



N° Ref :.....

Centre Universitaire de Mila

Institut des Sciences et de la Technologie

Département de Mathématique
et Informatique

Mémoire préparé En vue de l'obtention du diplôme de Master

en : - Filière : Informatique
- Spécialité : STIC
- Option : Professionnel

Thème

La reconnaissance de visages par combinaison de classificateurs

Préparé par : Boukemara Fadila

Soutenue devant le jury :

- Président :... Yesaadi SabrinaGrade :... Maitre-assistant classe A..
- Examineur :Dib Abderrahim....Grade :.....Maitre-assistant classe A
- Promoteur :Salima Nebti....Grade :.....Maitre de conférences classe B.....

Année universitaire : 2013/2014

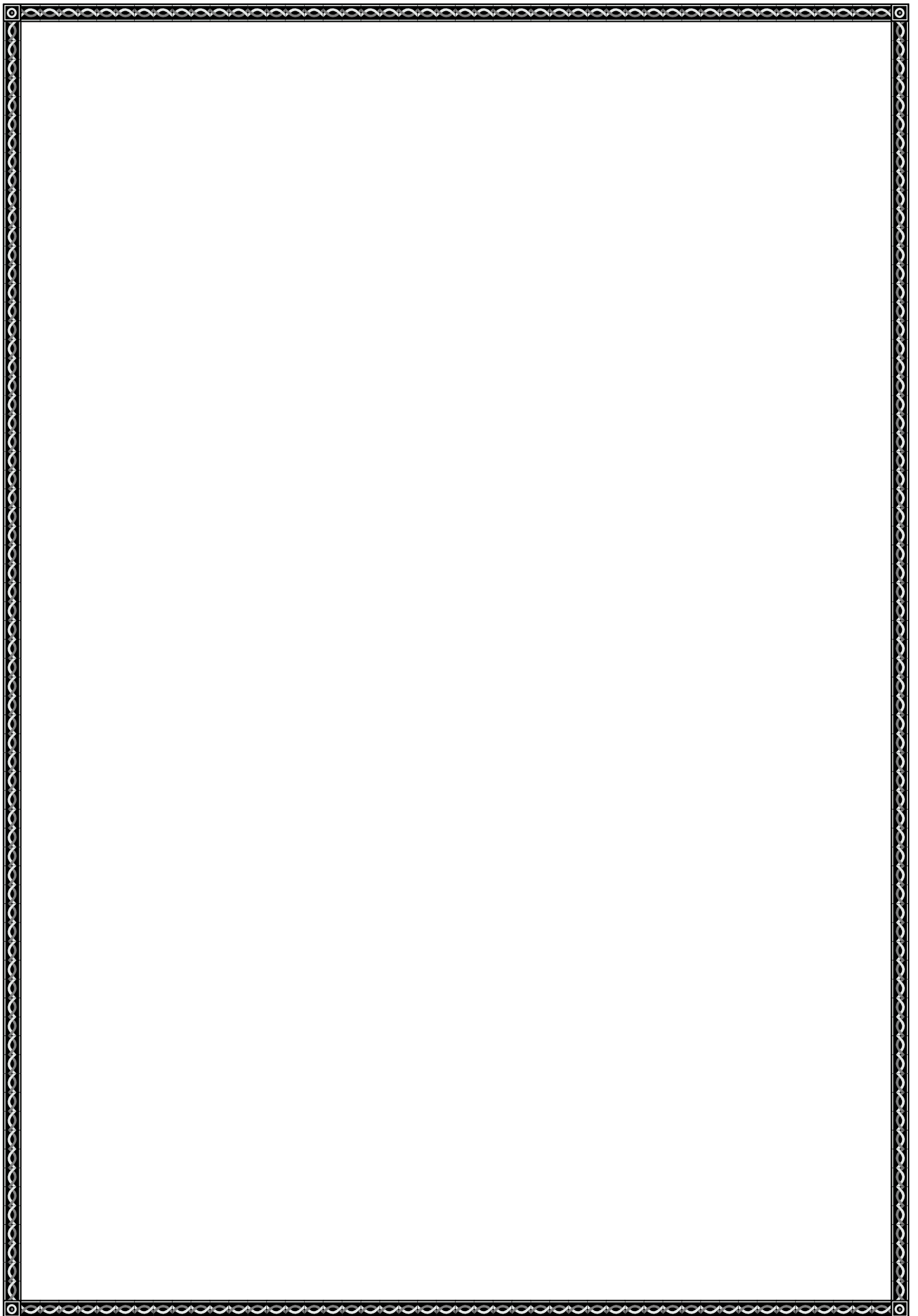


Table des matières

1. Introduction Générale :	6
2. Historique	6
3. Motivations	7
4 . Objectif	7
1. Résumé	8
2. La reconnaissance faciale :	9
2. Applications	9
2.1 Identification	9
2.2 Vérification	10
3. Architecture générale:	11
3.1 Capture :	11
3.2 Détection du visage :	12
3.3 Extraction de caractéristiques: 1	12
3.4 Comparaison des caractéristiques:	13
4. Principales difficultés de la reconnaissance faciale	13
4.1 Changement d'illumination	13
4.2 Variation de pose	14
4.3 Expressions faciales	14
4.4 Occultations partielles	15
5. Conclusion	15
1. Résumé :	16
2. La détection de visages	17
2.1 Approches basées sur les connaissances acquises	17
2.2 Approches basées sur le « Template-matching »	18
2.3 Approches basées sur l'apparence	19
2.4 Approches basées sur des caractéristiques invariantes	19
2.4.1 Détection de visages basée sur l'analyse de la couleur de la peau :	20
3 Techniques de reconnaissance de visage	20
3.1 Approches holistiques ou globales	21

3.2 Méthodes locales	23
3.2.1 Méthodes basées sur les caractéristiques locales	23
3.2.2 Méthodes basées sur l'apparence locale	24
