle de décision AP

Méthodes		BD	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
LGPP+2DDWT	2DPCA	ORL	98.67	98.52	98.50	98.50
		Yale	99.91	99.91	99.91	99.91
		AR	98.40	98.30	98.30	98.28
	2DLDA	ORL	99.20	99.20	99.20	99.25
		Yale	99.91	99.91	99.91	99.91
		AR	98.65	98.52	98.50	98.50

Tableau IV-9 : Résultats de la fusion des SVMs par règle de décision MV

Méthodes		BD	Précision (%)	Rappel (%)	F-mesure (%)	Exactitude (%)
LGPP+2DDWT	2DPCA	ORL	98.37	98.27	98.22	98.25
		Yale	99.91	99.91	99.91	99.91
		AR	98.40	98.30	98.30	98.28
	2DLDA	ORL	98.97	98.87	98.85	98.85
		Yale	99.91	99.91	99.91	99.91
		AR	98.64	98.52	98.50	98.50

En outre, les résultats illustrés dans le Tableau IV-9 montrent que la règle de décision MV a amélioré la performance des deux classificateurs en termes de F-Mesure et d'exactitude. En utilisant la méthode 2DPCA, ces deux paramètres atteignent respectivement : (98.22%, 98.25%) pour ORL, (99.91%, 99.91%) pour Yale et (98.30%, 98.28%) pour AR. De plus, en utilisant 2DLDA, ils atteignent respectivement : (98.85%, 99.85%) pour ORL ; (99.91%, 99.91%) pour Yale et (98.50%, 98.50%) pour AR. Certes, les améliorations des performances

réalisées par fusion sont expliquées par l'appartenance des instances males classées de chaque classificateur individuel. En effet, plus les classes de ces instances diffèrent d'un classificateur à l'autre, plus la probabilité qu'elles soient bien classées par fusion augmente.

### III.3. Comparaison avec d'autres méthodes

Dans cette section, nous présentons les résultats de comparaison de notre approche avec d'autres travaux de pointe. Ces travaux ont été choisis en raison de leurs méthodes étroitement liées aux composants de notre approche. En fait, nous avons pris en compte des travaux portant sur le LBP, les ondelettes, la réduction de la dimension et classification par l'apprentissage artificiel. De plus, étant donné que le temps d'exécution dépend des caractéristiques physiques de la machine d'exécution, nous limiterons la comparaison au taux de reconnaissance.

Tableau IV-10 : Comparaison avec des méthodes de l'état de l'art

ORL	Méthodes	DSNPE	DWT +PCA	DWT +LDA	2DPCA +PCA	2DPCA +LDA	GGZ + HOG	LGPP+2DDWT+SVM	
	Taux(%)	96	94.2	97.1	92.1	92.1	98	2DPCA	2DLDA
	Réfs	[148]	[149]	[149]	[150]	[150]	[151]	Notre approche	
Yale	Méthodes	Gabor + DSNPE	LPP	Gabor +2DPCA	B(2D) 2PCA	Gabor +LDA	GGZ +HOG	LGPP+2DDWT+SVM	
	Taux(%)	93.5	90.4	93.73	93.73	93.3	97.8	2DPCA	2DLDA
	Réfs	[148]	[152]	[153]	[154]	[155]	[151]	Notre approche	
AR	Méthodes	1DLBP + DWT	SCR	LLE +SVM	PCA+ SVM	2DPCA +Gabor	BG2D (2LRP)	LGPP+2DDWT+SVM	
	Taux(%)	98.3	97.5	95	92.67	96.15	96.82	2DPCA	2DLDA
	Réfs	[138]	[158]	[159]	[145]	[156]	[157]	Notre approche	

En effet, les résultats illustrés dans le Tableau IV-10 montrent que les performances de l'approche proposée sont significatives, particulièrement avec le jeu de données Yale. Ainsi, les taux atteints avec ORL sont très compétitifs avec ceux rapportés dans [148-151]. Ceux atteints avec Yale sont plus élevés que ceux rapportés dans [148,151-155]. De plus, ceux atteints avec AR sont comparables à ceux rapportés dans [138,145,156-159]. De plus, les

meilleurs résultats sont obtenus respectivement avec les deux combinaisons: LGPP+2DDWT+2DPCA+SVM et LGPP+2DDWT+2DLDA+SVM. Elles ont atteints respectivement (98. 4 et 98. 25 avec ORL), (99. 69 avec Yale) et (98. 09 et 98. 45 avec AR).

## **IV. Conclusion**

Dans ce chapitre, nous avons présenté le nouveau modèle probabiliste gradient local (LGPP). C'est un descripteur d'extraction de caractéristiques locales qui utilise les concepts gradient et l'intervalle de confiance pour procréer une nouvelle description de l'image où les zones quasi-homogènes et les zones de pics sont notoirement séparées. Tout d'abord, nous avons exprimé de nouveaux intervalles de confiance en utilisant les voisinages des gradients, ensuite nous avons utilisé ces expressions, de l'intervalle de confiance, pour formuler notre descripteur LGPP. En plus, nous avons montré que la qualité de représentation de l'image par LGPP augmente avec la taille de cet intervalle de confiance jusqu'à atteindre un état permanent pour une taille de  $8\sigma$ . Après, nous avons mis en évidence les principaux avantages de ce descripteur en l'appliquant sur des images affectées par des changements d'illumination, des déformations structurelles et d'ajout de bruits. Enfin, nous avons conçu un système de reconnaissance faciale intégrant les méthodes LGPP, 2DDWT, 2DPCA, 2DLDA, LIBSVM, SMO, F-MLP et CS-MLP. Dans ce système, chaque image de visage de taille  $112 \times 92$  est représentée par un vecteur descripteur de très faible taille, 168 attributs permettant d'optimiser le temps de classification. Finalement, nous avons mené des expériences utilisant les bases de données ORL, Yale et AR. Les résultats de ces expériences ont montré que toutes les méthodes appliquées durant le processus de la reconnaissance ont efficacement contribué à l'amélioration de la performance globale du système de reconnaissance en termes de taux de reconnaissance et temps d'exécution. Aussi, la comparaison de l'approche proposée avec les méthodes récentes de l'état de l'art a montré sa compétitivité et sa prépondérance.

Dans le chapitre suivant, nous proposons une nouvelle vision qui prend en considération la surface 3D, les interactions lumière-objets et la conversion énergie-pixels. C'est le modèle binaire à champ électrique Virtuel (Electric Virtual Binary Pattern, EVBP).

---

---

## **Chapitre V. Modèle Binaire à Champ Électrique Virtuel (Electric Virtual Binary Pattern, EVBP) : Application à la Biométrie Faciale**

---

---

Notre système visuel parvient à ressentir les surfaces 3D grâce au phénomène de la réflexion-diffuse de la lumière. Dans ce phénomène, chaque surface 3D est une union de plusieurs surfaces élémentaires sur lesquelles des rayons lumineux sont réfléchis et/ou dispersés dans différentes directions. Dans ce chapitre, nous nous basons sur cette conséquence ainsi que le principe du balayage de Poincaré pour concevoir notre nouveau descripteur local nommé modèle binaire à champ électrique virtuel (Electric Virtual Binary Pattern, EVBP). Afin de formuler ce descripteur, nous modélisons chaque voisinage de l'image par une grille de charges électriques virtuelles.

## **I. Introduction**

Dans les deux chapitres précédents, nous avons présenté et étudié nos deux variétés, LBPP et LGPP, utilisant l'estimation probabiliste des intervalles de confiance. Nous avons montré qu'elles ont comblé certaines lacunes liées au descripteur LBP d'origine. Dans ce chapitre, nous proposons une nouvelle variété qui se base sur l'ensemble des opérations réalisées lors de la construction de l'image d'un objet réel en partant de la topologie de sa surface. Cette nouvelle variété appelée modèle binaire à champ électrique virtuel (Electric Virtual Binary Pattern, EVBP) génère des micros motifs, tout en se basant sur la dualité « pixel - 'charge charges électriques virtuelles' » [160] et le principe de balayage évoqué dans le problème de l'uniformisation de Poincaré [161]. Dans ce cadre, chaque voisinage de l'image est considéré comme une grille de charges électriques virtuelles en équilibre électrostatique. Une telle description est requise du fait que chaque surface élémentaire de l'objet réel est représentée par un nombre de pixels au niveau de l'image obtenue par transposition de l'énergie lumineuse en énergie électrique, puis en valeurs de niveau de gris. Dans ce contexte, nous exprimons les interactions entre chaque paire de charges situées sur chacune des quatre directions spatiales usuelles de la grille (verticale, horizontale et diagonales). Ensuite, nous utilisons ces expressions dans la formulation du descripteur EVBP. Enfin, nous nous inspirons du principe du balayage pour déterminer les seuils de segmentation de la fonction de décision d'EVBP. L'impact immédiat de ce descripteur sur la description faciale réside dans son pouvoir de localisation des principaux traits du visage. Un autre avantage marquant de cette description est la restriction de la plage des histogrammes à 16 étiquettes d'histogramme seulement et réduisant la taille du vecteur descripteur eLBPH de 93,75%. Finalement, et dans le but d'évaluer les performances de la représentation faciale via EVBP, nous intégrons ce dernier dans notre système de reconnaissance faciale. Dans un premier temps, nous mettons en évidence l'intérêt de la réduction de la taille du vecteur descripteur eLBPH en réalisant des tests de validation utilisant des métriques de mesure de similarité entre histogrammes, y compris  $\chi^2$  et les deux normes euclidiennes ( $L^1$  et  $L^2$ ). Dans un deuxième temps, nous introduisons des méthodes de réduction de la dimension statistique. Les résultats obtenus à l'aide des bases de données ORL, Yale et AR ont démontré l'efficacité et la compétitivité de l'approche basée sur EVBP vis-à-vis des méthodes récentes de l'état de l'art.

## II. Principe du descripteur EVBP

### II.1. Présentation du descripteur EVBP

Notre descripteur EVBP vise à améliorer la représentation faciale en intégrant le modèle champs électrique virtuel à la description locale des images. Dans ce modèle, chaque voisinage de l'image est considéré comme étant une grille de charges électriques virtuelles en équilibre électrostatique. Dans ce contexte, la topologie locale de chaque élément de surface réel est décrite au niveau de l'image par l'ensemble des interactions entre les charges virtuelles de la grille qui lui correspond. Pour atteindre un tel objectif, nous joignons à chaque voisinage  $3 \times 3$  un repère lié aux quatre directions spatiales usuelles (axe vertical, axe horizontal et 2 axes diagonaux). Ensuite, et en mesurant les interactions entre ces charges virtuelles suivant chacune des quatre directions spatiales nous définissons des règles de décision qui seront utilisées pour formuler l'expression de notre descripteur EVBP.

### II.2. Modélisation du descripteur EVBP

En général, l'image numérique est une fonction à valeurs discrètes qui représente une scène par un pavage de pixels. Par conséquent, chaque pixel de l'image représente implicitement la quantité de lumière transmise par réflexion-diffuse sur un élément de surface local de l'objet. Cette quantité dépend principalement de sa géométrie locale (concavité et convexité) et des effets d'éclairage qui l'affectent (luminosité, contraste et bruit). En se basant sur cette idée, EVBP génère une nouvelle description où la présence ou l'absence de fortes variations du gradient autour d'un pixel est déterminée par le niveau de dissymétrie des interactions, suivant chacune des quatre directions de sa grille correspondante. Cette dissymétrie vient du fait que l'énergie reçu par chaque point du capteur dépend principalement de la géométrie locale de l'élément de surface réel et de ses caractéristiques optiques.

#### II.2.1. Modèle champ électrique virtuel

Dans la théorie électrostatique, chaque charge  $q_0$  placée en un point d'espace  $M_0$  produit une perturbation dans l'espace qui l'entoure appelée champ électrostatique. Ainsi, à chaque point  $M_k$  de l'espace, cette perturbation  $\vec{E}_{0k}$  est donnée par :

$$\vec{E}_{0k} = K q_0 \frac{\overrightarrow{M_0 M_k}}{\|\overrightarrow{M_0 M_k}\|^3} \quad \text{où, } K \text{ est une constante électrostatique} \quad (\text{V.1})$$

De plus, la présence des charges électriques  $q_0$  et  $q_k$  respectivement aux deux points  $M_0$  et  $M_k$  induit entre ces deux charges une force électrostatique mutuelle  $\vec{F}_{0k}$  exprimée par :

$$\vec{F}_{0k} = \vec{F}(q_k \rightarrow q_0) = q_k \vec{E}_{0k} = K q_0 q_k \frac{\overrightarrow{M_0 M_k}}{\|\overrightarrow{M_0 M_k}\|^3} \quad (\text{V.2})$$

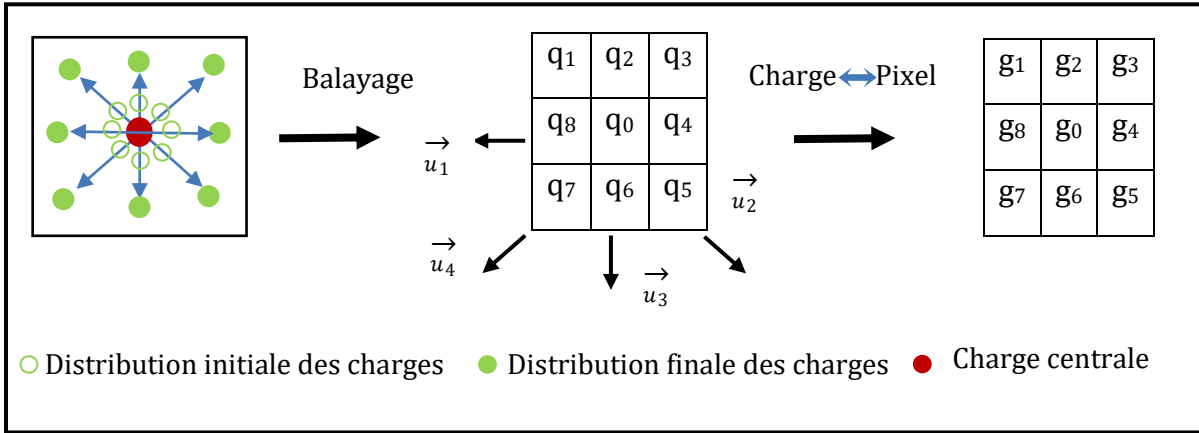


Figure V-1 : Analogie entre pixel et charge électrique virtuelle

Par ailleurs, la Figure V-1 montre que sur une grille  $3 \times 3$ , les forces électrostatiques entre la charge centrale  $q_0$  et ses huit charges voisines  $q_k$  ( $k = 1, \dots, 8$ ) sont exercées selon les quatre directions spatiales suivantes : horizontale ( $\vec{u}_1$ ), vertical ( $\vec{u}_3$ ), et diagonales ( $\vec{u}_2$  et  $\vec{u}_4$ ). Ainsi, à chaque charge  $q_k$  de la grille est associée une valeur en niveau de gris  $g_k$ . Cette association va nous permettre de généraliser les expressions suivantes pour formuler le descripteur EVBP.

Les forces horizontale ( $\vec{F}_h$ ) et verticale ( $\vec{F}_v$ ) exercées suivant les directions  $\vec{u}_1$  et  $\vec{u}_3$  sont respectivement exprimées par :

$$\vec{F}_h = \vec{F}_{04} + \vec{F}_{08} = \frac{Kq_0}{R^2} (q_8 - q_4) \vec{u}_1 \quad (\text{V.3})$$

$$\vec{F}_v = \vec{F}_{02} + \vec{F}_{06} = \frac{Kq_0}{R^2} (q_6 - q_2) \vec{u}_3 \quad (\text{V.4})$$

Les forces diagonales  $\vec{F}_{d1}$  et  $\vec{F}_{d2}$  exercées suivant les directions  $\vec{u}_2$  et  $\vec{u}_4$  sont exprimées par :

$$\vec{F}_{d1} = \vec{F}_{01} + \vec{F}_{05} = \frac{Kq_0}{(\sqrt{2}R)^2} (q_5 - q_1) \vec{u}_2 \quad (\text{V.5})$$



$$\vec{F}_{d2} = \vec{F}_{03} + \vec{F}_{07} = \frac{Kq_0}{(\sqrt{2}R)^2} (q_7 - q_3) \vec{u}_4 \quad (\text{V.6})$$

Par conséquent, la force électrostatique résultante  $\vec{F}_t$  associée à la charge  $q_0$  est exprimée par :

$$\vec{F}_t = \vec{F}_h + \vec{F}_v + \vec{F}_{d1} + \vec{F}_{d2} = \sum_{j=1}^4 F_j \vec{u}_j \quad (\text{V.7})$$

Afin de déterminer l'expression du descripteur EVBP d'un pixel de l'image, nous associons à son voisinage de  $3 \times 3$  pixels une grille élémentaire de  $3 \times 3$  charges électriques virtuelles. Ensuite, nous regroupons deux à deux et suivant chacune des quatre directions ( $\vec{u}_1, \vec{u}_3, \vec{u}_2$  et  $\vec{u}_4$ ) les huit charges voisines de cette grille. Enfin, nous calculons les expressions des quatre interactions électrostatiques ( $\vec{F}_h, \vec{F}_v, \vec{F}_{d1}$  et  $\vec{F}_{d2}$ ) qui régissent les quatre paires de charges précédemment définies. Ces expressions sont finalement utilisées pour la formulation mathématique du descripteur EVBP.