2.3.3 Méthodes Hybrides.....	27
4. L'extraction de caractéristiques.....	28
4.1 Les ondelettes de Gabor	29
4.2 Local Binary Pattern.....	30
5. Conclusion.....	33
1. Résumé.....	34
2 Présentation	35
3. Les Machines à Vecteur Support (SVMs).....	35
2.1 Le cas linéairement séparable.....	36
2.2 Le cas non-linéairement séparable	37
2.3 Classification Multi-classes.....	38
3 Pourquoi combiner SVM avec une technique d'extraction de paramètres.....	38
4 L'analyse en composantes principales (PCA).....	39
4.1 Principe générale	39
4.2 Étapes à suivre pour effectuer une analyse en composantes principales.....	40
4.3 Caractéristiques de PCA.....	41
4.4 Classification basée sur PCA-SVM	41
5. Les Modèles de Markov Cachés (MMC).....	41
5.1 Les éléments d'un MMC	41
5.2 Apprentissage d'un MMC.....	42
5.2.1 L'algorithme de Viterbi.....	42
5.2.2 L'algorithme Rabiner:	43
6. Méthode SVD.....	44
7. Combinaison de classificateurs.....	44
8. Types de sorties d'un classificateur.....	44
Les mesures de performances d'un classificateur.....	45
9. Les méthodes de combinaison des classificateurs	45
9.1 Méthodes de vote	45
1. Résumé.....	48

2. Conception:.....	49
2.1. Identification des acteurs du Système:	49
2.2. Identification des cas d'utilisation du Système:	49
2.2.1. Enregistrement :.....	49
2.2.2. Vérification :.....	50
2.2.3. Identification :.....	51
2.2.4. Calculs de performances :.....	52
Taux de Reconnaissance:.....	52
Algorithme de calcul du taux de reconnaissance :	53
3. Structure de Données et Implémentation :	53
3.1 Environnement du travail.....	53
3.1. Environnements matériel.....	53
3.2 Outils de développement :.....	54
3.2.1 Matlab 2011	54
4.1 Interface Authentification	54
4.2 Interface Présentation.....	55
Conclusion Générale:.....	59
Bibliographie :	60

Liste des Figures

Figure 1 : Scores de compatibilité pour différentes technologies biométriques.	9
Figure 2 : Schéma d'Identification du visage.....	10
Figure 3 : Schéma de vérification du visage	11
Figure 4:Schéma général d'un système de reconnaissance de visage.....	11
Figure 5:Normalisation géométrique du visage	12
Figure 6:Exemple de variation d'éclairage	14
Figure 7:Exemples de variation de poses.....	14
Figure 8:Exemples de variation d'expressions.	15
Figure 9:Modèle de visage composé de 16 régions	18
Figure 10:Les dix vues d'une personne dans la base de données ORL.....	22
Figure 11:Image de visage partitionnée en image	25
Figure 12: une image du visage originale, sa projection et image reconstruite	26
Figure 13:Exemple de représentation faciale en ondelettes de Gabor	29
Figure 14:Opérateur LBP.....	30
Figure 15: Trois voisinages pour des R et P différents	31
Figure 16:Représentation d'un visage par les histogrammes du code LBP	32
Figure 17:Discrimination linéaire.....	37
Figure 18:projection suivant le principe de la PCA.....	39
Figure 19:Une chaîne de Markov à 3 états avec les processus associés	42
Figure 20:Les méthodes de vote	46
Figure 21: diagramme de séquence pour l'enregistrement d'une personne	49
Figure 22:diagramme de séquence pour la vérification.....	50
Figure 23:diagramme de séquence pour identification.....	51
Figure 24:diagramme de séquence pour le calcul des performances du système.	52
Figure 25:Interface graphique d'authentification.....	54
Figure 26:Interface graphique de reconnaissance faciale	55
Figure 27:Le panneau face recognition	56
Figure 28:8Interface graphique de reconnaissance faciale générale	56

Figure 29:Panel HMM.....	57
Figure 30:Panel SVM.....	57
Figure 31:Panel Combinaison	58

Liste des Tables

Tableau 1:Comparaison des méthodes basées sur les caractéristiques locales ou globales.....	28
Tableau 2:Table récapitulative des résultats expérimentaux.....	58

1. Introduction Générale :

L'authentification de visage a gagné une attention considérable ces dernières années, par le besoin croissant de systèmes de vérification d'accès en utilisant plusieurs modalités (voix, image de visage, empreintes digitales, cartes à puce, etc.). De tels systèmes sont employés pour la vérification de l'identité d'un utilisateur en net, en utilisant un automate de banque, en entrant dans un lieu sécurisé, etc.. .

L'authentification de visage est différente de l'identification de visage (ou de la classification): dans des tâches d'authentification, le système sait a priori l'identité de l'utilisateur (par exemple par le code de sa carte à puce), et on doit vérifier cette identité; en d'autres termes, le système doit décider si l'utilisateur à priori est un imposteur ou pas. Dans l'identification de visage, l'identité a priori n'est pas connue : le système doit décider lesquelles des images stockées dans une base de données ressemblent plus à l'image à identifier; la décision n'est plus binaire.

Travailler avec des images approximatives n'est pas efficace : dans l'authentification de visage, plusieurs images d'une même personne peuvent être nettement différentes, en raison des changements de point de vue, en couleurs et l'illumination, ou tout simplement parce que le look du visage de ce même personne est différent de jour en jour.[1]

2. Historique

La reconnaissance faciale est une technique biométrique qui a une longue histoire.

Si l'empreinte digitale est la technique biométrique la plus ancienne et la plus répandue et qui a été adoptée par Scotland Yard dès 1901 et par la préfecture de Police de Paris, dès 1903 pour rechercher les criminels, la reconnaissance de visages a été développé par Benton et Van Allen en 1968 pour évaluer la capacité d'identification des visages non familiers. Il ne s'agit pas d'un test de reconnaissance mnésique de visages familiers ou non familiers, mais d'une épreuve consistante à apparier des photographies de visages non familiers présentés sous différents éclairages et selon des angles différents et nécessitant une bonne capacité d'intégration Visio-spatiale.[1]

Actuellement cette technique est utilisée principalement pour des raisons sécuritaires lors des vols, d'identification et de contrôle d'accès : la comparaison se fait alors entre deux photos: l'une est extraite de l'enregistrement des caméras de surveillance, l'autre dite « image de référence ou image cible » correspondant au suspect, elle a été prise dans des conditions idéales et de bonne qualité.

L'utilisation des techniques de reconnaissance faciale a connu un développement à grande échelle depuis le milieu des années 90, avec l'utilisation efficace de nouvelles technologies,

notamment l'ordinateur et sa capacité de traitement d'images ce qui lui a donné des nouvelles applications.