### II.2.2. Formulation du descripteur EVBP

En associant à chaque charge électrique virtuelle  $q_k$  une valeur de niveau de gris  $g_k$  ( $k=0, \dots, 8$ ), nous généralisons les équations (V.3 à V.6) pour formuler le descripteur EVBP au sein de chaque pixel de l'image. Ainsi, nous définissons l'opérateur EVBP pour chacune des quatre directions ( $\vec{u}_j$ ;  $j = 1 \dots 4$ ) par l'expression suivante :

$$EVBP_{R, \vec{u}_j}(g_j, g_{(j+4)}) = \frac{K g_0}{r_{0j}^2} |g_{(j+4)} - g_j| \text{ avec } j = 1, 2, 3, 4 \quad (\text{V.8})$$

$$r_{0j} = \|\vec{g}_0 \vec{g}_j\| = \begin{cases} R & \text{if } j = 2, 4 \\ \sqrt{2}R & \text{if } j = 1, 3 \end{cases} \quad \text{Où, } R \text{ est le rayon du voisinage} \quad (\text{V.9})$$

Les quatre expressions ainsi obtenues sont utilisées par la suite pour attribuer un code binaire de 4 bits à chaque pixel de l'image (Figure V-2). Chacun des quatre bits prend une valeur égale à « 0 » si la valeur de «  $EVBP_{R, \vec{u}_j}$  » est supérieure à un seuil donné  $T$ . Sinon il prend une valeur égale à « 1 ». D'où l'expression finale du descripteur EVBP :

$$EVBP_{8,1} = \sum_{j=1}^4 s(EVBP_{\vec{u}_j}(g_j, g_{(j+4)})) 2^{j-1} \quad (\text{V.10})$$

$$\text{Où, } s(x) = \begin{cases} 0, & x \geq T \\ 1, & x < T \end{cases}; \quad T \in \mathbf{R} \text{ (seuil relatif à EVBP)} \quad (\text{V.11})$$

Partant de l'expression de l'équation (V.10), les valeurs de niveaux de gris générées par EVBP sont toutes situées dans l'intervalle  $[0 : 15]$ . Un tel intervalle est assez condensé et ne permet aucune distinction entre les différentes zones du visage. Pour remédier à ce problème, nous proposons de transformer les valeurs de niveaux de gris générées pour qu'elles soient comprises dans l'intervalle  $0 : 255]$ . Pour ce faire, nous avons remplacé le terme  $2^{j-1}$  dans l'équation (V.10) par le terme  $2^{j-1}+2^{j+3}$  dans l'équation (V.12). Par conséquent, nous générons 16 nouvelles valeurs de niveaux de gris étalées sur toute la plage  $0 :255]$  et assurant au mieux la séparation entre les différentes zones de visage. D'où l'expression de EVBP :

$$EVBP_{8,1} = \sum_{j=1}^4 s(EVBP_{\vec{u}_j} (g_j, g_{(j+4)}))(2^{j-1} + 2^{j+3}) \quad (V.12)$$

Pour généraliser cette description aux différentes résolutions, nous adoptons le ré-échantillonnage des P pixels voisins dans l'expression suivante :

$$EVBP_{P,R} = \sum_{j=1}^{P/2} s(EVBP_{\vec{u}_j} (g_j, g_{(j+P/2)}))(2^{j-1} + 2^{j-1+P/2}) \forall (P, R) \quad (V.13)$$

### II.2.3. Seuil du descripteur EVBP

Afin de définir le seuil compatible avec l'hypothèse du modèle de champ électrique virtuel, nous faisons appel au principe du balayage. En fait, ce principe consiste à modéliser la construction d'une surface locale 3D par un phénomène de balayage des éléments de cette surface. En effet, ce phénomène suppose que la construction d'une surface passe par des états successifs en allant de l'état initial à l'état final. Dans l'état initial, tous les éléments de la surface sont agglomérés à proximité de son centre de gravité. Dans chaque état ultérieur ces éléments s'éloignent les uns des autres dans toutes les directions jusqu'à atteindre l'état final où ces éléments sont en équilibre. Les résultats d'éloignement sont évalués dans notre descripteur suivant les quatre directions usuelles. Appliquons ce principe de balayage à la grille  $3 \times 3$  supposée être formée par une charge ponctuelle positive  $q_0$  située au centre de la grille, et une distribution de charges négatives qui l'entoure. Initialement, la distribution des charges négatives est prise d'une façon aléatoire avec une charge totale ( $dQ_n$ ). Par conséquent, le principe de balayage consiste à distribuer les charges négatives vers les bords de la grille et suivant les quatre directions spatiales ( $\vec{u}_j; j = 1 \dots 4$ ). Cette opération se produit tout en

---

conservant la charge totale de la grille :  $dQ_t = dQ_n + q_0$ . En effet, pour la grille représentant une zone quasi-homogène, les charges négatives sont uniformément réparties suivant les quatre directions usuelles. Par conséquent, les interactions entre la charge centrale positive et les charges de bord négatives sont semblables. Ce qui se manifeste par un mot de code EVBP qui vaut généralement « 11111111 ». Par conséquent, EVBP code d'une manière semblable toutes les zones quasi-homogènes de l'image. Alors que pour une grille représentant une zone de pics, les charges négatives sont aléatoirement réparties suivant les quatre directions usuelles. Les interactions entre la charge centrale positive et les charges de bord négatives sont dissymétriques. Ce qui se manifeste par un mot de code EVBP dont la succession de ses quatre bits est aléatoire. Par conséquent, EVBP arrive à localiser toutes les zones de l'image présentant de fortes variations du gradient d'intensité. Ces zones de pics sont généralement associées aux éléments de surface de l'objet présentant de forte courbure (Figure V-4). Pour interpréter le mode de codage d'EVBP, nous associons à la valeur du seuil  $\mathbf{T}$  de la fonction de décisions de l'équation (V.11) une charge électrique virtuelle  $q_T$ . En plus, nous supposons que chaque charge voisine  $q_k$  est composée de la charge  $q_T$  additionnée à une perturbation  $\varepsilon_k$ . L'équation de conservation de la charge totale de la grille devient :

$$dQ_t = dQ_n + q_0 \Rightarrow dQ_t = q_0 + \sum_{k=1}^P q_k \quad (\text{V.14})$$

$$\exists(q_T, \varepsilon_k) / q_k = q_T + \varepsilon_k \Rightarrow q_0 + Pq_T + \sum_{k=1}^P \varepsilon_k = 0 \quad (\text{V.15})$$

Pour une grille 3×3 cette équation devient :

$$q_0 + 8q_T + \sum_{k=1}^8 \varepsilon_k = 0 \quad (\text{V.16})$$

En outre, le seuil  $q_T$  est déterminé par l'analyse de la description des zones quasi-homogène par un balayage quasi-idéal des charges virtuelles. Dans ce balayage, les charges virtuelles négatives migrent de la même façon du centre de la grille vers ses huit bords. Par conséquent, le terme  $\sum_{k=1}^8 \varepsilon_k = 0$ , et donc  $q_T = -q_0/8$ . De plus, cette valeur de  $q_T$  est traduite par la valeur en niveau de gris  $g_T = g_0/8$ . Ainsi, le choix du seuil local  $T$  du descripteur EVBP est tel que :  $T = (g_0/8) \cdot Kg_0 = Kg_0^2/8$ . En conséquence, ce seuil est utilisé par EVBP pour évaluer le degré de fluctuations géométriques et photoniques au niveau de l'élément de surface qui correspond au voisinage de l'image centré sur le pixel courant. La Figure V-2 illustre un exemple de calcul de

---

l'opérateur EVBP. Il consiste à convertir chaque voisinage en une grille de charges fictives et à utiliser le modèle électrique virtuel pour évaluer leurs interactions. Enfin, il engendre une nouvelle valeur en niveau de gris pour chaque pixel de l'image suivant l'équation (V.13).

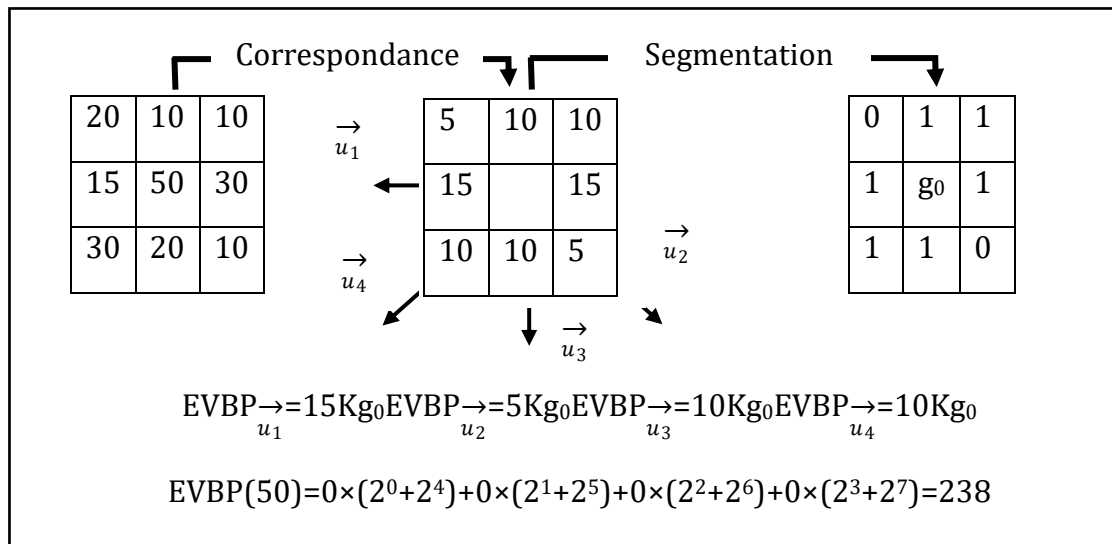


Figure V.2 : Exemple de codage du descripteur EVBP

### II.3. Les avantages du descripteur EVBP

En tant que variété de l'opérateur LBP, notre descripteur EVBP retient les mêmes avantages que ce dernier, spécialement en ce qui concerne la faible complexité et l'invariance aux changements monotone d'éclairage. En plus, EVBP présente d'autres avantages qui le qualifient pour être appliqué dans autant d'application en reconnaissance de formes et vision par ordinateur. Ces avantages sont généralement liés aux changements d'illumination, à l'addition des bruits et aux pouvoir de détection des zones présentant de fortes variations du gradient.

- Changement d'illumination : deux cas sont distingués : dans le premier cas, les changements d'illumination affectent en même temps le premier plan et le fond (changement global). Alors que dans le deuxième cas, le premier plan et/ou le fond sont affectés séparément (changement local). En ce qui concerne le changement global, la transformation (a) de la Figure V-3 montre que les deux descripteurs, LBP et EVBP, ont le comportement vis-à-vis ce type de changement. En effet, chacun des deux descripteurs conserve son motif initial (11111000 pour LBP et 01110111 pour EVBP). Mais, leurs

comportements vis-à-vis les changements locaux sont différents. En effet, les transformations (b et c) de la Figure V-3 montrent que seul EVBP conserve son motif initial (01110111), alors que celui de LBP change (00000000 et 00000111). Par conséquent, EVBP représente un outil efficace pour résoudre des problèmes où le traitement des effets d'éclairage est une procédure essentielle (analyse des séquences vidéo, analyse de mouvement et élimination des ombres etc.).

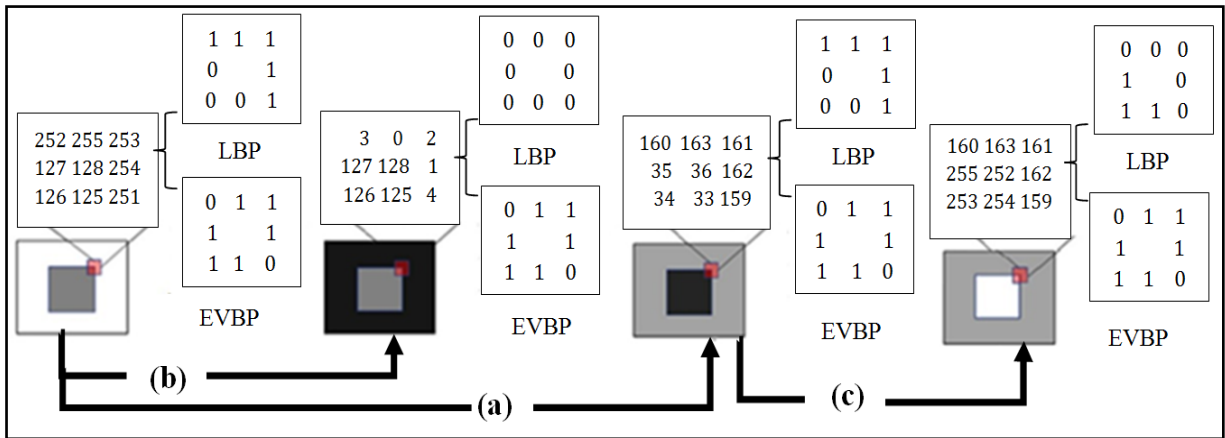


Figure V.3 : Comportement d'EVBP vis-à-vis les changements d'illumination

- Invariance aux bruits : en plus de ses performances envers les changements d'illumination, EVBP a un comportement remarquable vis-à-vis les bruits. En effet, nous avons mis en évidence cet effet en testant EVBP sur le bruit gaussien et celui du flou. Notamment, la Figure V-4- (a) montre que EVBP permet de corriger l'effet d'un bruit gaussien dont la valeur de l'écart-type est faible ou moyenne ( $\sigma < 10$ ). Cependant, cette qualité se dégrade tout en augmentant la valeur de l'écart-type ( $\sigma > 10$ ). De plus, et à partir de la Figure V-1- (b), nous remarquons qu'EVBP est efficace vis-à-vis les bruits du flou. En effet, il conserve la visibilité des principaux composants du visage surtout pour des rayons du masque flou inférieurs à cinq. Cette propriété permet à EVBP de traiter les problèmes où la dégradation de l'image est modérée (effet des conditions et des supports d'acquisition).
- Autres perspectives d'application : afin de découvrir le comportement du descripteur EVBP dans différents contextes du traitement d'images, nous avons appliqué ce dernier sur des images extraites de différents domaines de la vision par ordinateur.

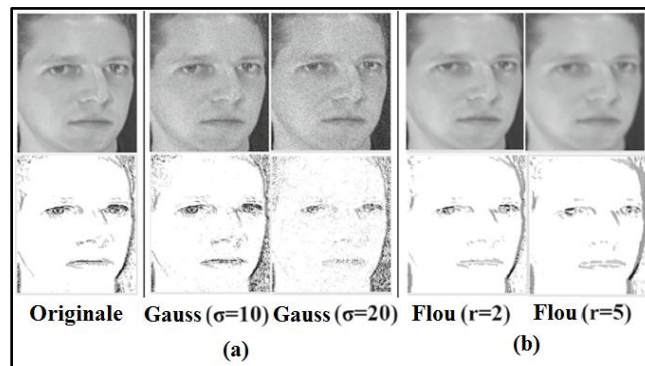


Figure V.4 : Comportement d'EVBP vis-à-vis des bruits Gaussien et flou

La Figure V.5 illustre les résultats de l'application d'EVBP sur des images contenant respectivement des visages, des textes, des contours et des objets divers. En effet, les trois exemples de la Figure V.5- (a) montrent que l'application d'EVBP sur des images de visages qui sont capturées sous différentes conditions permet de mieux localiser les principaux composants de la surface faciale. En outre, les trois exemples de la Figure V.5-(b) montrent qu'EVBP réalise un prétraitement utile pour l'analyse des documents textes. Ce prétraitement permet d'éliminer efficacement les différentes sortes d'anomalies qui peuvent affecter des images contenant des textes. En effet, EVBP assure la soustraction de l'arrière-plan dans le premier exemple, l'élimination des taches dans le second et le retrait des ombres dans le troisième. De même, les trois exemples de la Figure V.5-(c) montrent qu'EVBP arrive à détecter les zones présentant de fortes variations de gradient. En effet, dans ces trois exemples, EVBP détecte et accentue efficacement les contours et les bords des objets contenus dans l'image. Enfin, les trois exemples de la Figure V.5-(d) montrent qu'EVBP peut représenter un outil approprié pour les applications où la soustraction de l'arrière-plan est une étape indispensable. Dans ces trois exemples, les objets du premier-plan sont parfaitement segmentés. En plus des perspectives proposées dans les exemples de la Figure V.5, EVBP peut aussi être utilisé dans plusieurs autres applications où l'amélioration de l'extraction des caractéristiques de la texture est souhaitable. Spécialement le domaine de reconnaissance faciale. Dans ce domaine, la taille des vecteurs descripteurs influence les performances du système de reconnaissance. EVBP représente une solution prometteuse à ce problème.

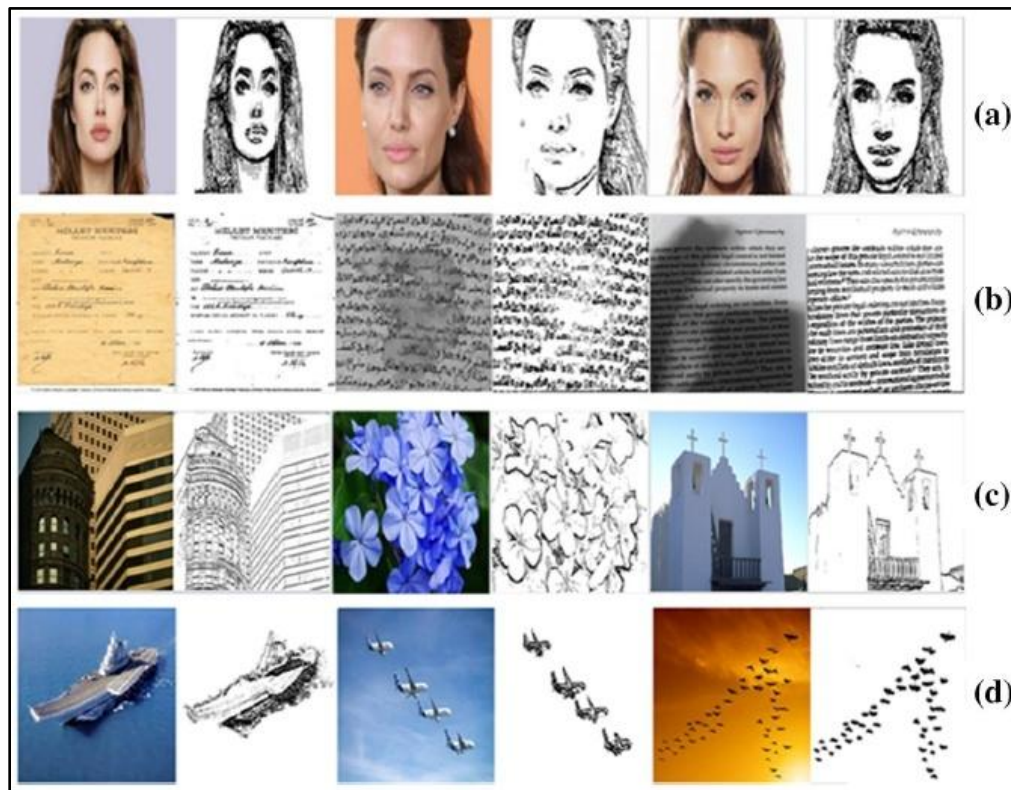


Figure V-5 : Autres perspectives d'application d'EVBP

Taille du vecteur caractéristique : dans les applications de la reconnaissance de visages, le descripteur LBP d'origine ainsi que ses différentes variétés utilisent l'algorithme des histogrammes LBP spatialement amélioré (Spatially enhanced Local Binary Pattern Histogram, eLBPH). Initialement proposé par Ahonen [16,56], eLBPH est un algorithme de référence découpe l'image en zones locales et concatène leurs histogrammes en un seul vecteur caractéristique représentant le visage. En effet, dans LBP d'origine eLBPH utilise un nombre de 28 étiquettes d'histogrammes, ce nombre est réduit à 59 étiquettes distinctes avec LBP uniforme ( $LBP^{u2}$ ). Le Tableau V-1, illustre le nombre étiquettes distinctes pour quelques variétés de LBP, y compris LTP, LDP et LGP. Pour notre descripteur EVBP, le mode de codage sur 4 bits permet de générer seulement 24 étiquettes d'histogrammes distinctes avec une réduction de 93,75% par rapport à LBP. En plus, les valeurs des étiquettes d'histogrammes engendrées par EVBP selon l'équation (V.12) sont tel que:  $\{\forall p; p = \text{sig}(b_j) \times (2^{j-1} + 2^{j+3}) \text{ avec } \text{sig}(b_j) = 0 \text{ ou } 1 \text{ et } j = 1,2,3,4\} = \{0, 17, 34, 51, 68, 85, 102, 119, 136, 153, 170, 187, 204, 221, 238, 255\}$ . Ces valeurs ont permis à eLBPH de construire des vecteurs

caractéristiques de faible taille. L'impact immédiat de cette description se manifeste au niveau de la phase classification de notre système de reconnaissance.