3. Motivations

S'investir dans le domaine de la reconnaissance de visages est sans doute motivé par la multiplicité et la variété des champs d'applications de celui-ci (contrôle d'accès, télésurveillance et identification des criminels...). L'intérêt pour ce domaine s'explique aussi par le fait que par rapport aux autres filières de reconnaissance (reconnaissance de l'empreinte digitale, de la voix, etc.), la reconnaissance de visages n'a pas atteint un niveau aussi avancé que ces autres filières, elle souffre de plusieurs anomalies. Ceci s'explique avant tout par la complexité de la forme en question (le visage) comparée avec les autres formes, des efforts importants sont fournis dans ce domaine de la recherche biométrique dont les besoins en sécurité (internet, e-Learning, criminologie, transactions bancaire) sont de plus en plus importants et où les enjeux économiques sont énormes. L'image faciale est la forme biométrique la plus communément employée pour l'identification et le contrôle d'accès par authentification de personnes. Les différentes formes d'information extraite du visage, ont poussé vers l'utilisation de plusieurs techniques, mais aucune d'elles n'est totalement efficace, d'où l'idée d'en combiné plusieurs est apparue pour profiter des avantages de chacune.

4 . Objectif

La reconnaissance de visages est une technique biométrique moderne, elle est très surveillante, l'objectif de ce travail est de concevoir un système de la reconnaissance faciale capable d'identifier une personne avec un taux de reconnaissance très élevé basé sur la combinaison de classificateurs efficaces notamment le SVM (Machine à vecteurs supports) et les HMMs (les modèles de Markov caché). Notre travail a étudié et évalué à travers des expériences l'efficacité de ces deux classificateurs et de leur fusion sur l'identification de visages.

Chapitre 1 :

La Reconnaissance Faciale

1. Résumé

Le but de ce chapitre est de donner un panorama des méthodes les plus significatives en reconnaissance 2D de visages. Au cours de ce chapitre, une brève présentation des méthodes les plus populaires utilisées en reconnaissance faciale est proposée.

2. La reconnaissance faciale :

La reconnaissance faciale est une tâche que les humains effectuent naturellement et sans effort dans leur vie quotidienne. La grande disponibilité d'ordinateurs puissants et peu onéreux ainsi que des systèmes informatiques embarqués ont suscité un énorme intérêt dans le traitement automatique des images et des vidéos numériques au sein de nombreuses applications, incluant l'identification biométrique, la surveillance, l'interaction homme-machine et la gestion de données multimédia. La reconnaissance faciale, en tant qu'une des technologies biométriques de base, a pris une part de plus en plus importante dans le domaine de la recherche, ceci étant dû aux progrès rapides des technologies telles que les appareils photo numériques, Internet et les dispositifs mobiles, le tout associé à des besoins de sécurité en augmentation. La reconnaissance faciale possède plusieurs avantages par rapport aux autres technologies biométriques : elle est naturelle et facile à utiliser.

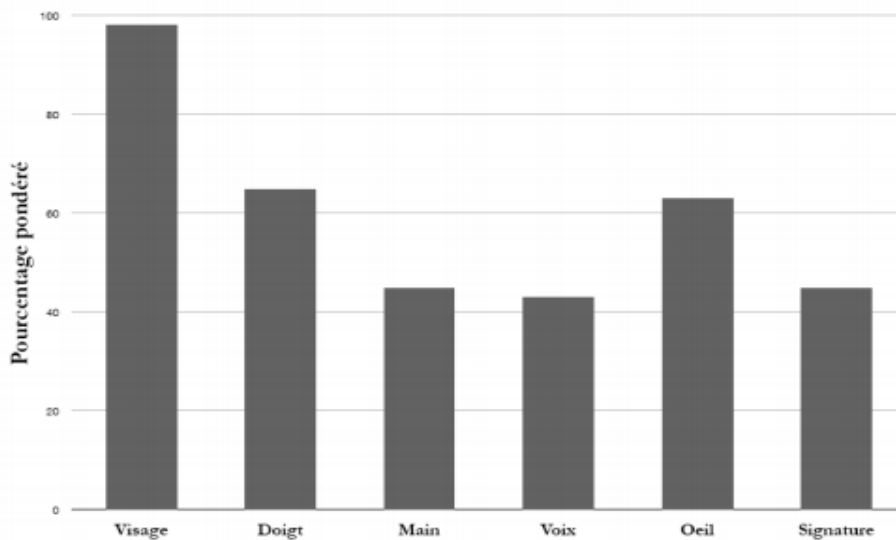


Figure 1 : Scores de compatibilité pour différentes technologies biométriques.

2. Applications

Les systèmes de reconnaissance peuvent être divisés en deux grandes catégories selon leur fonctionnalité :

2.1 Identification

Ayant un flux d'image, ou une image qui contient un visage, l'identification est la reconnaissance à partir d'une base de données sans autre connaissance a priori de l'identité de la personne. Cette tâche donne lieu à une comparaison de type 1-to-N : compare une image du flux avec les images enregistrées dans la base de donnée (figure 2). Une application est la

reconnaissance de personnes « dangereuses » enregistrées dans une WATCH LIST. Cette application est sujette à des contraintes de temps de calcul d'où l'enregistrement de caractéristiques du visage au lieu de l'image entière .Aussi l'enregistrement d'une base de donnée de visage est conditionné, en France par l'autorisation de la CNIL (Commission Nationale Informatique et Liberté)qui cautionne la protection des données personnelles. Cette autorisation n'est donnée qu'en fonction du fort besoin de sécurité nécessaire par l'application. [2]

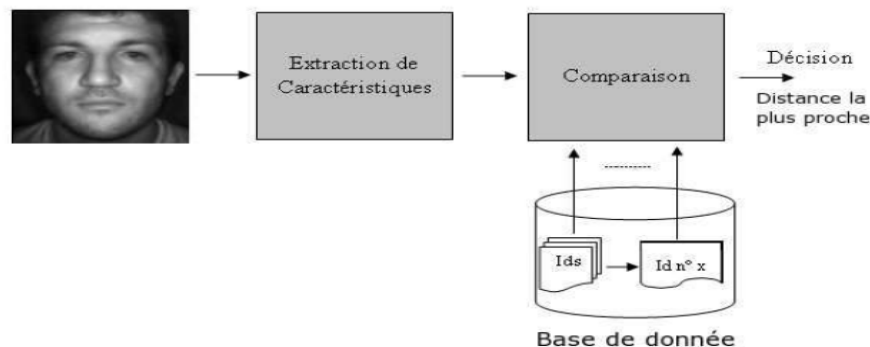


Figure 2 : Schéma d'Identification du visage

2.2 Vérification.

Ayant un flux d'image, ou une image qui contient un visage, la vérification est la reconnaissance de visage à partir d'une image de référence enregistrée au préalable dans une puce ou un fichier. C'est une vérification de l'identité revendiquée c'est-à-dire que personne déclare au préalable son identité (figure 3). Cette tâche donne lieu à une comparaison 1-to-1. Une des applications est le contrôle ou vérification d'identité à l'entrée des Etats- Unis par exemple (ex : le nouveau passeport biométrique). [2]

La référence biométrique peut alors être enregistrée dans une base de données mais aussi sur une carte à puce détenue par la personne .Dans ce cas la CNIL émet un avis moins restrictif à l'utilisation de la biométrie. Ainsi on peut trouver des appareils électronique personnels utilisant la biométrie tels que les ordinateurs portables, les coffres, les téléphones etc.

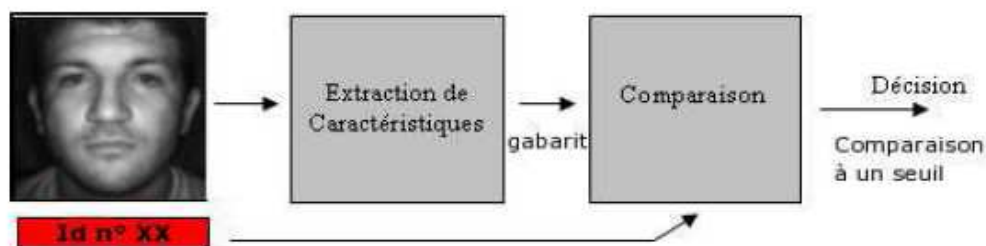


Figure 3 : Schéma de vérification du visage

3. Architecture générale:

Un système de reconnaissance de visage peut être considéré comme composé de 4 parties (figure 4). [2]