Tableau V-1 : Taille du vecteur eLBPH pour des variétés de LBP (56 blocs)

Méthodes	Nombre de motifs	Taille du vecteur eLBPH	Taux de réduction	Réfs
LBP	256	$7*8*256=14336$	0%	[55]
LBP <sup>u2</sup>	59	$7*8*59=3304$	76.95%	[55]
LTP	118	$7*8*118=6608$	53.91%	[122]
LGP	59	$7*8*59=3304$	76.95%	[123]
LDP	56	$7*8*56=3136$	78.12%	[131]
<b>EVBP</b>	<b>16</b>	$7*8*16=896$	<b>93.75%</b>	Notre approche

### III. Reconnaissance de visage par EVBP

#### III.1. Le système de reconnaissance faciale proposé

Dans notre système de reconnaissance, chaque visage détecté et normalisé est traité par le descripteur EVBP, puis par eLBPH pour générer un vecteur descripteur de taille réduite. L'ensemble des vecteurs descripteurs de chaque base de données est classifié en utilisant les métriques de mesure de la similarité des histogrammes.

##### III.1.1. Prétraitement

Dans cette phase un recadrage géométrique et une égalisation d'histogramme sont respectivement réalisés pour éliminer les éléments entravant la performance globale du système et assurer l'équiprobabilité des niveaux de gris de l'image.

##### III.1.2. Traitement par le descripteur EVBP

Dans cette phase, les images normalisées sont traitées par le descripteur EVBP. Ce dernier utilise la dualité entre le voisinage de pixels et la grille de charges virtuelles pour générer une description où la présence des interactions dissymétriques entre charges virtuelles se traduit par la localisation des principaux composants du visage. Des exemples montrant l'effet de la mise en correspondance pixel-charge sont illustrés dans la Figure V.7.



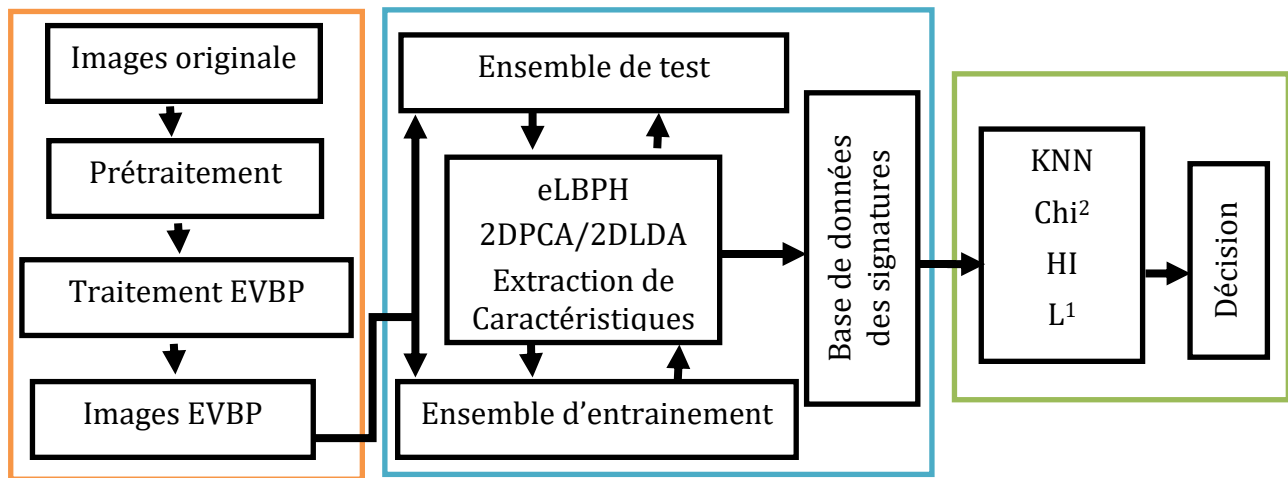


Figure V.6 : Diagramme blocs du système de reconnaissance proposé

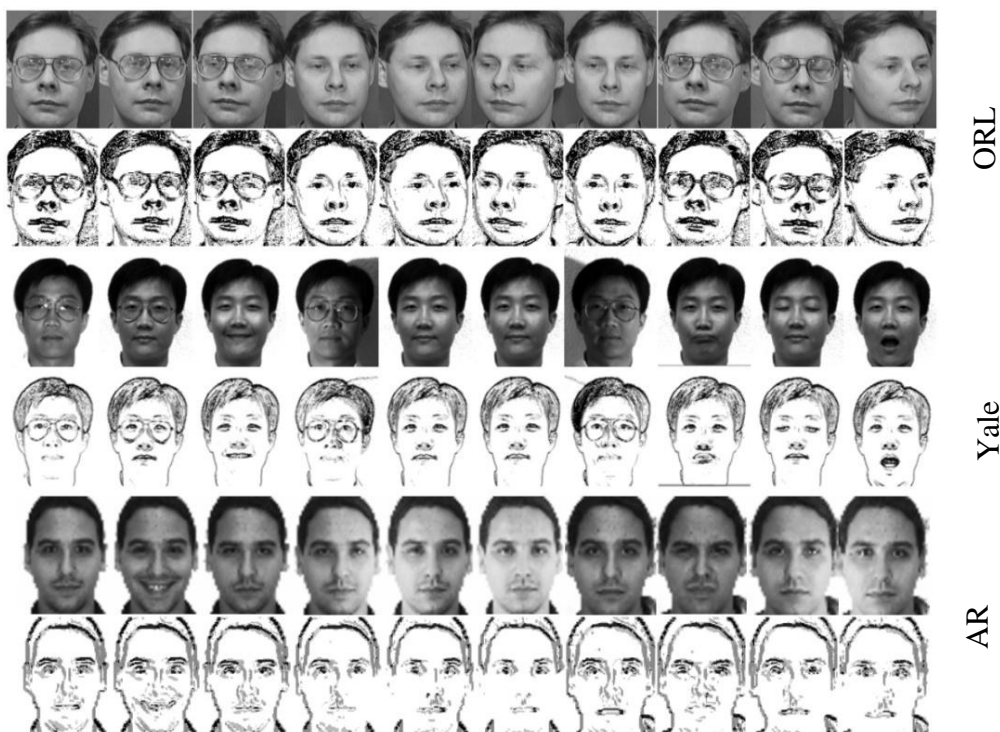


Figure V.7 : Exemples d'images obtenues par EVBP pour ORL, Yale et AR.

### III.1.1. Vecteurs caractéristiques (eLBPH)

Dans cette phase, l'algorithme eLBPH engendre un vecteur descripteur en concaténant les  $m$  histogrammes locaux des  $m$  blocs non chevauchants  $B_1 \dots B_m$  de chaque image. Afin de

caractériser ce vecteur descripteur deux paramètres sont à déterminer, le nombre de blocs ( $n_b$ ) et le nombre d'étiquettes d'histogramme ( $h$ ). La Figure V.8 montre un exemple d'extraction d'un vecteur caractéristique eLBPH représentant une image de visage  $112 \times 92$ , de ORL, par un vecteur de taille  $L = n_b \times h = 8 \times 7 \times 16 = 896$  coefficients.

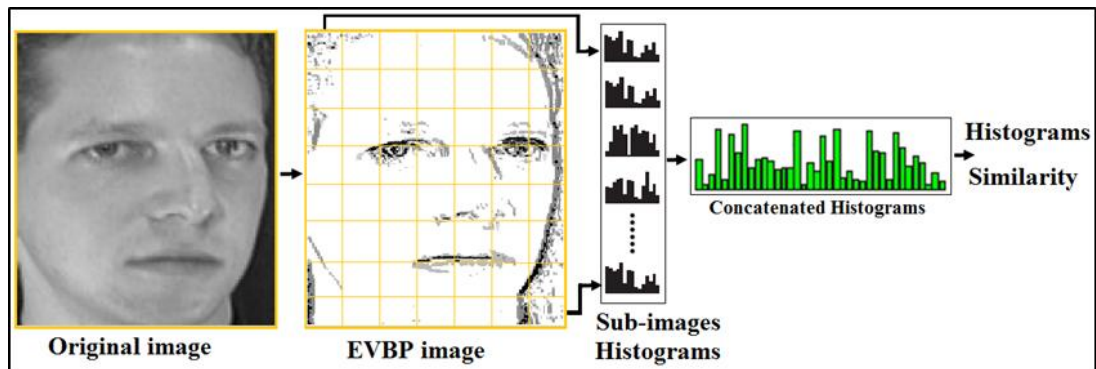


Figure V.8 : Les étapes d'extraction du vecteur caractéristique eLBPH.

### III.1.2. Classification

Dans cette phase, le classificateur k-plus proche voisin (KNN) utilise les techniques de correspondance entre histogrammes, notamment, Chi-square ( $\chi^2$ ), intersection des histogrammes (HI), norme Euclidien ( $L^2$ ), et distance de Manhattan ( $L^1$ ) pour mesurer la similarité entre les vecteurs eLBPHs précédemment engendrés pour chaque base de données.

## III.2. Les résultats expérimentaux

Dans cette sous-section, nous évaluons notre système de reconnaissance faciale à travers deux phases expérimentales utilisant les bases de données ORL, Yale et AR. Les expériences de la première phase sont menées dans le but de déterminer les trois paramètres suivants : la taille optimale du vecteur descripteur, la métrique adéquate et le nombre optimal d'images par classe pris pour l'arrière-plan du système. Les expériences de la deuxième phase sont menées dans le but de valider le modèle résultat des premières expériences et de déterminer la combinaison idéale entre les méthodes, utilisées dans chaque phase de notre système, qui donne la meilleure performance du processus global de reconnaissance. Dans la première phase expérimentale, les expériences menées sont répétées 100 fois. Dans chaque répétition, les images de chaque base de données appartenant à  $n_c$  classes sont aléatoirement subdivisées

en un ensemble d'entraînement et un ensemble de test. Le premier comporte  $n_{tr}$  images par classe ; le second comporte  $n_{ts}$  images par classe. Ensuite, les valeurs optimales des paramètres critiques de chaque méthode seront choisies de manière à répondre au compromis entre le taux de reconnaissance et le temps d'exécution. Finalement et dans la deuxième phase expérimentale, les données sont analysées par l'algorithme de classification artificiel KNN en mode de validation croisée. Les performances de ce classificateur sont évaluées suivant le taux d'exactitude, la précision, le rappel et la F-mesure.

### **III.2.1. Taille optimale des blocs**

Le fonctionnement de l'algorithme eLBPH dépend du choix des valeurs des deux paramètres  $n_b$  et  $p$ , représentant le nombre et la taille des blocs non chevauchants de l'image. Pour déterminer leurs valeurs optimales nous avons menés des expériences où nous varions la valeur de  $p$  et nous mesurons pour chaque valeur le taux de reconnaissance et le temps d'exécution. Les résultats illustrés dans le Tableau V-2, représentent les valeurs enregistrées en termes de taux de reconnaissance et du temps d'exécution pour les cinq tailles de blocs ( $p=8,10, 12,14$  et  $16$ ) et en utilisant la métrique  $\text{Chi}^2$  pour la mesure de similarité des vecteurs de chacune des bases de données ORL, Yale et AR. Dans ces expériences, le temps d'exécution représente la durée écoulée le long de toutes les phases de reconnaissance et pour toute la base de données envisagée.

- La base de données ORL est subdivisée en deux ensembles : un pour l'arrière-plan avec  $n_{tr}=5$  et l'autre pour les tests avec  $n_{ts}=5$ . Les résultats montrent que la taille blocs  $14 \times 14$  permet d'atteindre un taux de reconnaissance de  $96.50 \pm 0.8\%$  en  $4.22$  s au lieu de  $95.50 \pm 0.8\%$  en  $7.09$  s pour la norme JPEG ( $8 \times 8$  pixels). Soit, un gain de  $40.0\%$  en termes de temps d'exécution et un gain de  $1.03\%$  en termes de taux de reconnaissance.
- La base de données Yale est pareillement subdivisée avec  $n_{tr}=5$  et  $n_{ts}=6$ . Les résultats montrent que la taille blocs  $14 \times 14$  permet d'atteindre un taux de reconnaissance de  $98.89 \pm 0.8\%$  en  $1.32$ s au lieu de  $98.50 \pm 0.9\%$  en  $2.06$ s pour la norme JPEG ( $8 \times 8$  pixels). Soit, un gain de  $35.92\%$  en termes de temps d'exécution et un gain de  $0.40\%$  en termes de taux de reconnaissance.

- La base de données AR est aussi subdivisée avec  $n_{tr}=n_{ts}=7$ . Les résultats montrent que la taille blocs  $14 \times 14$  permet d'atteindre un taux de reconnaissance de  $97.10 \pm 1.1\%$  en 28.77s au lieu de  $97.20 \pm 1.1\%$  en 52.62s pour la norme JPEG ( $8 \times 8$  pixels). Soit, un gain de 45.32% en termes de temps d'exécution et une perte de 0.1% en termes de taux de reconnaissance.

A partir des résultats précédents et en tenant compte du compromis taux de reconnaissance et temps d'exécution, nous choisissons la taille  $14 \times 14$  comme taille optimale des blocs pour la suite de l'expérimentation.

Tableau V-2 : Taille optimal des blocs

BD	ORL ( $n_{tr}=5, n_{ts}=5$ )		Yale ( $n_{tr}=5, n_{ts}=6$ )		AR ( $n_{tr}=7, n_{ts}=7$ )	
	Taux (%)	Temps(s)	Taux (%)	Temps(s)	Taux (%)	Temps(s)
$8 \times 8$	$95.50 \pm 0.8$	7.09	$98.50 \pm 0.9$	2.06	$97.20 \pm 1.1$	52.62
$10 \times 10$	$95.20 \pm 0.8$	5.97	$98.20 \pm 0.8$	1.82	$96.60 \pm 1.2$	41.45
$12 \times 12$	$95.80 \pm 0.9$	4.91	$98.10 \pm 0.8$	1.53	$96.10 \pm 1.3$	32.86
$14 \times 14$	$96.50 \pm 0.8$	4.22	$98.89 \pm 0.8$	1.32	$97.10 \pm 1.1$	28.77
$16 \times 16$	$95.90 \pm 0.9$	3.78	$97.78 \pm 0.8$	1.22	$96.60 \pm 1.3$	24.82

### III.2.2. Métrique adéquate

Afin de déterminer la métrique adéquate de mesure de similarité compatible avec la représentation faciale EVBP, nous menons des expériences où nous comparons les résultats obtenus à l'aide des quatre métriques :  $\chi^2$ , HI,  $L^2$  et  $L^1$ . Les Tableau V-3, Tableau V-4 et Tableau V-5 représentent respectivement les résultats enregistrés en termes de taux de reconnaissance et du temps d'exécution pour les trois bases de données ORL, Yale et AR. En effet, pour chaque base de données, nous appliquons ces quatre métriques pour mesurer respectivement la similarité entre les vecteurs descripteurs eLBPH engendrés respectivement à partir d'EVBP et de LBP<sup>u2</sup>.

#### a. Résultats de la base ORL

ORL est subdivisée avec  $n_{tr}=5$  et  $n_{ts}=5$ . Les résultats illustrés au Tableau V-3 montrent que les métriques utilisées donnent des bons résultats avec EVBP en termes de taux de

reconnaissance et temps d'exécutions. En effet, ce taux atteint respectivement les valeurs (93.00%, 95.50%, 96.20% et 96.50%) avec les métriques (HI,  $L^2$ ,  $L^1$  et  $\text{Chi}^2$ ). Notons que, la métrique  $L^1$  réduit au maximum le temps d'exécution avec 28.98% comparée à la métrique  $\text{Chi}^2$ , tout en restant comparable en termes de taux de reconnaissance avec une perte de 0.31% seulement.

Tableau V-3 : Métrique adéquate pour ORL

ORL ( $n_{tr}=5$ , $n_{ts}=5$ , $p=14$ )		$\text{Chi}^2$	HI	$L^1$	$L^2$
EVBP	Taux (%)	96.50	93.00	96.2	95.50
	Temps(s)	4.21	3.6	2.99	2.86
LBP <sup>u2</sup>	Taux (%)	94.00	91.00	91.50	89.5
	Temps(s)	27.24	18.86	17.65	16.44

### b. Résultats de la base Yale

Yale est subdivisée avec  $n_{tr}=5$  et  $n_{ts}=6$ . Les résultats illustrés au Tableau V-4 montrent que les métriques utilisées donnent des bons résultats avec EVBP en termes de taux de reconnaissance et temps d'exécutions. En effet, Ce taux atteint respectivement les valeurs (93.33%, 97.78%, 98.89% et 98.89%) avec les métriques (HI,  $L^2$ ,  $L^1$  et  $\text{Chi}^2$ ). Notons que la métrique  $L^1$  réduit de 29.14% le temps d'exécution par rapport à  $\text{Chi}^2$ , tout en restant comparable en termes de taux de reconnaissance avec une perte de 0.22% seulement.

Tableau V-4 : Métrique adéquate pour Yale

Yale ( $n_{tr}=5$ , $n_{ts}=6$ , $p=14$ )		$\text{Chi}^2$	HI	$L^1$	$L^2$
EVBP	Taux (%)	98.89	93.33	98.67	97.78
	Temps(s)	1.75	1.29	1.24	1.2
LBPu2	Taux (%)	85.56	82.22	85.56	81.11
	Temps(s)	3.65	3.27	3.53	3.05

### c. Résultats de la base AR

Pour AR est similairement subdivisée avec  $n_{tr}=7$  et  $n_{ts}=7$ . Les résultats illustrés au Tableau V-5 montrent que les métriques utilisées donnent des bons résultats avec EVBP qui atteint respectivement les taux de reconnaissance 89.86%, 90.29%, 96.51% et 97.15 avec les

métriques HI,  $L^2$ ,  $L^1$  et  $\text{Chi}^2$ . Notons que la métrique  $L^1$  réduit au maximum le temps d'exécution avec un taux de 38.54% comparée à  $\text{Chi}^2$ , tout en restant comparable en termes de taux de reconnaissance avec une diminution de 0.66% seulement.

Tableau V-5 : Métrique adéquate pour AR

AR ( $n_{tr}=7, n_{ts}=7, p=14$ )		$\text{Chi}^2$	HI	$L^1$	$L^2$
EVBP	Taux (%)	97.15	89.86	96.51	90.29
	Temps(s)	29.5	23.81	18.13	13.77
LBP <sup>u2</sup>	Taux (%)	78.86	73.43	75.00	51.00
	Temps(s)	81.46	47.38	41.45	36.28

A partir des résultats précédents, la distance Manhattan ( $L^1$ ) représente la métrique la plus qualifiée pour la phase de classification de notre système de reconnaissance. Un autre facteur important pour l'évaluation de notre approche est lié aux valeurs de  $n_{tr}$  et  $n_{ts}$ .

### III.2.3. Effet de l'ensemble d'apprentissage

Afin de confirmer la robustesse du descripteur EVBP, nous menons des expériences avec la taille blocs  $14 \times 14$  et la distance Manhattan ( $L^1$ ). Dans ces expériences, nous faisons varier respectivement le nombre d'images par classe pris pour l'arrière-plan  $n_{tr}$  et celui pris pour le test  $n_{ts}$ . Les graphiques de la Figure V.9 montrent les résultats obtenus par les bases de données ORL, Yale et AR avec différentes valeurs du couple  $(n_{tr}, n_{ts})$ .

- La base de données ORL est subdivisée en deux ensembles, d'arrière-plan et de test, dont les tailles dépendent des valeurs du couple  $(n_{tr}, n_{ts})$  définie  $\{ \forall i/1 \leq i \leq 9 ; n_{tr}=i \text{ et } n_{ts}=10-i \}$ . Les résultats illustrés dans la Figure V.9-(a) montrent que le choix de la taille blocs  $14 \times 14$  et de la distance Manhattan ( $L^1$ ) permet d'atteindre un bon taux de reconnaissance qui augmente pour des valeurs de  $n_{tr}$  allant de 1 à 9.
- Yale est subdivisée tel que :  $\{ \forall i/1 \leq i \leq 10 ; n_{tr}=i \text{ et } n_{ts}=11-i \}$ . Les résultats illustrés dans la Figure V.9-(b) montrent que le choix de la même taille blocs  $14 \times 14$  et de la métrique ( $L^1$ ) permet d'atteindre un bon taux de reconnaissance qui accroît avec  $n_{tr}$  en allant de 1 à 10.
- Quant à AR, elle est subdivisée tel que :  $\{ \forall i/1 \leq i \leq 13 ; n_{tr}=i \text{ et } n_{ts}=14-i \}$ . Les résultats illustrés dans la Figure V.9-(c) montrent que le choix de la taille blocs  $14 \times 14$  et de la

distance Manhattan ( $L^1$ ) permet d'atteindre un bon taux de reconnaissance qui augmente avec la valeur de  $n_{tr}$  allons de 1 à 13. En conclusion des résultats précédents, chaque base de données est caractérisée par le triplet (taille bloc, métrique,  $n_{tr}$ ) qui donne des résultats optimaux. Nous prenons pour la suite les triplets :  $(14 \times 14, L^1, 5)$  pour ORL,  $(14 \times 14, L^1, 6)$  pour Yale, et  $(14 \times 14, L^1, 7)$  pour AR.

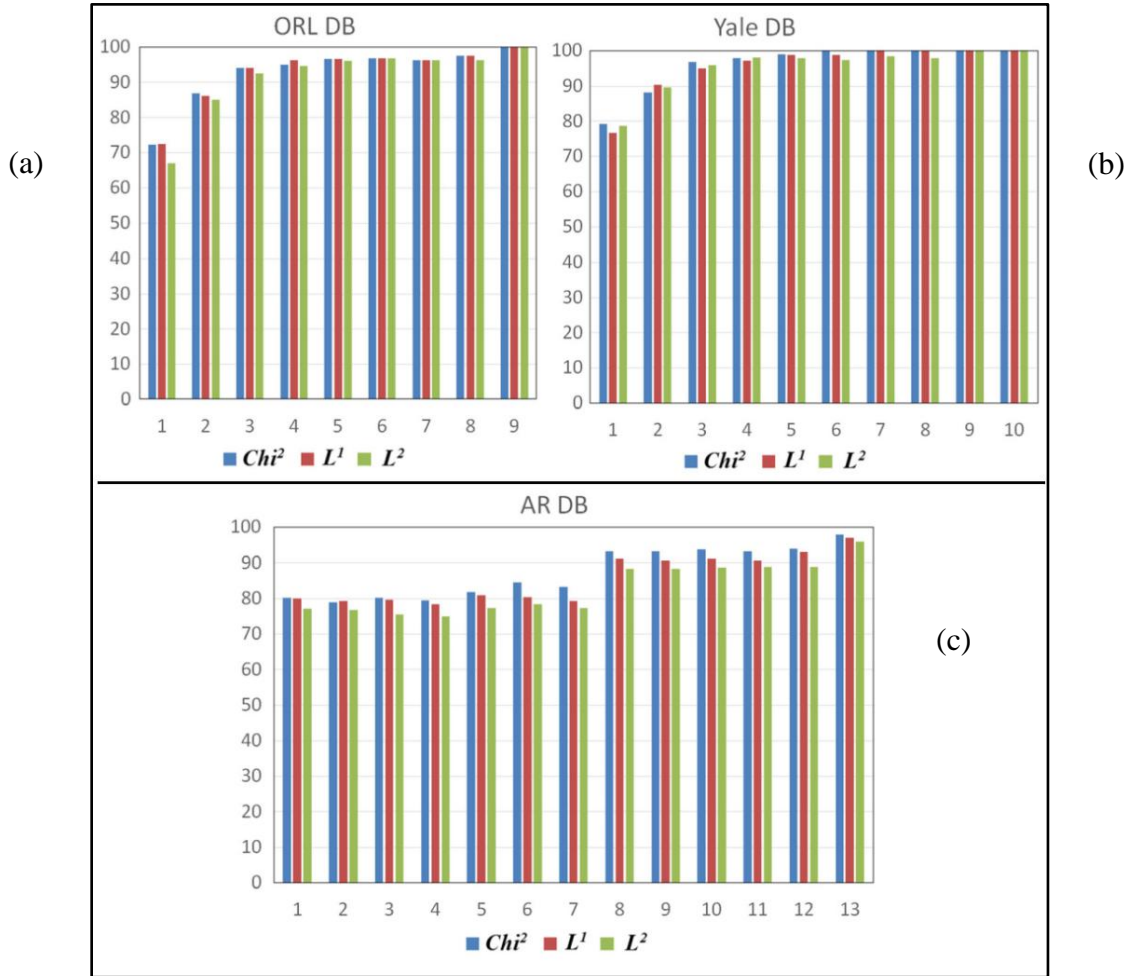


Figure V.9 : Taux de reconnaissance de différents couples  $(n_{tr}, n_{ts})$

### III.2.4. Evaluation par apprentissage artificiel

Dans cette sous-section, nous réalisons des expériences qui prennent en compte les deux premiers arguments de chaque triplet. Pour chaque base de données, nous appliquons l'algorithme (KNN) en mode de validation croisée de  $k$ -folders ( $kf=4, 6, 8, 10$ ) pour évaluer la classification de l'ensemble de données composé de vecteurs descripteurs eLBPH de taille  $L=$

$8 \times 7 \times 16 = 896$ . Le Tableau V-6 illustre les résultats obtenus pour les trois bases de données ORL, Yale et AR. Ces résultats représentent les valeurs des paramètres suivants : taux des vrais positifs (TP), taux des faux positifs (FP), précision, rappel, F-mesure et l'exactitude. Ces paramètres sont des outils de mesure de performance de différents classificateurs à apprentissages.

Tableau V-6 : Résultats de classification par KNN ( $L^1$ )

BD	TP	FP	Précision	Rappel	F-mesure	Exactitude
ORL	0.963	0.001	0.971	0.963	0.964	96.32 %
Yale	0.976	0.002	0.977	0.976	0.976	97.57%
AR	0.957	0.001	0.960	0.957	0.957	95.71%

Pour ORL, les valeurs prises par les différents paramètres sont significatives, spécialement celles de la F-mesure et de l'exactitude. Ces derniers atteignent respectivement les valeurs 96.40% et 96.32%. De même, pour Yale, les valeurs des paramètres sont elles aussi significatives. La F-mesure et l'exactitude atteignent respectivement les valeurs 97.60% et 97.57%. Quant à AR, ces valeurs sont elles aussi significatives. La F-mesure et l'exactitude atteignent respectivement les valeurs 95.70% et 95.71%.

### III.2.5. EVBP et méthodes de réduction de la dimension

Dans cette sous-section, nous comparons les performances des méthodes de réduction de la dimension usuelles lorsqu'elles sont appliquées respectivement sur les trois bases de données (ORL, Yale et AR) et leurs transformées EVBP. Le Tableau V-7 récapitule les résultats de la réduction de la dimension des méthodes PCA, LDA, 2DPCA et 2DLDA. La performance de chaque méthode est mesurée en termes de taux de reconnaissance et taille de la base propre de projection (termes entre parenthèses dans le Tableau V-7). En effet, ces résultats montrent qu'EVBP offre des avantages supplémentaires aux méthodes PCA, LDA, 2DPCA et 2DLDA. Particulièrement avec les deux bases de données Yale et AR.



Tableau V-7 : Combinaison d'EVBP et les méthodes de réduction de la dimension

Bases	ORL ( $n_{tr}=5$ , $n_{ts}=5$ )		Yale ( $n_{tr}=5$ , $n_{ts}=6$ )		AR ( $n_{tr}=7$ , $n_{ts}=7$ )	
	EVBP	Originale	EVBP	Originale	EVBP	Originale
PCA	91(17)	92.5(17)	97.78(14)	84.44(14)	95(40)	77.53(40)
LDA	91.5(39)	91(39)	98.89(14)	94.44(14)	95.5(99)	94.5(99)
2DPCA	96(6)	95(6)	98.89(4)	88.89(10)	96.86(6)	93.43(10)
2DLDA	96.5(8)	96.5(8)	98.89(4)	95.56(6)	98.86(6)	96.43(10)

- ORL est subdivisée en deux ensembles tel que :  $n_{tr}=5$  et  $n_{ts}=5$ , les résultats obtenus avec la base de données formée par les images d'origine et celle formée par les images EVBP sont comparables. Ainsi, les meilleurs d'entre eux sont obtenus par 2DPCA et 2DLDA et atteignent respectivement les taux 96.00% et 96.50%.
- Yale est subdivisée en deux ensembles avec :  $n_{tr}=5$  et  $n_{ts}=6$ , les résultats obtenus avec la base de données formée par les images EVBP sont meilleurs que ceux obtenus avec celle formée par les images d'origine. Ainsi, les meilleurs d'entre eux sont obtenus par 2DPCA et 2DLDA qui ont atteint le même taux 98.89% et avec la même dimension 4. En effet, EVBP a amélioré le taux de reconnaissance de 10.11% pour 2DPCA et 3.36% pour 2DLDA.
- De même, AR est subdivisée en deux ensembles avec :  $n_{tr}=7$  et  $n_{ts}=7$ , les résultats obtenus avec la base de données formée par les images EVBP sont eux aussi meilleurs que ceux obtenus avec celle formée par les images d'origine. Ainsi, les meilleurs d'entre eux sont aussi obtenus par 2DPCA et 2DLDA qui ont atteint respectivement les taux 96.86% et 98.86% avec la même dimension 6. En effet, EVBP a amélioré le taux de reconnaissance de 3.54% pour 2DPCA et 2.46% pour 2DLDA.

### III.3. Comparaison avec les méthodes de l'état de l'art

Finalement, nous comparons les techniques proposées avec des méthodes récentes de l'état de l'art. Le Tableau V-8 résume les différents résultats obtenus pour les bases de données ORL, Yale et AR.

Tableau V-8 : Comparaison avec des méthodes de l'état de l'art

ORL	Méthodes	LBP	DWT+ PCA	DWT+ LDA	PCA+ 2DPCA	LDA+ 2DPCA	HSF	EVBP		
								eLBPH	2DPCA	2DLDA
	Taux(%)	91	94.2	97.1	92.1	92.1	98	96.5	96.0	96.5
Réfs	[147]	[149]	[149]	[150]	[150]	[162]	Notre approche			
Yale	Méthodes	Gabor+ LBP+LPQ	LPP	Gabor+ mod2DPCA	B(2D) 2PCA	ANMM	EVBP			
							eLBPH	2DPCA	2DLDA	
	Taux(%)	90.7	93.3	90.4	92.7	93.73	98.89	98.89	98.89	
Réfs	[163]	[164]	[81]	[154]	[164]	Notre approche				
AR	Méthodes	ULBP+ LPQ	SCR	LLE + SVM	PCA+ SVM	2DPCA	EVBP			
							eLBPH	2DPCA	2DLDA	
	Taux(%)	92.7	97.5	95	92.67	96.1	97.15	96.86	98.86	
Réfs	[146]	[165]	[159]	[145]	[166]	Notre approche				

En effet, pour ORL, les résultats obtenus par la fusion de EVBP et les trois méthodes d'extraction de caractéristiques (eLBPH, 2DPCA et 2DLDA) montrent la compétitivité de la représentation faciale EVBP. Les taux atteints sont élevés que ceux rapportés dans [147,150] et comparables à ceux rapportés dans [149,162]. De même, pour Yale la compétitivité de telle représentation faciale par EVBP est largement justifiée. En effet, les taux atteints par la même fusion sont largement supérieurs à ceux rapportés dans [81, 154, 163-164]. Concernons AR, la fusion de EVBP avec les deux méthodes (eLBPH et 2DPCA) atteint des taux supérieurs à ceux rapportés dans [145-146] et reste compétitive comparée aux méthodes évoquées dans [159, 165-166].

## **IV. Conclusion**

Dans ce chapitre, nous avons présenté une méthode d'extraction de caractéristiques basée sur l'intensité: le modèle binaire virtuel électrique (EVBP). Ce dernier engendre une nouvelle description de l'image, particulièrement celle du visage, en se basant sur les deux concepts: dualité « pixel-charge électrique virtuelle » et principe de balayage évoqué par Poincaré. En effet, EVBP interprète la forte variation de gradient au niveau des caractéristiques pertinente du visage par une dissymétrie des interactions entre charges électriques virtuelles des grilles correspondantes. Cette dissymétrie est due à l'irrégularité de la distribution de l'énergie projetée, par réflexion-diffuse, de l'objet réel vers le capteur.

Afin d'étudier l'impact d'EVBP sur le processus de reconnaissance, nous avons mené des expériences approfondies sur trois bases de données, ORL, Yale, et AR, nous avons souligné qu'EVBP est plus efficace que LBP. Il ressort des résultats obtenus que la description faciale par EVBP a atteint d'excellentes performances en termes de haute précision et exécution rapide. Ceci est dû à la forte discrimination et à la taille réduite des vecteurs descripteurs générés par EVBP. Il est important de noter que cette description a permis d'utiliser une métrique de mesure de similarité,  $L^1$ , à la place de la métrique,  $\text{Chi}^2$ , appropriée par différentes variétés de LBP, ce qui a réduit d'avantage le temps d'exécution de notre système. Aussi, la comparaison de l'approche proposée avec les méthodes récentes de l'état de l'art a montré sa compétitivité et sa prépondérance.

Dans le chapitre suivant, nous nous intéresserons à une autre description locale qui prend en compte la redondance de l'information au sein de chaque voisinage et interprète cette redondance comme un problème d'optimisation primale-duale.

---

---

## **Chapitre VI. Modèle Optimal Binaire Local (Optimal Local Binary Pattern, Optim-LBP) : Application à la Biométrie Faciale**

---

---

Les phénomènes de notre environnement peuvent être des problèmes de programmation non linéaire dont l'explication fait appel à des concepts d'optimisation. L'objectif principal de tels concepts réside dans la réalisation de solution optimale avec un minimum d'énergie. Dans ce chapitre, nous formulons la description locale au sein d'une image comme un problème de programmation quadratique séquentielle qui vise à déterminer les poids optimaux de contribution des pixels voisins dans la construction de chaque pixel courant. Dans ce cadre, nous définissons notre fonction objectif en se basant sur le degré d'homogénéité mesuré par l'entropie de Shannon.

## **I. Introduction**

La représentation de l'image en termes de pixels peut être vue comme un problème statistique dont les variables sont les liens reliant les pixels aux seins de chaque voisinage. À l'hypothèse que les éléments de chaque voisinage,  $3 \times 3$ , ont un comportement quasi-similaire vis-à-vis les différentes transformations, géométriques et d'illumination, usuelles, nous formulons le problème d'encodage de l'image comme un problème d'optimisation où chaque pixel courant dépend implicitement des contributions de ses plus proches voisins. Dans ce contexte, nous proposons un nouveau descripteur basé sur l'apparence locale appelé modèle d'optimisation binaire local (Local Binary Optimization Pattern, Optim-LBP) qui génère une nouvelle description de l'image, en particulier celle du visage. Afin de définir la fonction objectif régissant notre distribution de l'information sur les éléments d'un voisinage, notre méthode, intitulée Optim-LBP, mesure le degré d'homogénéité de chaque voisinage en termes de la valeur et de la redondance locale de chacun de ses pixels en utilisant l'entropie de Shannon. Notre méthode porte sur une famille de modèles dont chaque variable représente le degré de participation d'un pixel voisin dans la constitution du pixel courant. Le choix optimal de ces variables est réalisé en assurant une perte minimale de l'information après substitution de l'intensité exacte par la valeur approchée.

Plusieurs méthodes d'optimisation ont été proposées et regroupées en deux classes. Les méthodes de la première classe sont caractérisées par des exigences sur la régularité de la fonction objectif et de ses contraintes, ainsi que des exigences sur la hessienne. Celles de la deuxième classe n'exigent aucune régularité des composants du problème étudié. Deux courants marquent la première classe, l'approche primale et l'approche duale. Parmi ces méthodes, on trouve la descente du gradient et ses variétés, les méthodes de pénalité, les lagrangiennes et les lagrangiennes augmentées [167]. Les méta-heuristiques, elles, sont des techniques inspirées des phénomènes naturels consistant à chercher un optimum local acceptable tout en agissant sur des populations de solutions initialisés arbitrairement. Parmi ces méthodes, on trouve la recuit simulé [168], la recherche tabou [169], les algorithmes génétiques [170], la colonie des fourmilles [171] et les Essaims Particulaires [172]. Enfin, et en se basant sur la résolution du problème d'optimisations proposé, Optim-LBP engendre une

nouvelle description faciale où chaque pixel courant est codé en fonction des poids de contribution de ses voisins.

## II. Méthodes d'optimisation

### II.1. Optimisation d'une fonction régulière

#### II.1.1. Optimisation sans contraintes

Les algorithmes d'optimisation sans contraintes cherchent à résoudre les problèmes de minimisation des fonctions régulières  $f$  en déterminant les points qui les minimisent et sans aucune condition préalable sur les relations entre les éléments de l'espace des réalisations.

$$\min_{x \in \mathbb{R}^n} f(x) \quad (\text{VI.1})$$

- Les méthodes de descente: sont les techniques plus connues de l'approche primale, notamment, la descente du gradient et ses variétés qui ne diffèrent que par leurs façons de choisir le pas de déplacement et la direction de descente; ces méthodes sont adoptées par la suite aux problèmes avec contraintes en termes d'un bouqué de méthodes connues sous le nom de directions réalisables [165]. Dans la majorité de ces méthodes, la direction de descente est prise comme le vecteur opposé du gradient et le pas de la descente de la direction courante est déterminé par l'algorithme de rebroussement suivant :

**Données :** un point  $x$ , une direction de descente  $d$  et deux réels  $\alpha$  et  $\beta$ .

**Résultat :** le pas de descente  $t > 0$

$t = 1$  ;

Tant que  $f(x + td) > f(x) + \alpha \langle \nabla f(x), d \rangle$  faire :

Réduire  $t$  tel que  $t = \beta t$  et  $0 < \beta < 1$

fin tant que.

- Méthode de descente du gradient : elle consiste à suivre itérativement la direction de descente jusqu'à atteindre un minimum local. Les étapes de la version naïve de cette méthode est telle que :

**Données :** un point initial  $x_0$ , un seuil de tolérance  $\varepsilon > 0$ .

**Résultat :** un point  $x$  proche du minimum  $x^*$

$k = 0$  ;  $x = x_0$  ;

Tant que  $\|\nabla f(x)\| > \varepsilon$  faire :

Calculer la direction :  $d_k = -\nabla f(x_k)$

Déterminer le pas de descente  $t_k > 0$  par la méthode de rebroussement

$$x_{k+1} = x_k + t_k \cdot d_k$$

$k = k+1$  ;

fin tant que.

- Méthode de Newton : c'est un algorithme de descente qui consiste à minimiser itérativement le développement de Taylor au second ordre de la fonction  $f$ .