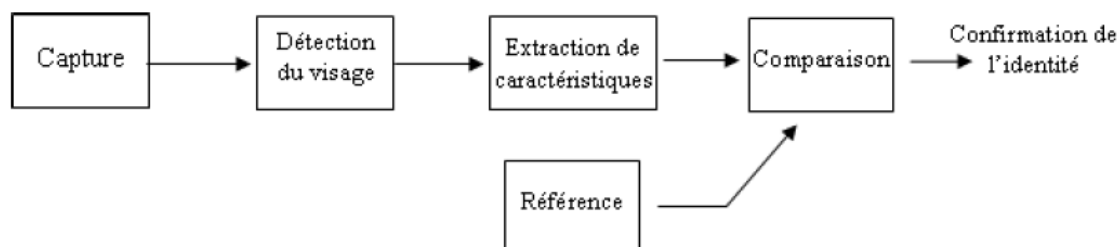


Figure 4: Schéma général d'un système de reconnaissance de visage

3.1 Capture : C'est la première étape dans le processus. C'est aussi l'acquisition des informations et leur transfert vers l'unité de traitement. Elle est une étape très importante dans les systèmes de reconnaissance. En effet, avoir des images de bonne qualité en référence améliore les performances du système de reconnaissance. Il faut réussir à capturer l'information pertinente sans bruit. Il existe plusieurs types de capteurs pour l'acquisition du visage qui se distinguent par leur mode de fonctionnement, leurs échelles spectrales et leur mode d'acquisition.

On trouve sur le marché les capteurs d'image classiques à 2D tels que : les CCD (Couple charged device) ou CMOS pour capturer des images dans le spectre visible et/ou proche-infrarouge, ou les capteurs thermiques qui permettent une acquisition dans l'infrarouge. Il existe des capteurs qui nous donnent une image avec l'information 3D, cela se fait par des scanners 3D, où la mesure de la profondeur est réalisée grâce à un rayon laser balayant la scène ou par stéréo vision. Chaque type de capteur présente des avantages et des inconvénients.

Dans la reconnaissance de visage on peut utiliser les capteurs 3D par exemple afin de se libérer des problèmes de pose. Mais leur prix excessif ne permet pas leur utilisation à grande échelle. Les capteurs spectroscopiques proches infrarouges sont utilisés pour éliminer les problèmes d'illumination. [2]

3.2 Détection du visage : Après avoir capturé la scène contenant un visage, la deuxième étape consiste à l'extraire de l'image. Cela peut se faire par détection de la couleur de la peau, ou par des méthodes détectant les différentes caractéristiques du visage par des descripteurs locaux. Cette étape est plus délicate lorsque l'image acquise contient plusieurs visages ou un fond non uniforme qui crée une texture perturbant la bonne segmentation du visage. Cette étape est aussi dépendante de la qualité des images acquises. Après la segmentation du visage, on peut filtrer ou améliorer la qualité par des prétraitements qui sont appliqués au visage extrait. On peut effectuer des normalisations géométriques et photométriques. Ces prétraitements sont nécessaires pour éliminer ou limiter les variations de pose ou d'illumination.

Un prétraitement photométrique tend à uniformiser l'éclairage dans une image et ainsi minimiser l'influence de l'illumination. Cela peut être effectué soit par des méthodes simples telle que l'égalisation d'histogramme, une correction gamma ou par des méthodes plus complexes tel que le lissage anisotropique [6] ou la méthode retinex. Une normalisation géométrique est un ajustement du visage pour qu'il ait une dimension donnée. La taille du visage est généralement donnée par la distance inter-oculaire. La figure 5 montre la normalisation géométrique.

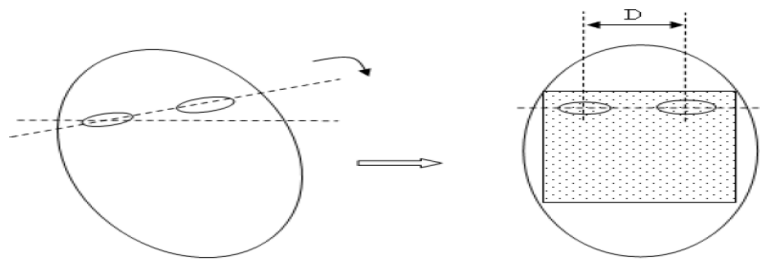


Figure 5: Normalisation géométrique du visage

Tous ces prétraitements tendent à éliminer les variations de formes du visage (rotation, taille) et à avoir une image uniformément éclairée. Ainsi, l'image du visage ressemble à une image acquise dans de bonnes conditions, comparable à celles de l'image de référence. On diminuera donc les chances d'un faux rejet.[2]

3.3 Extraction de caractéristiques: le but est d'extraire les caractéristiques du visage qui peuvent le rendre à la fois différent des autres visages et robuste aux variations du visage de la même personne. C'est l'information nécessaire pour que le visage d'une personne soit distinct des autres personnes mais ressemble au même visage acquis en différentes conditions d'éclairage et de pose. Les anciens travaux de la reconnaissance de visage [7][8], la représentation du visage devrait passer par l'utilisation de la bouche, des yeux, du nez, de leurs positions relatives et de leurs géométrie. Cependant, cette procédure a montré des limites [9]. Il faut alors une analyse plus sophistiquée du visage afin de localiser autres caractéristiques plus robustes. Certaines

méthodes n'utilisent que la détection des yeux pour normaliser le visage et ensuite font une analyse globale du visage (algorithmes de type CPA, LDA, etc.)