**Données :** un point initial  $x_0$ , un seuil de tolérance  $\varepsilon > 0$ .

**Résultat :** un point  $x$  proche du minimum  $x^*$

$k = 0$  ;  $x = x_0$  ;

Calculer le premier pas de descente :

$$d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

$$C_k = -\langle d_k, \nabla f(x_k) \rangle$$

Tant que  $C_k > \varepsilon^2$  faire :

Déterminer le pas de descente  $t_k > 0$  par la méthode de rebroussement

$$x_{k+1} = x_k + t_k d_k$$

$k = k+1$  ;

Calculer le nouveau pas de descente :

$$d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k) \text{ et } C_k = -\langle d_k, \nabla f(x_k) \rangle$$

fin tant que.

Plusieurs autres algorithmes d'optimisation sans contraintes ont été proposés notamment la méthode du gradient à pas optimal, la méthode du gradient conjugué, la méthode de Davidon Fletcher Powe et la méthode de relaxation [167].

### II.1.2. Optimisation avec contraintes

Les méthodes d'optimisation avec contraintes cherchent à minimiser la fonction objectif  $f$  sous réserve de certaines relations logiques entre ses différentes variables. Les méthodes de pénalisation et les méthodes de programmation quadratique séquentielle sont les plus connues de cette classe.

$$\begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ C_{eg}(x) = 0 \\ C_{in}(x) \leq 0 \end{cases} \quad (VI.2)$$

- Méthodes de pénalisation : elles consistent à transformer le problème d'optimisation avec contraintes en un problème d'optimisation sans contrainte en injectant une fonction de pénalisation de contraintes  $p$ .

$$\min_{x \in \mathbb{R}^n} (f(x) + \sigma p(x)) \quad \text{avec} \quad \sigma > 0 \quad (VI.3)$$

Parmi méthodes de pénalisation, nous distinguons les méthodes de pénalisation extérieure et celles de pénalisation intérieure. Dans les méthodes de pénalisation extérieure, l'optimum  $x^*$  est approché vers l'extérieur en pénalisant la fonction  $f$  à l'extérieur de l'ensemble admissible. Dans les méthodes de pénalisation intérieure l'optimum  $x^*$  est approché d'une manière itérative vers l'intérieur de l'ensemble admissible [167].

- Méthode de programmation quadratique séquentielle : cette technique est utilisée dans le cadre des problèmes d'optimisation non linéaire sous contraintes dont la forme est :

$$\begin{cases} \min_{x \in \mathbb{R}^n} f(x) & \text{sous} \\ g_i(x) = 0 & \text{pour } i = 1, \dots, n \\ h_j(x) \leq 0 & \text{pour } j = n + 1, \dots, m \\ x^{inf} \leq x \leq x^{sup} \end{cases} \quad (VI.4)$$



C'est une technique primale-duale qui transforme le problème étudié en une série de problèmes de minimisation locale, tout en se basant sur les quantificateurs de Lagrange [167].

## II.2. Les méta-heuristiques

Les méta-heuristiques sont des méthodes d'optimisation qui cherchent le point le plus proche du minimum d'une fonction parmi un nombre fini de points candidats. Deux grandes classes des méta-heuristiques sont distinguées, notamment les méta-heuristiques à solution unique (Mono-attracteur) et les méta-heuristiques à population (Multi-attracteurs).

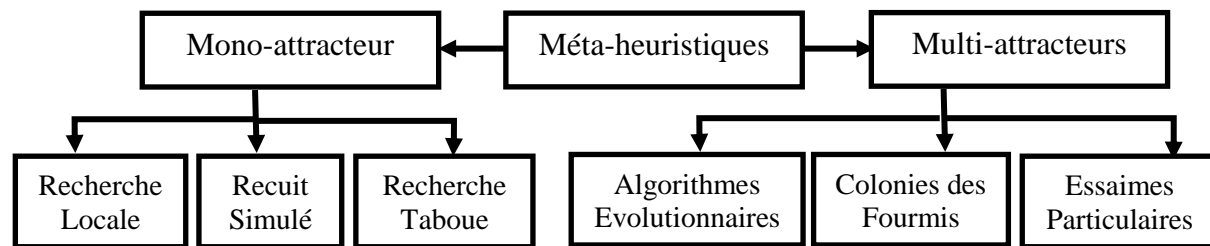


Figure VI.1 : Les méta-heuristiques les plus connues pour l'optimisation

- Recuit Simulé : cette ancienne méta-heuristique permet de résoudre le problème de minimum local en s'appuyant sur l'algorithme de Metropolis-Hastings qui décrit l'évolution d'un système thermodynamique par la statistique de Boltzmann [168]. En outre, l'évolution d'un système vers son état d'équilibre est simulée suivant la probabilité d'acceptation de chaque transformation élémentaire subie par le système. Chaque transformation entraînant une diminution de l'énergie sera définitivement admise ; elle sera acceptée, avec une probabilité entre 0 et 1, dans le cas contraire,.
- Recherche Taboue : c'est une technique de recherche locale qui se distingue par son recours à un historique des solutions récemment visitées ou taboues. Cela permet de ne pas rester piégé dans un minimum local ou d'y retourner trop rapidement. Dans cette technique, la taille de la liste taboue est une donnée primordiale empiriquement déterminée [171].
- Essaim de particules : l'optimisation par essaim de particules est une technique évolutionnaire qui développe une solution optimale du problème, en utilisant une population de solutions candidates. Chaque particule est caractérisée, lors de son évolution

dans l'espace de recherche, par sa vitesse  $v(t)$  et sa position  $x(t)$  [172]. Afin de régler ces deux paramètres chaque particule doit obéir aux règles suivantes:

- Chaque individu doit prendre une décision.
- Chaque individu doit être avéré de sa propre position locale et des informations sur les individus de son voisinage.
- Chaque individu doit rester proche des autres individus, aller dans une même direction, et/ou voler à la même vitesse.

Après la présentation des principaux algorithmes d'optimisation, nous passons à la conception et à l'étude de notre descripteur local traitant les pixels d'une image comme des réalisations, et leur distribution comme un problème de programmation quadratique séquentielle formulé comme un problème dual d'optimisation.

### **III. Modèle optimal binaire local (Optim-LBP)**

Dans cette section, nous proposons notre modèle optimal binaire local (Optimal Local Binary Pattern, Optim-LBP) qui consiste à générer une nouvelle description de l'image, en particulier celle du visage en faisant appel à la minimisation sous contraintes d'une fonction objectif qui détermine les contributions optimales des pixels de chaque voisinage.

#### **III.1. Modélisation du descripteur Optim-LBP**

##### **III.1.1. Présentation**

Dans leurs descriptions d'une image, LBP d'origine et certains de ses variétés utilisent des fonctions de décision déterministes qui entraînent des pertes d'informations. Ces pertes sont généralement liées à l'apparence locale de l'image. En effet, LBP ne peut remplir dans la plupart des cas la condition de sélection qui préserve l'apparence locale de chaque voisinage. Il peut engendrer le même modèle LBP pour des voisinages assez différents. La Figure VI.2 illustre le cas de deux voisinages représentant la frontière entre des zones différentes et une zone homogène. Malgré la différence totale des apparences locales de ces deux voisinages, ils sont codés par le même modèle LBP. Afin de combler, autant que possible, les lacunes de LBP préalablement mentionnées, nous avons proposé notre descripteur Optim-LBP. L'idée

principale de ce descripteur consiste à générer une nouvelle description de l'image, en particulier celle du visage, en faisant appel à un problème de minimisation sous contraintes d'une fonction objectif définie à partir de la composition de chaque voisinage courant. Ainsi, lors de recodage du pixel courant, ce problème de minimisation est élaboré dans le but de déterminer les poids optimaux des contributions de tous les pixels du voisinage courant. Cela est réalisé par la génération et l'étude de la transition du voisinage courant de son état actuel et un état transitoire où la valeur approchée de chaque pixel est déduite de sa valeur actuelle à un facteur près.

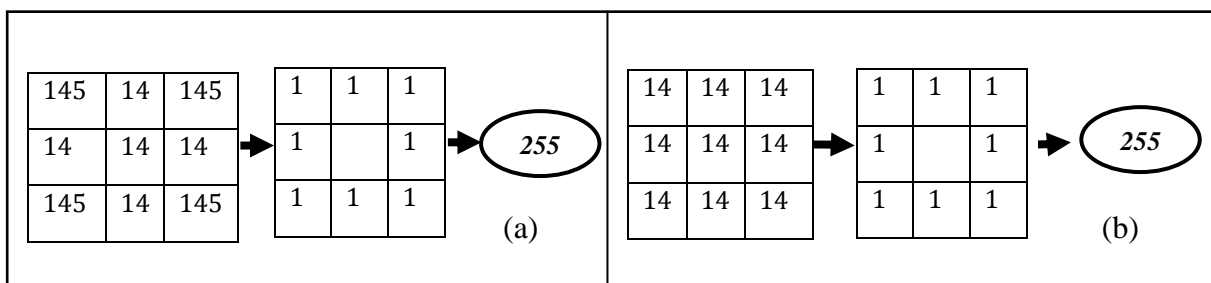


Figure VI.2 : Effet de l'aspect binaire de la fonction de décision

### III.1.2. Formulation de Optim-LBP

Avant de passer à la formulation du descripteur Optim-LBP, nous définissons les paramètres suivants :

- Variables : supposons le voisinage,  $N_c = \{g_0 \dots g_{p-1}, g_c\}$ , centré sur le pixel courant  $c$ , et qui est caractérisé dans son état initial,  $E_i$ , par l'entropie  $S_c$ . Les éléments de ce voisinage vont transiter vers l'état final,  $E_f$ , caractérisé par l'entropie  $S'_c$ , pour engendrer le voisinage,  $N'_c = \{g'_0 \dots g'_{p-1}, g'_c\}$ . La relation qui existe entre les deux états,  $E_i$  et  $E_f$ , est définie par les poids des contributions des pixels voisins  $x_i \in [0, 1]$  tel que  $g'_i = x_i \cdot g_i$ . Ces derniers sont utilisés pour définir la fonction objectif à minimiser.
- Fonction objectif: la détermination du vecteur,  $x = (x_0 \dots x_{p-1}, 1)$ , représentant les poids optimaux des contributions des pixels voisins nécessite la définition d'un critère d'optimisation qui modifie l'information locale de chaque pixel et conserve l'information globale du voisinage lors du passage d'un état à un autre. Cela revient à définir une fonction objectif qui mesure le degré d'homogénéité d'un voisinage en utilisant l'erreur

quadratique entre l'entropie initiale  $S_c$  liée au voisinage  $N_c$  et l'entropie finale  $S'_c$  liée au voisinage  $N'_c$ .

$$S_{err} = (S_c - S'_c)^2 \quad (VI.5)$$

Dans son travail de base, Shannon [173] propose de mesurer la quantité d'information d'un système en utilisant une nouvelle expression de l'entropie  $S$  :

$$S = \sum_i -p_i \log p_i \quad (VI.6)$$

Où,  $p_i$  est la probabilité de redondance de l'élément  $i$  du système.

- Modèle : pour déterminer la contribution optimale de chaque pixel voisin, nous calculons l'entropie initiale  $S_c$  et l'entropie finale  $S'_c$ .

$$S_c = \sum_i -p_i \log p_i \quad (VI.7)$$

$$S'_c = \sum_i -p'_i \log p'_i = \sum_i -x_i p_i \log x_i p_i \quad (VI.8)$$

Le vecteur optimal  $x$  composé de contributions de pixels voisins est une solution du modèle d'optimisation suivant :

$$\begin{cases} \text{Min } S_{err}(x) & \text{sous contraintes,} \\ \sum_i x_i \leq P - 1 \\ x \in [0; 1]^{P-1} \end{cases} \quad (VI.9)$$

La fonction objectif du modèle proposé est définie comme suit :

$$S_{err} = \left( \sum_{i=0}^{P-1} (1 - x_i) p_i \log(1 - x_i) p_i \right)^2 \quad (VI.10)$$

Par conséquent, la minimisation de cette fonction suivant le modèle sous contraintes proposé génère les valeurs optimales des poids,  $x_i$ . Ces poids sont utilisés par la suite pour calculer les modèles Optim-LBP par l'équation suivante :

$$\text{Optim-LBP}_{P,R} = \text{floor} \left( \sum_{j=0}^{P-1} 2^j x_j \right) \quad (VI.11)$$

Où,  $P$  représente le nombre de pixels voisins et  $x = \{x_j, j=0, \dots, P-1\}$  le vecteur des solutions optimales de l'équation (VII.7).

---

### **III.1.3. Analyse et discussion**

Généralement, chaque image peut être considérée comme association de quatre composantes principales. La première composante est formée de l'arrière-plan représentant la scène, la seconde est composée des structures ajoutées à cette scène ; la troisième composante forme l'ensemble des zones limites représentant les contours de séparation entre les structures ajoutées et la scène ; la dernière composante représente les effets de fluctuations causées par les conditions de capture de l'image. Afin de séparer au maximum ces quatre composantes, nous avons proposé notre descripteur Optim-LBP qui consiste à modéliser chaque pixel courant comme association de sa valeur réelle à un bruit aléatoire, à estimer l'importance de ce bruit à travers une fonction d'évaluation et à redistribuer ce bruit aléatoire sur tous les pixels du voisinage courant. Pour atteindre cette exigence, nous avons fait évoluer le système, voisinage courant, d'un état initial  $E_i$  caractérisé par l'entropie  $S_c$  à un état final  $E_f$  caractérisé par l'entropie  $S'_c$ . Nous avons ensuite minimisé l'erreur quadratique entre ces valeurs d'entropie. Ce processus permet de calculer les poids optimaux des contributions des pixels voisins en de leurs valeurs et leurs redondances, et d'utiliser en suite ces contributions dans la reformulation d'Optim-LBP. Le premier avantage d'Optim-LBP est sa capacité de coder chaque pixel courant en fonction de la redondance de ses voisins ; le deuxième est lié à son comportement vis-à-vis des changements usuels affectant l'image.

En effet, les deux exemples de la Figures VII-3-a montrent que malgré la différence des valeurs des pixels des deux voisinages, les poids des contributions des pixels voisins sont identiques, 0.87 et les deux valeurs de niveaux de gris se valent 221. De même, les deux exemples de la Figures VII-3-b montrent que l'effet de la valeur du pixel courant sur les poids de contribution reste très limité et les valeurs de niveaux de gris engendrées sont très comparables (246 et 247). Concernant les changements globaux ou locaux d'illumination, la Figures VII-4 illustre deux exemples de changements d'illumination où les poids des contributions des pixels voisins restent du même ordre et donnent des valeurs en niveaux de gris similaires (121 et 122).

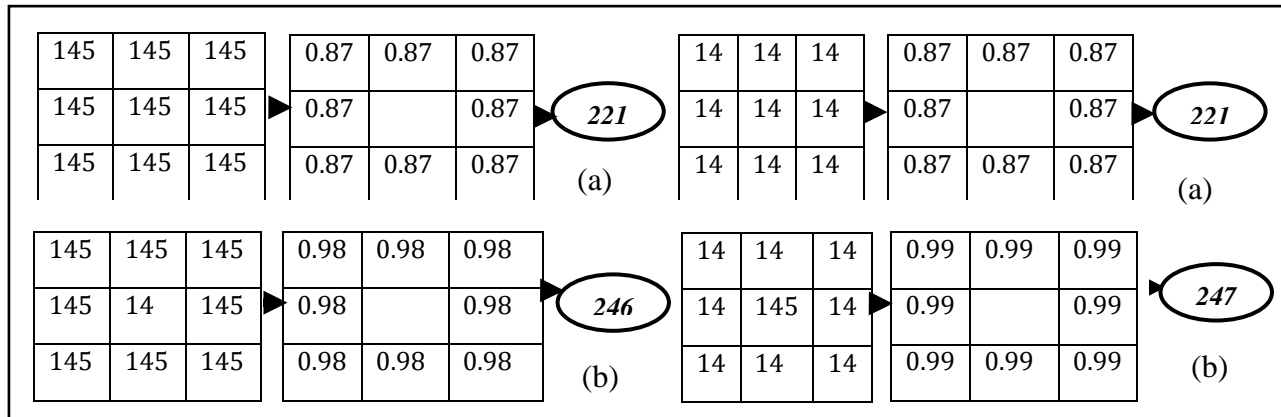


Figure VI-3 : Exemples de comportement d'Optim-LBP via différents voisinages

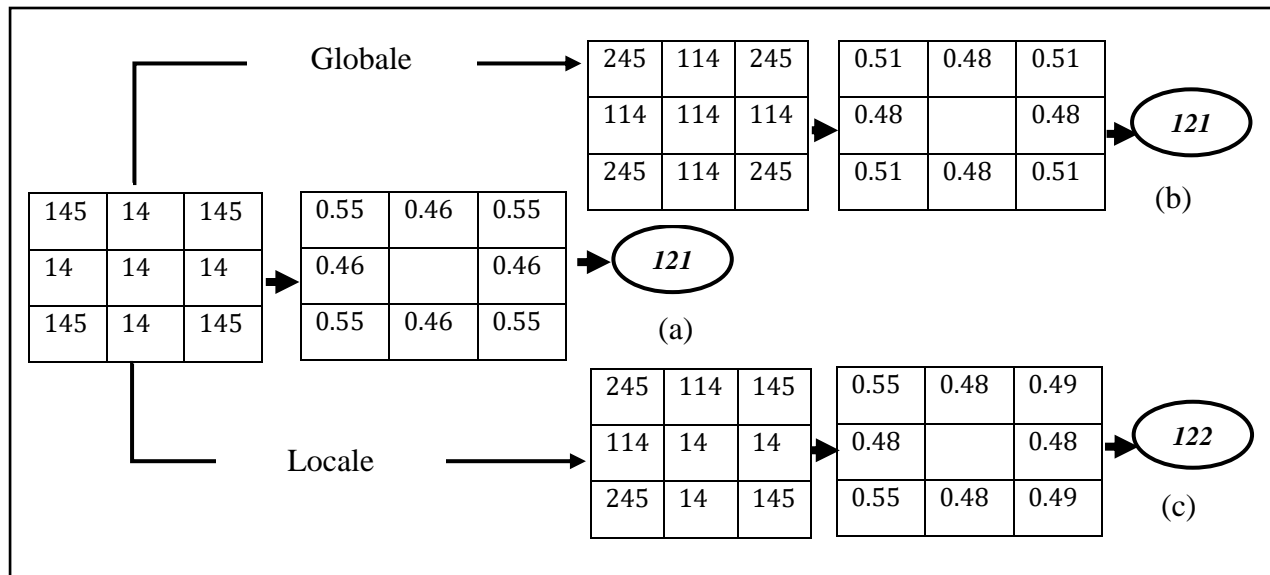


Figure VI-4 : Comportement de Optim-LBP vis-à-vis changement d'illumination

Le troisième avantage d'Optim-LBP est mis en évidence par l'analyse de la forme des histogrammes des images originales et des images obtenues par Optim-LBP. Selon la Figures VII-5, les histogrammes des images originales présentent de nombreuses fluctuations. Ils diffèrent d'une personne à une autre et pour les différentes vues d'une même personne. En revanche, les histogrammes des images Optim-LBP sont très localisés et présentent un pic au niveau de gris 127. Ce qui permet de subdiviser chaque image en deux classes correspondent respectivement aux zones quasi-homogènes et aux zones de pics. De plus la forme des histogrammes des vues différentes d'une même personne sont presque identiques et permettent

de grossièrement leurs classification. La section suivante est consacrée à l'implémentation du système de reconnaissance facile basé sur Optim-LBP. La mesure des performances de ce système est réalisée à travers une série d'expériences utilisant les bases de données ORL, Yale et AR.

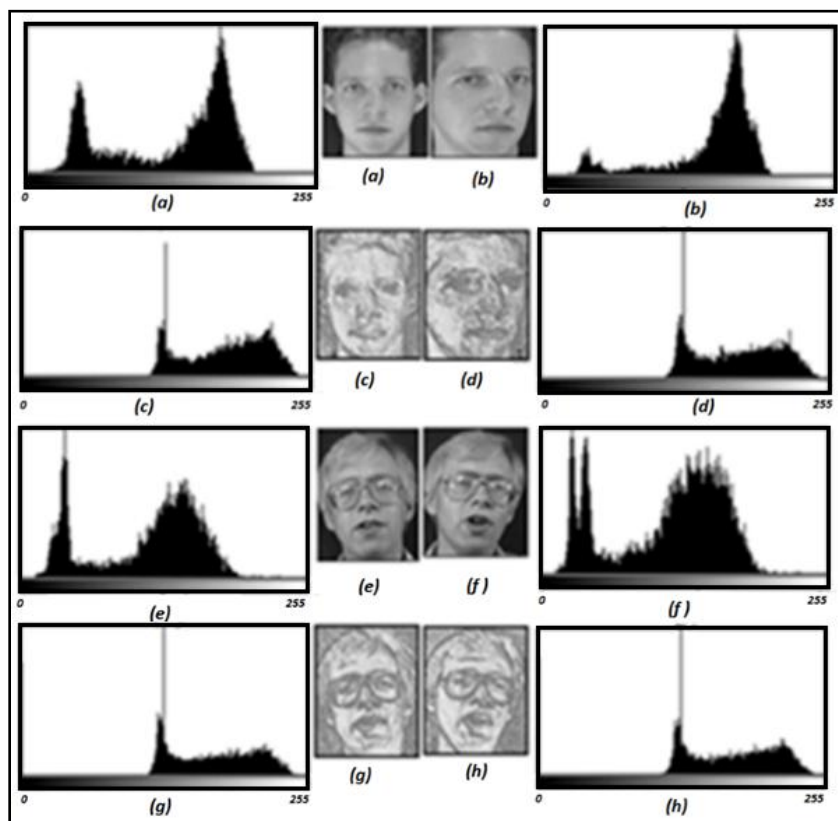


Figure VI-5: Exemples d'histogrammes d'Optim-LBP

## III.2. Reconnaissance de visage par Optim-LBP

### III.2.1. Présentation du système

Dans chaque système de reconnaissance faciale, trois opérations clés sont exécutées successivement. Dans la première opération, une nouvelle description faciale est générée; cette dernière est utilisée dans la deuxième opération pour engendrer des vecteurs descripteurs de faibles dimensions. Dans la troisième opération, l'ensemble des vecteurs descripteurs est utilisé pour choisir le classificateur le plus approprié en termes de paramètres de performances.

---

Afin d'engendrer une description faciale assez discriminante, Optim-LBP encode chaque image de visage de manière à localiser les différents composants qui marquent la topologie de la surface faciale, notamment les yeux, le nez et la bouche (Voir, Figure IV.6). Certes, la taille de ces représentations reste assez importante, mais leur réduction par la méthode 2DLDA procure des vecteurs descripteurs très discriminants et à faible dimension et de haute discrimination. Enfin, nous choisissons le classificateur approprié parmi les quatre versions de SVM suivantes: la première version est l'algorithme SVM optimisé nommée SMO ; les trois versions restantes sont des SVMs utilisant le noyau polynomial défini par:  $K(x,y) = (x^t y + c)^d$  avec  $c, d \in \mathbf{R}$ . Soient, SVM1( $c=1$  et  $d=1$ ), SVM2( $c=1$  et  $d=2$ ) et SVM3( $c=1$  et  $d=3$ ).

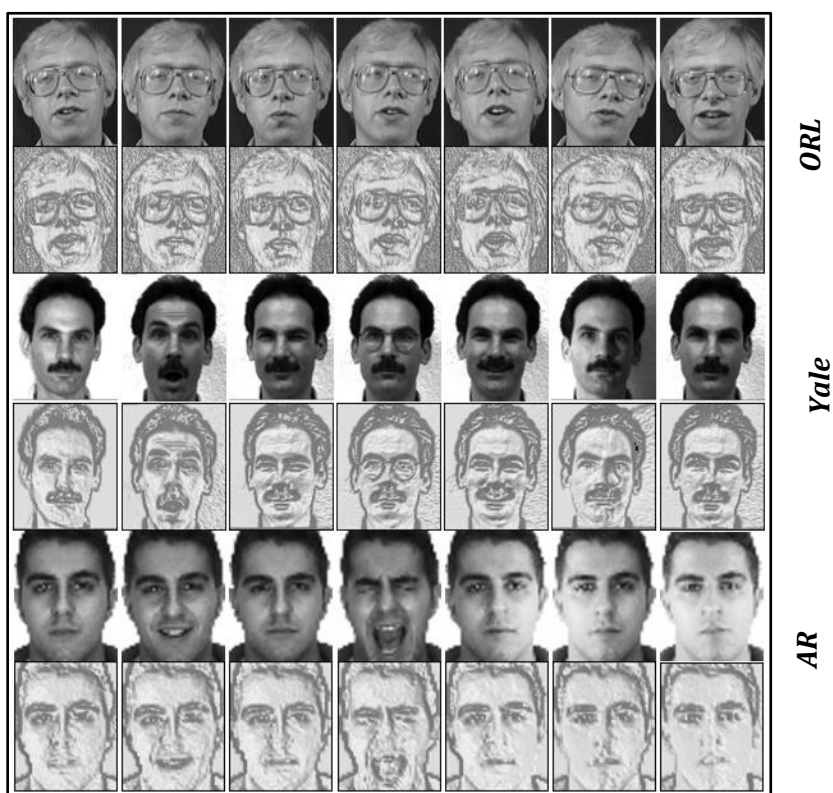


Figure VI.6 : Images Optim-LBP pour ORL, Yale et AR.

### III.2.1. Expérimentations

Dans cette section, nous mettons en évidence l'efficacité de l'approche proposée en réalisant une série d'expériences où les représentations Optim-LBP des images des trois bases de données ORL, Yale et AR sont utilisées. Dans ces expériences, nous étudions l'effet de la

---



dimension de la base de projection sur la performance du système en termes de taux de reconnaissance. Ensuite, nous étudions les performances de différents classificateurs en termes de paramètres d'évaluation de l'apprentissage artificiel, notamment, la précision, le rappel la F-mesure et l'exactitude; et nous choisissons le classificateur le plus adéquat à notre système de reconnaissance.

### III.2.1.1. Effet de la dimension de la base de projection

Dans cette section, nous mettons en œuvre l'effet de la dimension de la base de projection sur la performance du système de reconnaissance. Pour y parvenir, nous menons des expériences où nous appliquons la méthode 2DLDA sur les images des bases de données ORL, Yale et AR. Nous subdivisons aléatoirement chaque base de données en deux ensembles l'un pour l'arrière-plan l'autre pour le test. L'ensemble d'arrière-plan est formé par  $n_{tr}$  images par classe, celui du test contient les  $n_{ts}$  images par classe restantes. Ensuite, nous mesurons les taux de reconnaissance pour différentes valeurs de la dimension de la base de projection, tout en fixant  $n_{tr}$  et  $n_{ts}$ . Les résultats ainsi obtenus sont illustrés dans le Tableau VI 1. À partir de ces résultats, nous pouvons conclure que pour les trois bases de données utilisées, 2DLDA atteint de bons taux de reconnaissance avec de petites valeurs de la dimension. En effet, pour la base ORL( $n_{tr}=5$ ,  $n_{ts}=5$ ) le taux de reconnaissance atteint,  $96.38 \pm 1.3\%$ , avec la dimension ( $dim=6$ ). Pour la base Yale( $n_{tr}=5$ ,  $n_{ts}=6$ ) le taux de reconnaissance atteint,  $98.11 \pm 1.1\%$ , avec la dimension ( $dim=8$ ). Quant à AR( $n_{tr}=7$ ,  $n_{ts}=7$ ), le taux de reconnaissance,  $94.44 \pm 0.9\%$ , est atteints avec la dimension ( $dim=10$ ). Pour chaque base de données, nous avons utilisé sa dimension optimale ( $dim$ ) pour engendrer un ensemble des vecteurs descripteurs de taille ( $112 \times dim$ ). Les ensembles ainsi obtenus seront classifiés par les classificateurs SMO, SVM1, SVM2 et SVM3 évoqués au (para-III.2.1 du chapitre 2).

Tableau VI-1 : Dimensions optimales des bases de données ORL, Yale et AR

Dim	4	6	8	10	12	16	20
ORL( $n_{tr}=5, n_{ts}=5$ )	96.25 $\pm$ 1.7	96.38 $\pm$ 1.3	96.04 $\pm$ 0.9	96.04 $\pm$ 1.7	95.49 $\pm$ 1.8	94.90 $\pm$ 1.3	94.97 $\pm$ 1.8
Yale ( $n_{tr}=5, n_{ts}=6$ )	97.44 $\pm$ 1.5	97.88 $\pm$ 1.5	98.11 $\pm$ 1.1	97.66 $\pm$ 1.5	97.44 $\pm$ 1.1	97.84 $\pm$ 1.0	97.94 $\pm$ 1.1
AR ( $n_{tr}=7, n_{ts}=7$ )	93.66 $\pm$ 1.1	93.71 $\pm$ 0.9	94.00 $\pm$ 0.9	94.44 $\pm$ 0.9	94.43 $\pm$ 1.1	93.98 $\pm$ 1.1	93.73 $\pm$ 1.0

### III.2.1.2. Evaluation des classificateurs

Dans cette section, nous menons des expériences de classification à validation croisée en utilisant quatre variétés SMO, SVM1, SVM2 et SVM3. En outre, l'évaluation de la performance de chacun des classificateurs est faite sur la base de différents paramètres d'apprentissage déduits de la matrice de confusion de chaque classificateur [119-120]. Les résultats obtenus sont enregistrés dans les tableaux VI-2, VI-3 et VI-4. En effet, les résultats du Tableau VI-2 montrent que pour ORL, les quatre classificateurs ont atteint de bonnes performances avec une faible supériorité de SVM1 et SVM2. Ces dernières ont enregistré une F-mesure de 0.985 et une exactitude de 98.5%. Ceux du Tableau VI-3 montrent que pour Yale, ces classificateurs ont fini à des valeurs élevées avec une faible supériorité de SMO et SVM1 qui ont atteints une F-mesure de 0.994 et une exactitude de 99.39%. Enfin, les résultats du Tableau VI-4 montrent que pour AR, l'efficacité des quatre classificateurs a atteint respectivement une F-mesure de 0.96 et une l'exactitude de l'ordre de 96.1%. A partir des résultats précédemment obtenus, nous concluons que la description faciale par Optim-LBP engendre des vecteurs descripteurs très discriminants favorisant une bonne réduction de la dimension par 2DLDA et une meilleure classification par les SVMs.

Tableau VI-2 : Classification artificielle de la base de données ORL

Classificateurs	Précision	Rappel	F-mesure	Exactitude
SMO	0.982	0.98	0.98	98.25
SVM1	0.986	0.985	0.985	98.5
SVM2	0.986	0.985	0.985	98.5
SVM3	0.984	0.983	0.982	98.25

Tableau VI-3 : Classification artificielle de la base de données Yale

Classificateurs	Précision	Rappel	F-mesure	Exactitude
SMO	0.994	0.994	0.994	99.39
SVM1	0.994	0.994	0.994	99.39
SVM2	0.988	0.988	0.988	98.78
SVM3	0.982	0.982	0.982	98.18

Tableau VI-4 : Classification artificielle de la base de données AR

Classificateurs	Précision	Rappel	F-mesure	Exactitude
SMO	0.964	0.961	0.96	96.25
SVM1	0.964	0.961	0.96	96.1
SVM2	0.964	0.961	0.96	96.1
SVM3	0.964	0.961	0.96	96.1

### III.2.2. Comparaison avec les travaux antérieurs

La comparaison l'approche proposée aux travaux antérieurs étroitement liés cette dernière sera basée les valeurs de l'exactitude atteints dans l'expérimentation. Les résultats de cette comparaison sont illustrés dans Tableau VI-5. En effet, ces résultats montrent que les performances de l'approche proposée sont significatives et les taux atteints avec les bases ORL, Yale et AR sont très compétitifs à ceux rapportés dans [174-178].

Tableau VI-5 : Comparaison avec des méthodes de l'état de l'art

ORL	Méthodes	LBP+ +SVM	LBP-URI +SVM	2DPCA +SVM	LBP-RI +SVM	PCA +SVM	Optim-LBP+ 2DLDA+SVM1
	Taux(%)	93.7	96.9	97.3	92.1	95.7	98.5
	Réfs	[174]	[174]	[175]	[174]	[175]	Notre approche
Yale	Méthodes	ICA +SVM	HMM-LBP	2DPCA +SVM	PCA+ R-Logistic	PCA +SVM	Optim-LBP+ 2DLDA+SVM1
	Taux(%)	99.39	99.39	92.70	99.39	99.39	99.39
	Réfs	[178]	[177]	[176]	[177]	[177]	Notre approche
AR	Méthodes	Line edge map	HMM+G- HOG	2DPCA	ICA+ SVM	PCA +SVM	Optim-LBP+ 2DLDA+SVM1
	Taux(%)	96.43	98.12	96.10	94.00	92.67	96.0
	Réfs	[178]	[179]	[178]	[178]	[178]	Notre approche

## **IV. Conclusion**

Dans ce chapitre, nous avons présenté le nouveau modèle Optimal binaire local (Optim-LBP). Ce modèle consiste à formuler le problème de recodage des pixels d'une image à partir d'un problème de minimisation sous contraintes primal-dual. Dans ce problème, la fonction objectif est définie comme erreur quadratique de l'entropie de Shannon du voisinage courant pris dans deux état différents (initial et transitoire). En effet, les valeurs des pixels sont directement utilisées dans l'état initial, alors que dans l'état transitoire, ces valeurs sont pondérées les poids de contributions des pixels voisins. La détermination du vecteur formé par ces les poids de contributions représente l'objectif de la résolution du problème de minimisation sous contraintes proposé. Ainsi, nous avons utilisé la description faciale via Optim-LBP pour concevoir un système de reconnaissance faciale où la méthode 2DLDA permet de réduire l'espace des projections et les algorithmes SVMs servent à la classification. Les expériences approfondies réalisé à l'aide des bases ORL, Yale et AR ont prouvé l'efficacité de l'approche proposée. Ainsi, les meilleurs résultats correspondent à l'algorithme SVM1 qui a atteint respectivement 98.5% avec ORL, 99.39% avec Yale et 96.0% avec AR.

## *Conclusion générale et perspectives*

---

---

Les performances d'un système biométrique, en particulier la reconnaissance faciale, dépendent principalement de trois facteurs: la précision de la reconnaissance, le besoin de sauvegarde et la complexité de calcul. Ainsi, les objectifs à atteindre sont les suivants : taux de reconnaissance élevé, temps d'exécution réduit tout en mettant moins l'accent sur les besoins en stockage. Afin d'atteindre ces objectifs, une séquence de traitement partant de la détection du visage à la prise de décision doit être réalisée. Dans cette séquence, nous devons engendrer des descriptions faciales spécifiques permettant l'extraction efficace des caractéristiques faciales, la réduction au maximum de la dimension des vecteurs descripteurs et la meilleure classification de ces derniers.

Les travaux réalisés au cours de cette thèse impliquent la conception de nouveaux descripteurs locaux visant l'amélioration du processus d'extraction des caractéristiques locales de l'image, particulièrement celle du visage. Ces descripteurs nommés respectivement LBPP, LGPP, EVBP et Optim-LBP ont été utilisés pour générer de nouvelles représentations du visage en assurant la meilleure localisation des traits marquants de la surface faciale. Certes, ces représentations visent améliorer les performances de la reconnaissance faciale en minimisant la taille des vecteurs descripteurs et en assurant un niveau élevé de discrimination. L'idée principale qui relie ces quatre descripteurs locaux est liée à la coexistence des zones quasi-homogènes et des zones de pics dans chaque image. Ainsi, la séparation entre ces deux types de zones nécessite un traitement spécifique qui prend en compte la caractérisation des principales composantes du visage par la forte courbure de leur géométrie locale. Dans ce sens, chacun de nos quatre descripteurs traite cette particularité en utilisant des concepts qui ont été inspirés du comportement local des pixels d'un voisinage de l'image. En effet, notre premier descripteur, LBPP, utilise des résultats de la théorie des probabilités notamment la loi des grands nombres et le théorème central limite pour juger chaque pixel de l'image en fonction de ses pixels voisins et de l'intervalle de confiance qui les régit. Dans le même sens, notre deuxième descripteur, LGPP, applique ce même principe de l'intervalle de confiance à des images des gradients pour améliorer de plus la localisation des zones de pics. Notre troisième descripteur, EVBP, parte d'un autre concept qui se base sur le processus de construction de l'image à partir de deux concepts la réflexion-diffuse et le principe de

---

---

balayage. La réflexion-diffuse est une interaction lumière-élément de surface réel qui transmet une certaine énergie lumineuse vers le capteur; cette énergie sera transposée en termes de pixels au niveau de l'image. En fin, notre quatrième descripteur, Optim-LBP, considère chaque voisinage de l'image comme un espace de solutions d'un problème d'optimisation primal-dual. Dans ce sens, les poids de contributions des pixels voisins dans la construction du pixel courant sont vus comme solutions du problème de minimisation sous contraintes d'une fonction objectif définie en fonction du degré d'homogénéité de chaque voisinage. Afin de mettre en évidence l'efficacité des descripteurs conçus, nous les avons mis au cœur d'un système de reconnaissance faciale où nous les avons combinés avec les méthodes de *Machine-Learning* les plus populaires, notamment 2DPCA, 2DLDA, 2DDCT, 2DDWT, SVMs et MLPs. Dans ce contexte, plusieurs combinaisons (descripteur local/méthode de réduction de la dimension/algorithmes de classification) sont proposées et testées. Les résultats atteints par application de telles combinaisons sur différentes bases de données y compris ORL, Yale et AR ont prouvé que la réunion des propriétés de toutes ces méthodes a effectivement atteint son objectif en termes de taux de reconnaissance et temps d'exécution.

Dans cette thèse, nous avons précédé nos propres contributions par deux chapitres de l'état de l'art. Dans le premier chapitre, nous avons présenté les systèmes biométriques et plus spécialement la biométrie faciale. Ensuite, nous avons décrit brièvement l'état de l'art de la reconnaissance faciale. Le deuxième chapitre était une présentation détaillée des principales méthodes qui sont utilisées dans la conception de notre système de reconnaissance faciale, notamment les méthodes d'extraction des caractéristiques, les méthodes de réduction de la dimension et les algorithmes de la classification. Nos propres contributions sont regroupées en quatre chapitres dont chacun d'entre eux présente en détaille le principe, les avantages et l'intégration de l'un de nos descripteurs dans le processus de la reconnaissance faciale.