3.4 Comparaison des caractéristiques: les algorithmes de comparaison diffèrent selon les caractéristiques précédemment extraites. On trouve dans la littérature plusieurs approches : basées sur le calcul de distance, le calcul de similarité .D'autres méthodes sont basées sur la classification des attributs caractéristiques utilisant un classificateur spécifique tels que le SVM, le classificateur bayésien ou utilisant plusieurs classificateurs telle que la méthode Adaboost. [2]

4. Principales difficultés de la reconnaissance faciale

Pour le cerveau humain, le processus de la reconnaissance de visages est une tâche visuelle de haut niveau. Bien que les êtres humains puissent détecter et identifier des visages dans une scène sans beaucoup de peine, construire un système automatique qui accomplit de telles tâches représente un défi sérieux. Ce défi est d'autant plus grand lorsque les conditions d'acquisition des images sont très variables. Il existe deux types de variations associées aux images de visages : inter et intra classe. La variation interclasse est limitée à cause de la ressemblance physique entre les individus. Par contre la variation intra-classe est vaste. Elle peut être due à plusieurs facteurs que nous citons ci-dessous [5]

Les variations intra-classe telles que les variations d'illumination, de pose ou d'expression détériorent les performances d'un système de reconnaissance. Le visage d'une personne X peut ressembler plus à celui d'une personne Y qu'à lui-même si l'on change les conditions d'acquisition. Ces effets sont surtout notables dans le cas des images tirées des séquences vidéo où l'environnement d'acquisition est incontrôlable. Ainsi, il est nécessaire de trouver des méthodes efficaces permettant de minimiser les effets de ces différentes variations.

Les algorithmes de normalisation photométrique permettent d'atténuer l'effet d'illumination. Ils ont permis ainsi d'améliorer considérablement les performances de la reconnaissance faciale dans des cas spécifiques acquis en lumière visible [85]. Une seconde voie consiste à utiliser l'acquisition active. Elle permet aussi d'éliminer ces variations d'illuminations et ainsi d'améliorer les performances du système sans augmenter la complexité algorithmique. Cela permet une meilleure intégration de l'algorithme sur composant programmable et ainsi d'effectuer la reconnaissance en temps réel avec une consommation d'énergie raisonnable.

Nous allons présenter dans ce qui suit quelques types de ces variations intra-classe [2].

4.1 Changement d'illumination

L'apparence d'un visage dans une image varie énormément en fonction de l'illumination de la scène lors de la prise de vue. Les variations d'éclairage rendent la tâche de reconnaissance de

visage très difficile. En effet, le changement d'apparence d'un visage du à l'illumination, se révèle parfois plus critique que la différence physique entre différent individus, et peut entraîner une mauvaise classification des images d'entrée. Ceci a été expérimentalement observé dans Adini et al [10] où les auteurs ont utilisé une base de données de 25 individus.[5] L'identification de visage dans un environnement non contrôlé reste donc un domaine de recherche ouvert. Les évaluations FRVT [Bla01] ont révélé que le problème de variation d'illumination constitue un défi majeur dans la reconnaissance faciale.



Figure 6:Exemple de variation d'éclairage

4.2 Variation de pose Le taux de reconnaissance de visage baisse considérablement quand des variations de pose sont présentes dans les images. Cette difficulté a été démontrée par des tests d'évaluation élaborés sur les bases FERET et FRVT [11] [12]. La variation de pose est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Quand le visage est de profil dans le plan image (orientation $< 30^\circ$), il peut être normalisé en détectant au moins deux traits faciaux (passant par les yeux). Cependant, lorsque la rotation est supérieure à 30° , la normalisation géométrique n'est plus possible (voir figure 6, figure 7).



Figure 7:Exemples de variation de poses

4.3 Expressions faciales

Un autre facteur qui affecte l'apparence du visage est l'expression faciale (voir figure 8).