En effet, notre premier descripteur LBPP est mis dans le troisième chapitre. Ce descripteur génère de nouvelle description de l'image où les zones présentant de fortes variations du gradient sont parfaitement localisées en particulier les principales composantes de la surface faciale. Nous avons montré que cette localisation augmente avec la taille de l'intervalle de confiance jusqu'à atteindre un état stable caractérisée par une valeur de  $k$  valant 4. De plus, nous avons montré que ce descripteur représente un outil fort pour divers applications y compris la reconnaissance faciale. Dans ce sens, nous avons intégré LBPP dans la conception

de notre système de reconnaissance où plusieurs méthodes de réduction de la dimension notamment 2DPCA, 2DLDA et 2DDCT ; et de classification y compris les SVMs et les MLPs sont utilisées. Nous avons mis en évidence l'intérêt de la description faciale par LBPP dans la réduction de la dimension et la classification via une série d'expériences réalisées sur les bases de données ORL, Yale et AR. Les résultats obtenus avec les différentes combinaisons {LBPP(k)/2DPCA, 2DLDA et 2DDCT/SVM et MLP} ont montré la supériorité de notre approche comparée à celles de l'état d'art. Les meilleurs résultats sont effectivement obtenus avec (LBPP(k=4)/2DDCT/SVM).

Dans le même cadre, nous avons consacré le quatrième chapitre à notre deuxième descripteur, LGPP. Dans ce chapitre nous avons généré une description faciale plus améliorée que la précédente en appliquons le concept de l'intervalle de confiance sur des voisinages des gradients. Nous avons conclu que cette description contribue efficacement à l'amélioration des performances du système de reconnaissance faciale. Des expériences réalisées à l'aide des trois bases de données ORL, Yale et AR ont montré que les différentes combinaisons {LGPP(k)/2DPCA, 2DLDA et 2DDWT/SVM et MLP} ont abouti à de très bons résultats comparé à l'état d'art. Elle atteint de meilleurs résultats avec {LGPP(k=4)/2DPCA, 2DLDA et 2DDWT /SVM}.

Dans le cinquième chapitre, notre troisième descripteur EVBP est mis en évidence. Il génère une nouvelle description de l'image en se basant sur la dualité « pixel - 'charge électrique virtuelle' » et le principe de balayage évoqué par Poincaré. Dans cette description la géométrie locale de chaque élément de surface réel est défini par la mise en correspondance entre chaque voisinage de l'image et une grille de charges électriques virtuelles. La formulation des interactions électriques entre ces charges virtuelles suivant les quatre directions usuelles de la grille ont conduit au rétrécissement de la plage des étiquettes d'histogrammes et donc à la réduction des vecteurs descripteurs de l'algorithme eLBPH. Les expériences réalisées à l'aide des trois bases de données ORL, Yale et AR ont montré que la combinaison {EVBP/eLBPH/L<sup>1</sup>} a abouti à de très bons résultats comparés à l'état d'art.

Enfin, notre sixième chapitre est consacré à notre descripteur Optim-LBP. Dans ce chapitre nous avons mis en évidence l'efficacité de la mesure des contributions partielles des pixels voisins dans le recodage du pixel courant. Nous avons formulé ce problème comme une minimisation sous contraintes de l'erreur quadratique de l'entropie de Shannon lorsque les

---

éléments de chaque voisinage transitent de leur état actuel (initial) vers un état transitoire (final). Les expériences réalisées à l'aide des trois bases de données ORL, Yale et AR ont montré que la combinaison {Optim-LBP/2DLDA/SVMs} a abouti à de très bons résultats comparés à l'état d'art.

Nous affirmons à ce stade que nos descripteurs sont :

- Capables d'atteindre de meilleures performances en termes de taux de reconnaissance et de temps d'exécution en les combinant avec des méthodes de *Machine-Learning*.
- Capables d'assurer de meilleures réductions de la dimension lorsqu'ils sont combinés avec les méthodes 2DPCA, 2DLDA, 2DDCT et 2DDWT.
- Capables d'assurer de meilleures discriminations des vecteurs descripteurs et ainsi d'accélérer le processus de classification.
- Capables d'engendrer des descriptions en temps réel des images issues de différents domaines où l'élimination de l'arrière plan est essentielle.

Plusieurs axes de recherches nous semblent prometteurs comme perspectives de ce travail. Ils concernent l'adaptation de nos descripteurs aux domaines suivants :

- Reconnaissance de visages 3D : la pose et l'extraction des régions d'intérêts sont les principaux défis de la reconnaissance 3D qui peuvent être surmontés par l'introduction des variétés riemanniennes et de la géométrie projective.
- Autres modalités biométriques : nombreuse sont les modalités biométriques où nos descripteurs peuvent être utilisés; particulièrement la biométrie de l'iris.
- Analyse de mouvement: dans ce type d'applications, nos descripteurs peuvent assurer l'élimination efficace de l'arrière-plan permettant ainsi un suivi efficace des objets du premier plan.

Finalement, nous envisageons d'introduire l'apprentissage en profondeur (*Deep-Learning*) dans la phase de classification afin de pouvoir l'appliquer dans les domaines précédemment évoqués.



## *Bibliographie de l'auteur*

---

---

Revue internationale :

- Dahmouni, A., K. Elmoutaouakil, and K. Satori. "A New Local Binary Probabilistic Pattern (LBPP) and Subspace methods for face recognition." *Wseas Transactions On Computers* 14 (2015): 588-597.
- Dahmouni, A., Aharrane, N., K. Elmoutaouakil, and K. Satori. "MULTI-CLASSIFIERS FACE RECOGNITION SYSTEM USING LBPP FACE REPRESENTATION." *INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING INFORMATION AND CONTROL* 13 5 (2017): 1721-1733.
- Dahmouni, A., K. El Moutaouakil, and K. Satori. "A Biometric Solution of Education Environment Using Face Recognition System." *International Journal of Imaging and Robotics™* 18.1 (2018): 73-84.
- Dahmouni, A., K. Elmoutaouakil, and K. Satori. "Face description using electric virtual binary pattern (EVBP): application to face recognition." *Multimedia Tools and Applications*. Springer. (2018): 1-19.
- Dahmouni, A., Aharrane, N., K. Elmoutaouakil, and K. Satori. "A Face Recognition Based Biometric Solution in Education." *Pattern Recognition and Image Analysis*. Springer. 28.4 (2018): 758-770.
- Aharrane, N., Dahmouni, A., K. Elmoutaouakil, and K. Satori. et al. "A robust statistical set of features for Amazigh handwritten characters." *Pattern Recognition and Image Analysis*. Springer. 27.1 (2017): 41-52.

Conférences internationales :

- Dahmouni, A., K. Elmoutaouakil, and K. Satori. "Robust face recognition using local gradient probabilistic pattern (LGPP)". In: *Proceedings of the Mediterranean Conference on Information & Communication Technologies*. Springer. 2015. 2016. p. 277-286.
- Dahmouni, A., Aharrane, N., K. Elmoutaouakil, and K. Satori. "Face recognition using local binary probabilistic pattern (LBPP) and 2D-DCT frequency decomposition".

---

In: Computer Graphics, Imaging and Visualization (CGiV). IEEE. 2016 13th International Conference on. IEEE, 2016. p. 73-77.

- Dahmouni, A., Aharrane, N., K. Elmoutaouakil, and K. Satori."A New Hybrid Face Recognition System via Local Gradient Probabilistic Pattern (LGPP) and 2D-DWT". In: Europe and MENA Cooperation Advances in Information and Communication Technologies. Springer. 2017. p. 269-278.
- Dahmouni, A., K. Elmoutaouakil, and K. Satori."A Cloud Face Recognition System using A New Optimal Local Binary Pattern". In: Proceedings of the 2nd international Conference on Big Data, Cloud and Applications. ACM. 2017. p. 39.
- Dahmouni, A., Aharrane, N., K. Elmoutaouakil, and K. Satori."Clustering and Jarque-Bera Normality Test to Face Recognition. *Elsevier*. 2018, vol. 127, p. 246-255.
- Aharrane, N., Dahmouni, A., K. Elmoutaouakil, and K. Satori." End-to-end system for printed Amazigh script recognition in document images." In Advanced Technologies for Signal and Image Processing (ATSIP), *IEEE*. 2017 International Conference on (pp. 1-6).

---

---

## *Références*

---

---

- [1]Delac, Kresimir, and Mislav Grgic. "A survey of biometric recognition methods." 46th International Symposium Electronics in Marine. Vol. 46. 2004.
- [2]Faundez-Zanuy, Marcos. "Signature recognition state-of-the-art." IEEE aerospace and electronic systems magazine 20.7 (2005): 28-32.
- [3]Daugman, John. "New methods in iris recognition." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 37.5 (2007): 1167-1175.
- [4]Jain, A., Bolle, R., & Pankanti, S. (Eds.). (2006). Biometrics: personal identification in networked society (Vol. 479). Springer Science & Business Media.
- [5]Zhang, David. "Palmprint identification." Biometric Solutions: For Authentication in an E-World 697 (2012): 243.
- [6]Jain, Anil K., and Stan Z. Li. Handbook of face recognition. New York: springer, 2011.
- [7]Khefif, Bouchra. Mise au point d'une application de reconnaissance faciale. Diss. 2013.
- [8]Yapp Yapp, Robin. "Brazilian police to use 'robocop-style' glasses at world cup." The Telegraph. Retrieved March 5 (2011): 2106.
- [9]Ichino, M. A. N. A. B. U. "General metrics for mixed features the Cartesian space theory for pattern recognition." Systems, Man, and Cybernetics, 1988. Proceedings of the 1988 IEEE International Conference on. Vol. 1. IEEE, 1988.
- [10] Turk, Matthew, and Alex Pentland. "Eigenfaces for recognition." Journal of cognitive neuroscience 3.1 (1991): 71-86.
- [11] Liu, Chengjun, and Harry Wechsler. "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition." IEEE Transactions on Image processing 11.4 (2002): 467-476.
- [12] Dagher, Issam. "Highly-compacted DCT coefficients." Signal, image and video processing 4.3 (2010): 303-307.
- [13] Ramesha, K., and K. B. Raja. "Face recognition system using discrete wavelet transform and fast PCA." Information Technology and Mobile Communication. Springer, Berlin, Heidelberg, 2011. 13-18.

- 
- 
- [14] Manjunath, Bangalore S., and Wei-Ying Ma. "Texture features for browsing and retrieval of image data." *IEEE Transactions on pattern analysis and machine intelligence* 18.8 (1996): 837-842.
- [15] Déniz, Oscar, et al. "Face recognition using histograms of oriented gradients." *Pattern Recognition Letters* 32.12 (2011): 1598-1603.
- [16] Ahonen, Timo, Abdenour Hadid, and Matti Pietikainen. "Face description with local binary patterns: Application to face recognition." *IEEE Transactions on Pattern Analysis & Machine Intelligence* 12 (2006): 2037-2041.
- [17] Perlibakas, Vytautas. "Face recognition using principal component analysis and log-gabor filters." *arXiv preprint cs/0605025* (2006).
- [18] Wiskott, Laurenz, et al. "Face recognition by elastic bunch graph matching." *International Conference on Computer Analysis of Images and Patterns*. Springer, Berlin, Heidelberg, 1997.
- [19] Brahmam, Sheryl, et al., eds. *Local binary patterns: new variants and applications*. Springer Berlin Heidelberg, 2014.
- [20] Pietikäinen, Matti, et al. *Computer vision using local binary patterns*. Vol. 40. Springer Science & Business Media, 2011.
- [21] Jain, Anil K., Arun Ross, and Salil Prabhakar. "An introduction to biometric recognition." *IEEE Transactions on circuits and systems for video technology* 14.1 (2004): 4-20.
- [22] Hanley, James A. "Receiver operating characteristic (ROC) curves." *Wiley StatsRef: Statistics Reference Online* (2014).
- [23] El-Abed, Mohamad. *Évaluation de système biométrique*. Diss. Université de Caen, 2011.
- [24] Tan, Xiaoyang, and Bill Triggs. "Enhanced local texture feature sets for face recognition under difficult lighting conditions." *International Workshop on Analysis and Modeling of Faces and Gestures*. Springer, Berlin, Heidelberg, 2007.
- [25] Sarode, Jagdish P., and Alwin D. Anuse. "Face recognition under pose variations." *Int. J. Comput. Sci. Inf. Technol* 5.3 (2014): 2689-2693.
- 
-

- 
- [26] Khan, N. U. "A comparative analysis of facial expression recognition techniques." Advance Computing Conference (IACC), 2013 IEEE 3rd International. IEEE, 2013.
- [27] Tan, Xiaoyang, et al. "Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble." IEEE Transactions on Neural Networks 16.4 (2005): 875-886.
- [28] Faux, Francis. Détection et suivi de visage par la théorie de l'évidence. Diss. Pau, 2009.
- [29] Bianco, Simone, Francesca Gasparini, and Raimondo Schettini. "Adaptive skin classification using face and body detection." IEEE Transactions on Image Processing 24.12 (2015): 4756-4765.
- [30] Wang, Yi-Qing. "An analysis of the Viola-Jones face detection algorithm." Image Processing On Line 4 (2014): 128-148.
- [31] Rowley, Henry A., Shumeet Baluja, and Takeo Kanade. "Neural network-based face detection." Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on. IEEE, 1996.
- [32] Vaswani, Namrata, and Rama Chellappa. "Principal components null space analysis for image and video classification." IEEE Transactions on Image Processing 15.7 (2006): 1816-1830.
- [33] De la Torre, Fernando, et al. "Representational oriented component analysis (ROCA) for face recognition with one sample image per training class." null. IEEE, 2005.
- [34] Moghaddam, Baback. "Principal manifolds and probabilistic subspaces for visual recognition." IEEE Transactions on Pattern Analysis and Machine Intelligence 24.6 (2002): 780-788.
- [35] Wang, Jie, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos. "Selecting discriminant eigenfaces for face recognition." Pattern Recognition Letters 26.10 (2005): 1470-1482.
- [36] Belhumeur, Peter N., João P. Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. Yale University New Haven United States, 1997.
-

- 
- 
- [37] Chen, Li-Fen, et al. "A new LDA-based face recognition system which can solve the small sample size problem." *Pattern recognition* 33.10 (2000): 1713-1726.
- [38] Draper, Bruce A., et al. "Recognizing faces with PCA and ICA." *Computer vision and image understanding* 91.1-2 (2003): 115-137.
- [39] Ekenel, Hazim Kemal, and Bülent Sankur. "Feature selection in the independent component subspace for face recognition." *Pattern Recognition Letters* 25.12 (2004): 1377-1388.
- [40] Xue, Yun, et al. "A modified non-negative matrix factorization algorithm for face recognition." null. IEEE, 2006.
- [41] Visani, Muriel, Christophe Garcia, and Jean-Michel Jolion. "Bilinear discriminant analysis for face recognition." *International Conference on Pattern Recognition and Image Analysis*. Springer, Berlin, Heidelberg, 2005.
- [42] Yang, Ming-Hsuan. "Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods." *fgv*. IEEE, 2002.
- [43] Yang, Jian, et al. "KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition." *IEEE Transactions on pattern analysis and machine intelligence* 27.2 (2005): 230-244.
- [44] Borg, Ingwer, and P. Groenen. "Modern multidimensional scaling: theory and applications." *Journal of Educational Measurement* 40.3 (2003): 277-280.
- [45] Yang, Ming-Hsuan. "Face recognition using extended isomap." *Image Processing. 2002. Proceedings. 2002 International Conference on*. Vol. 2. IEEE, 2002.
- [46] De la Porte, J., et al. "An introduction to diffusion maps." *Proceedings of the 19th Symposium of the Pattern Recognition Association of South Africa (PRASA 2008)*, Cape Town, South Africa. 2008.
- [47] Saul, Lawrence K., and Sam T. Roweis. "An introduction to locally linear embedding." unpublished. Available at: <http://www.cs.toronto.edu/~roweis/lle/publications.html> (2000).
- [48] Donoho, David L., and Carrie Grimes. "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data." *Proceedings of the National Academy of Sciences* 100.10 (2003): 5591-5596.
- [49] Brunelli, Roberto, and Tomaso Poggio. "Face recognition: Features versus templates." *IEEE transactions on pattern analysis and machine intelligence* 15.10 (1993): 1042-1052.
- 
-

- 
- 
- [50] Heisele, Bernd, et al. "Face recognition: component-based versus global approaches." *Computer vision and image understanding* 91.1 (2003): 6-21.
- [51] Lades, Martin, et al. "Distortion invariant object recognition in the dynamic link architecture." *IEEE Transactions on computers* 3 (1993): 300-311.
- [52] Majumder, Goutam, and Mrinal Kanti Bhowmik. "Gabor-fast ICA feature extraction for thermal face recognition using linear kernel support vector machine." *Computational Intelligence and Networks (CINE), 2015 International Conference on*. IEEE, 2015.
- [53] Dhieb, Thameur, et al. "Beta-Elliptic Model for Writer Identification from Online Arabic Handwriting." *Journal of Information Assurance & Security* 11.5 (2016).
- [54] Ojala, Timo, Matti Pietikäinen, and David Harwood. (1996) "A comparative study of texture measures with classification based on featured distributions." *Pattern recognition* 29 (1): 51-59.51-59.
- [55] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. (2002) "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on pattern analysis and machine intelligence* 24 (7): 971-987.
- [56] T. Ahonen, and M. Pietikäinen, Soft histograms for local binary patterns, In *Proceedings of the Finnish signal processing symposium, FINSIG, Vol.5, No.9, 2007*, pp. 1.
- [57] Pang, Yanwei, et al. "Efficient HOG human detection." *Signal Processing* 91.4 (2011): 773-781.
- [58] Vu, Ngoc-Son, and Alice Caplier. "Face recognition with patterns of oriented edge magnitudes." *European conference on computer vision*. Springer, Berlin, Heidelberg, 2010.
- [59] Vu, Ngoc-Son, and Alice Caplier. "Mining patterns of orientations and magnitudes for face recognition." *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011.
- [60] Gao, Xinbo, et al. "Face sketch synthesis algorithm based on E-HMM and selective ensemble." *IEEE Transactions on Circuits and Systems for Video Technology* 18.4 (2008): 487-496.
- [61] Moghaddam, Baback, and Alexander P. Pentland. "Face recognition using view-based and modular eigenspaces." *Automatic Systems for the Identification and Inspection of Humans*. Vol. 2277. International Society for Optics and Photonics, 1994.
- [62] The ORL face database at the AT&T <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase>.
- 
-

- 
- [63] The Yale Face Database, <http://vision.ucsd.edu/content/yale-face-database>.
- [64] Martinez, A., Benavente, R. (1998). AR face database. CVC technical report 24.
- [65] M. Martinez and R. Benavente, "The AR Face Database," [http://rv11.ecn.purdue.edu/~aleix/aleix\\_face\\_DB.html](http://rv11.ecn.purdue.edu/~aleix/aleix_face_DB.html), 2003.
- [66] Jonathon Phillips, Harry Wechsler, Jeffery Huang, and Patrick J Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(5):295–306, (1998).
- [67] Jonathon Phillips, Hyeonjoon Moon, Syed Rizvi, Patrick J Rauss, et al. The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1090–1104, (2000).
- [68] UMIST VISAGE: <http://images.ee.umist.ac.uk/danny/database.html>
- [69] Multi-PIE : <https://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html>
- [70] Extended Yale B : [vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html](http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html)
- [71] XM2VTS: <http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/>
- [72] VMU: <http://www.nist.gov/itl/iad/ig/frgc.cfm>
- [73] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.
- [74] Gary Huang, Marwan Mattar, Honglak Lee, and Erik G Learned-Miller. Learning to align from scratch. In *Advances in Neural Information Processing Systems*, pages 764–772, 2012.
- [75] Van Der Maaten, Laurens, Eric Postma, and Jaap Van den Herik. "Dimensionality reduction: a comparative." *J Mach Learn Res* 10 (2009): 66-71.
- [76] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces", *IEEE Transactions on Pattern Analysis and machine Intelligence*, Vol. 12, No. (1), pp. 103- 108, January 1990.
- [77] Belahcène, M., M. A. Benatia, and S. Djedi. "Vérification de Visage utilisant Gabor et PCA/KPCA." *Science des matériaux (Laboratoire LARHYSS)* 4 (2015).
- [78] Yang, Jian, et al. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." *IEEE transactions on pattern analysis and machine intelligence* 26.1 (2004): 131-137.
-



- 
- 
- [79] Le, Thai Hoang, and Len Bui. "Face recognition based on SVM and 2DPCA." arXiv preprint arXiv: 1110.5404 (2011).
- [80] Nhat, Vo Dinh Minh, and SungYoung Lee. "Kernel-based 2DPCA for face recognition." *Signal Processing and Information Technology, 2007 IEEE International Symposium on*. IEEE, 2007.
- [81] Yan, H., et al. "Face recognition based on Gabor wavelet transform and modular 2dpca." *Proc. PEEE*. 2015.
- [82] Zhao, Wenyi, Rama Chellappa, and P. Jonathon Phillips. *Subspace linear discriminant analysis for face recognition*. Computer Vision Laboratory, Center for Automation Research, University of Maryland, 1999.
- [83] Yu, Hua, and Jie Yang. "A direct LDA algorithm for high-dimensional data—with application to face recognition." *Pattern recognition* 34.10 (2001): 2067-2070.
- [84] Visani, Muriel, Christophe Garcia, and Jean-Michel Jolion. "Two-dimensional-oriented linear discriminant analysis for face recognition." *Computer Vision and Graphics*. Springer, Dordrecht, 2006. 1008-1017.
- [85] Rouabhia, Chahrazed, and Hicham Tebbikh. "Efficient face recognition based on weighted matrix distance metrics and 2DPCA algorithm." *Archives of Control Sciences* 21.2 (2011): 207-221.
- [86] Li, Ming, and Baozong Yuan. "2D-LDA: A statistical linear discriminant analysis for image matrix." *Pattern Recognition Letters* 26.5 (2005): 527-532.
- [87] Nousath, S., G. Hemantha Kumar, and P. Shivakumara. "(2D) 2LDA: An efficient approach for face recognition." *Pattern recognition* 39.7 (2006): 1396-1400.
- [88] I. Dagher, "Highly compacted DCT coefficients, *Signal, image and video processing*, vol.4, no.3, pp. 303-30, 2010.
- [89] Ramesha, K., and K. B. Raja. "Face recognition system using discrete wavelet transform and fast PCA." *Information Technology and Mobile Communication*. Springer, Berlin, Heidelberg, 2011. 13-18.
- [90] Eickeler, Stefan, Stefan Müller, and Gerhard Rigoll. "Recognition of JPEG compressed face images based on statistical methods." *Image and Vision Computing* 18.4 (2000): 279-287.
- 
-

- 
- [91] F. Gmira, S. Hraoui, A. Saaidi, A.J. Oulidi, K. Satori, Securing the architecture of the JPEG compression by an dynamic encryption, In Intelligent Systems and Computer Vision (ISCV)IEEE, 2015, pp.1-6.
- [92] Omaia, Derzu, JanKees vd Poel, and Leonardo V. Batista. "2D-DCT distance based face recognition using a reduced number of coefficients." Computer Graphics and Image Processing (SIBGRAPI), 2009 XXII Brazilian Symposium on. IEEE, 2009.
- [93] Huang, Z.H., Li, W.J., Shang, J., Wang, J., Zhang, T., 2015, Non-uniform patch based face recognition via 2DDWT. Image and Vision Computing, 37, p.12-19.
- [94] Ayyavoo, Thamizharasi, and Jayasudha John Suseela. "Illumination pre-processing method for face recognition using 2D DWT and CLAHE." IET Biometrics 7.4 (2017): 380-390.
- [95] Cornuéjols, Antoine, and Laurent Miclet. Apprentissage artificiel : concepts et algorithmes. Editions Eyrolles, 2011.
- [96] De Beauville, Jean-Pierre Asselin, and Fatima-Zohra Kettaf. Bases théoriques pour l'apprentissage et la décision en reconnaissance des formes. Cépaduès, 2005.
- [97] Singh, Shio Kumar, M. P. Singh, and D. K. Singh. "A survey on network security and attack defense mechanism for wireless sensor networks." International Journal of Computer Trends and Technology 1.2 (2011): 9-17.
- [98] Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." nature 323.6088 (1986): 533.
- [99] Ciresan, Dan Claudiu, et al. "Convolutional neural network committees for handwritten character classification." Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE, 2011.
- [100] Sattari, M. Taghi, Halit Apaydin, and Fazli Ozturk. "Flow estimations for the Sohu Stream using artificial neural networks." Environmental Earth Sciences 66.7 (2012): 2031-2045.
- [101] Garrigues, Grant E., et al. "Gene expression clustering using self-organizing maps: analysis of the macrophage response to particulate biomaterials." Biomaterials 26.16 (2005): 2933-2945.
- [102] Du, Ke-Lin, and M. N. S. Swamy. "Recurrent neural networks." Neural Networks and Statistical Learning. Springer, London, (2014). 337-353.
-

- 
- [103] Lin, Shang-Hung, Sun-Yuan Kung, and Long-Ji Lin. "Face recognition/detection by probabilistic decision-based neural network." *IEEE transactions on neural networks* 8.1 (1997): 114-132.
- [104] Widrow, Bernard, and Marcian E. Hoff. *Adaptive switching circuits*. No. TR-1553-1. Stanford Univ Ca Stanford Electronics Labs, 1960.
- [105] Minsky, Marvin, and Seymour Papert. "Perceptron Expanded Edition." (1969).
- [106] Hopfield, John J. "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the national academy of sciences* 79.8 (1982): 2554-2558.
- [107] Rossi, Fabrice, and Briec Conan-Guez. "Multi-layer perceptron on interval data." *Classification, Clustering, and Data Analysis*. Springer, Berlin, Heidelberg, (2002). 427-434.
- [108] Pfister, Marcus, and Raul Rojas. "Speeding-up backpropagation-A comparison of orthogonal techniques." *Neural Networks. IJCNN'93-Nagoya. Proceedings of 1993 International Joint Conference on*. Vol. 1. IEEE, 1993.
- [109] Rajabioun, Ramin. "Cuckoo optimization algorithm." *Applied soft computing* 11.8 (2011): 5508-5518.
- [110] Tosun, Ömür. "Cuckoo search algorithm." *Encyclopedia of Business Analytics and Optimization*. IGI Global, (2014). 558-564.
- [111] Chatterjee, Sankhadeep, et al. "Hybrid modified Cuckoo Search-Neural Network in chronic kidney disease classification." *Engineering of Modern Electric Systems (EMES), 2017 14th International Conference on*. IEEE, 2017.
- [112] Vapnik, Vladimir, and Sayan Mukherjee. "Support vector method for multivariate density estimation." *Advances in neural information processing systems*. 2000.
- [113] Rajput, G. G., and Rajeshwari Horakeri. "Unconstrained Kannada Handwritten Character Recognition Using Multi-level SVM Classifier." *International Conference on Pattern Recognition and Machine Intelligence*. Springer, Berlin, Heidelberg, 2013.
- [114] Haridas, Nikhila, V. Sowmya, and K. P. Soman. "GURLS vs LIBSVM: performance comparison of kernel methods for hyperspectral image classification." *Indian Journal of Science and Technology* 8.24 (2015).
-

- 
- [115] Chang, Chih-Chung. "" LIBSVM: a library for support vector machines," ACM Transactions on Intelligent Systems and Technology, 2: 27: 1--27: 27, 2011." <http://www.csie.ntu.edu.tw/~cjlin/libsvm> 2 (2011).
- [116] Platt, John. "Sequential minimal optimization: A fast algorithm for training support vector machines." (1998).
- [117] Abe, Shigeo. "Fusing sequential minimal optimization and Newton's method for support vector training." International Journal of Machine Learning and Cybernetics 7.3 (2016): 345-364.
- [118] Yong, Zhou, Li Youwen, and Xia Shixiong. "An improved KNN text classification algorithm based on clustering." Journal of computers 4.3 (2009): 230-237.
- [119] D.M. POWERS, EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, and MARKEDNESS & CORRELATION. 2011.
- [120] Liu CL, Fujisawa H (2008) Classification and learning methods for character recognition: Advances and remaining problems. Studies in Computational Intelligence 90:139–161. doi: 10.1007/978-3-540-76280-5\_6.
- [121] Kittler, Josef, et al. "On combining classifiers." IEEE transactions on pattern analysis and machine intelligence 20.3 (1998): 226-239.
- [122] X Tan, B Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans, Image Process, 19(6), 1635–1650 (2007).
- [123] Jun, Bongjin, and Daijin Kim. "Robust face detection using local gradient patterns and evidence accumulation." Pattern Recognition 45.9 (2012): 3304-3316.
- [124] Nanni, Loris, Alessandra Lumini, and Sheryl Brahmam. "Survey on LBP based texture descriptors for image classification." Expert Systems with Applications 39.3 (2012): 3634-3641.
- [125] Brahmam, Sheryl, et al., (Eds.). (2014) "Local binary patterns: new variants and applications." Springer Berlin Heidelberg.
- [126] Pietikäinen, Matti, et al. Computer vision using local binary patterns. Vol. 40. Springer Science & Business Media, 2011.
- [127] Wang, Xian, et al. "The face recognition algorithm based on improved lbp." Opto-Electronic Engineering 39.7 (2012): 109-114.
-

- 
- 
- [128] Hafiane, Adel, Kannappan Palaniappan, and Guna Seetharaman. "Joint adaptive median binary patterns for texture classification." *Pattern Recognition* 48.8 (2015): 2609-2620.
- [129] Heikkilä, Marko, Matti Pietikäinen, and Cordelia Schmid. "Description of interest regions with center-symmetric local binary patterns." *Computer vision, graphics and image processing*. Springer, Berlin, Heidelberg, 2006. 58-69.
- [130] Fu, Xiao-feng, and Wei Wei. "Facial expression recognition based on multi-scale centralized binary pattern." *Control Theory & Applications* 26.6 (2009): 629-633.
- [131] Jabid, Taskeed, Md Hasanul Kabir, and Oksam Chae. "Local directional pattern (LDP) for face recognition." *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on*. IEEE, 2010.
- [132] Nguyen, Huu-Tuan, and Alice Caplier. "Elliptical local binary patterns for face recognition." *Asian Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2012.
- [133] Guo, Zhenhua, Lei Zhang, and David Zhang. "A completed modeling of local binary pattern operator for texture classification." *IEEE Transactions on Image Processing* 19.6 (2010): 1657-1663.
- [134] Kittler J, Hater M, Duin RPW (1996) Combining classifiers. In: *Proceedings - International Conference on Pattern Recognition*. pp 897–901.
- [135] Papakostas, George A., et al. "Moment-based local binary patterns: a novel descriptor for invariant pattern recognition applications." *Neurocomputing* 99 (2013): 358-371.
- [136] Ahonen, Timo, et al. "Recognition of blurred faces using local phase quantization." *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008.
- [137] ZHANG, Lun, CHU, Rufeng, XIANG, Shiming, et al. Face detection based on multi-block lbp representation. In : *International conference on biometrics*. Springer, Berlin, Heidelberg, 2007. p. 11-18.
- [138] Benzaoui, Amir, and Abdelhani Boukrouche. "1DLBP and PCA for face recognition." *Programming and Systems (ISPS), 2013 11th International Symposium on*. IEEE, 2013.
- [139] Tang, Hengliang, et al. "3D face recognition using local binary patterns." *Signal Processing* 93.8 (2013): 2190-2198.
- [140] Thadewald, Thorsten, and Herbert Büning. "Jarque–Bera test and its competitors for testing normality—a power comparison." *Journal of Applied Statistics* 34.1 (2007): 87-105.
- 
-

- 
- [141] Razali, Nornadiah Mohd, and Yap Bee Wah. "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests." *Journal of statistical modeling and analytics* 2.1 (2011): 21-33.
- [142] Jarque, Carlos M., and Anil K. Bera. "Efficient tests for normality, homoscedasticity and serial independence of regression residuals." *Economics letters* 6.3 (1980): 255-259.
- [143] Ajitha, S., et al. "Face recognition system using Combined Gabor Wavelet and DCT approach." *Recent Trends in Information Technology (ICRTIT), 2014 International Conference on. IEEE, 2014.*
- [144] Yang, Jian, et al. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." *IEEE transactions on pattern analysis and machine intelligence* 26.1 (2004): 131-137.
- [145] Déniz, Oscar, M. Castrillon, and Mario Hernández. "Face recognition using independent component analysis and support vector machines." *Pattern recognition letters* 24.13 (2003): 2153-2157.
- [146] Ghaforiyan, Hoda, and Mehran Emadi. "Human face recognition under pose variation with fusion geometric methods." *International Academic Journal of Science and Engineering* 3.1 (2016): 1-10.
- [147] Yu, Wei, et al. "An improved LBP algorithm for texture and face classification." *Signal, Image and Video Processing* 8.1 (2014): 155-161.
- [148] Lu, Gui-Fu, Zhong Jin, and Jian Zou. "Face recognition using discriminant sparsity neighborhood preserving embedding." *Knowledge-Based Systems* 31 (2012): 119-127.
- [149] Huang, Zheng-Hai, et al. "Face recognition based on pixel-level and feature-level fusion of the top-level's wavelet sub-bands." *Information Fusion* 22 (2015): 95-104.
- [150] Ren, Huorong, and Hongxin Ji. "Nonparametric subspace analysis fused to 2DPCA for face recognition." *Optik-International Journal for Light and Electron Optics* 125.8 (2014): 1922-1925.
- [151] Fathi, Abdolhossein, Pendar Alirezazadeh, and Fardin Abdali-Mohammadi. "A new Global-Gabor-Zernike feature descriptor and its application to face recognition." *Journal of Visual Communication and Image Representation* 38 (2016): 65-72.
- [152] Dornaika, Fadi, and Alireza Bosaghzadeh. "Exponential local discriminant embedding and its application to face recognition." *IEEE transactions on cybernetics* 43.3 (2013): 921-
-

---

934.

[153] H. Yan, P. Wang, W. D. Chen, and J. Liu, "Face recognition based on gabor wavelet transform and modular 2 DPCA," in Proc. PEEE (Hong Kong, 2015), pp. 245–248.

[154] Mashhoori, Ali, and Mansoor Zolghadri Jahromi. "Block-wise two-directional 2DPCA with ensemble learning for face recognition." *Neurocomputing* 108 (2013): 111-117.

[155] S. Du and R. K. Ward, "Improved face representation by nonuniform multilevel selection of Gabor convolution features," *IEEE Trans. Syst., Man, Cyberneti.-Part B: Cybernet.* 39(6) (2009) : 1408–1419.

[156] Rai Preeti and Pritee Khanna, "An illumination, expression, and noise invariant gender classifier using two-directional 2DPCA on real Gabor space," *J. Visual Lang. Comput.* 26 (2015): 15–28.

[157] Xia Wei, Shouyi Yin, and Peng Ouyang, "A high precision feature based on LBP and Gabor theory for face recognition," *Sensors* 13.4 (2013): 4499–4513.

[158] Wright, John, et al. "Robust face recognition via sparse representation." *IEEE transactions on pattern analysis and machine intelligence* 31.2 (2009): 210-227.

[159] Chude-Olisah, Chollette Chiazor. *New Face Recognition Descriptor Based on Edge Information for Surgically-altered Faces in Uncontrolled Environment*. Diss. Universiti Teknologi Malaysia, 2015.

[160] Bouda, B., et al. "Gray-level corner detection by virtual electric field model." 2007 International Symposium on Computational Intelligence and Intelligent Informatics. IEEE, 2007.

[161] Béguin, François. "Henri Poincaré et ses théoremes d'uniformisation."

[162] Lin, Guangfeng, et al. "Visual feature coding based on heterogeneous structure fusion for image classification." *Information Fusion* 36 (2017): 275-283.

[163] Zhou, Shu-Ren, Jian-Ping Yin, and Jian-Ming Zhang. "Local binary pattern (LBP) and local phase quantization (LBQ) based on Gabor filter for face representation." *Neurocomputing* 116 (2013): 260-264.

[164] Dornaika, F. (Ed.). *Advances in Face Image Analysis: Theory and applications*. Bentham Science Publishers. (2016).

[165] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans on Pattern Analysis and Machine Intelligence* 31(2) (2009)

---

---

210–227.

[166] Yang, Jian, et al. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." *IEEE transactions on pattern analysis and machine intelligence* 26.1 (2004): 131-137.

[167] Minoux, Michel. *Programmation mathématique: théorie et algorithmes*. Dunod, 1983.

[168] Aarts, E. H. L., and v. Laarhoven (1985), *Statistical cooling: A general approach to combinatorial optimization problems*, *Philips J. Res.*, 40(4), 193-226.

[169] Dubuc, Ghislain, et al. "Une heuristique de recherche avec tabous pour la conception de réseaux de distribution de contenu électronique." *INFOR: Information Systems and Operational Research* 45.3 (2007): 175-185.

[170] Omolaye, Philip O., Joseph M. Mom, and Gabriel A. Igwue. "A Holistic Review of Soft Computing Techniques." *Applied and Computational Mathematics* 6.2 (2017): 93.

[171] Mirjalili, Seyedali. "The ant lion optimizer." *Advances in Engineering Software* 83 (2015): 80-98.

[172] Kennedy, James. "Particle swarm optimization." *Encyclopedia of machine learning*. Springer, Boston, MA, 2011. 760-766.

[173] Shannon, Claude Elwood. "A mathematical theory of communication." *Bell system technical journal* 27.3 (1948): 379-423.

[174] Vishwakarma, Sapna, and Krishan Kant Pathak. "Face recognition using LBP Coefficient Vectors with SVM Classifier." *International Journal of Engineering Trends and Technology (IJETT)* 9.2 (2014): 96-100.

[175] Le, Thai Hoang, and Len Bui. "Face recognition based on SVM and 2DPCA." *arXiv preprint arXiv:1110.5404* (2011).

[176] Yan, H., et al. "Face recognition based on gabor wavelet transform and modular 2dPCA." *Proc. PEEE*. 2015.

[177] Zhou, Changjun, et al. "Face recognition based on PCA and logistic regression analysis." *Optik-International Journal for Light and Electron Optics* 125.20 (2014): 5916-5919.

[178] Chihouai, Mejda, et al. "A survey of 2D face recognition techniques." *Computers* 5.4 (2016): 21.

[179] Elgarrai, Zineb, et al. "Offline Face Recognition System Based on Gabor-Fisher Descriptors and Hidden Markov Models." *IJIMAI* 4.1 (2016): 11-14.