La déformation du visage qui est due aux expressions faciales est localisée principalement sur la partie inférieure du visage. L'information faciale se situant dans la partie supérieure du visage reste quasi invariable. Elle est généralement suffisante pour effectuer une identification. Toutefois, étant donné que l'expression faciale modifie l'aspect du visage, elle entraîne forcément une diminution du taux de reconnaissance. L'identification de visage avec expression faciale est un problème difficile qui est toujours d'actualité et qui reste non résolu. L'information temporelle

fournit une connaissance additionnelle significative qui peut être utilisée pour résoudre ce problème [13].



Figure 8:Exemples de variation d'expressions.

4.4 Occultations partielles

Le visage peut être partiellement masqué par des objets dans la scène, ou par des accessoires tels que lunettes, écharpe... Dans le contexte de la biométrie, les systèmes proposés doivent être non intrusifs c'est-à-dire qu'il ne doit pas compter sur la coopération avec d'autres systèmes. Par conséquent, il est important de savoir reconnaître des visages partiellement occultés. Gross et al [13] ont étudié l'impact du port des lunettes du soleil, et du cache-nez occultant la partie inférieure du visage sur la reconnaissance faciale. Ils ont utilisé la base de données AR [14]. Leurs résultats expérimentaux semblent indiquer que, dans ces conditions, les performances des algorithmes de reconnaissance restent faibles.

5. Conclusion

La reconnaissance de visage est une technique biométrique très utilisée dans les applications de contrôle de frontières, dans la sécurité des établissements et des zones urbaines et dans l'identification des conducteurs. C'est une technique commune, populaire, simple et qui offre beaucoup d'avantages.

Chapitre 2 :

La détection de visages et l'extraction de caractéristiques

1. Résumé :

Dans ce chapitre, nous essayons d'expliquer, en premier, le principe de la détection de visages et les enjeux à prendre en considération et puis nous faisons un survol des techniques d'extraction de caractéristiques, tout en essayant de faire apparaître leurs avantages et inconvénients.

2. La détection de visages

La détection de visages est la première étape dans le processus de reconnaissance faciale. Son efficacité a une influence directe sur les performances du système de reconnaissance de visages. Il existe plusieurs méthodes pour la détection de visages, certaines utilisent la couleur de la peau, la forme de la tête, l'apparence faciale, alors que d'autres combinent plusieurs de ces caractéristiques.

Les méthodes de détection de visages peuvent être subdivisées en quatre catégories [15] :

2.1 Approches basées sur les connaissances acquises

Ces méthodes s'intéressent aux parties caractéristiques du visage comme le nez, labouche et les yeux. Elles sont basées sur la définition des règles strictes à partir des rapports entre les caractéristiques faciales. Ces méthodes sont conçues principalement pour la localisation de visage. Dans [16], Kotropoulous et Pitas utilisent une méthode à base dérègles. Tout d'abord, les caractéristiques du visage sont localisées à l'aide de la méthode de projection proposée par Kanade[17] pour détecter les contours d'un visage. Soit $I(x,y)$ l'intensité lumineuse du pixel (x,y) de l'image $I(mn)$, les projections horizontale et verticale de cette image sont définies par

$$HI(x) = \sum_{y=1}^n I(x,y) \text{ et } VI(x) = \sum_{x=1}^m I(x,y) .$$

Le profil horizontal de l'image originale est calculé en premier. Les deux minimas locaux sont déterminés, ils correspondent aux bords gauche et droit du visage. Ensuite, le profil vertical t à son tour calculé. Les minima locaux de ce profil vertical correspondent aux positions de la bouche, du nez et des yeux. L'inconvénient de cette méthode est qu'elle n'arrive pas à détecter le visage lorsque ce dernier se trouve sur un arrière-plan complexe.

Yang and Huang [18] quant à eux, ont étudié les évolutions des caractéristiques du visage en fonction de la résolution. Ces méthodes réduisent progressivement la résolution de l'image du visage, par sous-échantillonnage ou par moyenne, pour faire disparaître les traits macroscopiques du visage. Ainsi, pour une résolution faible, la région du visage de vient uniforme. Yang et Huang se sont basés sur cette observation pour proposer une méthode hiérarchique de détection de visages. En commençant par les images à faible résolution, un ensemble de visages candidats est déterminé à l'aide d'un ensemble de règles permettant de rechercher les régions uniformes dans une image. Les visages candidats sont ensuite vérifiés en cherchant l'existence de traits faciaux proéminents grâce au calcul des minimas locaux à des résolutions supérieures. Une caractéristique intéressante de cette technique « descendante » de la recherche de zone d'intérêt (informations globales vers des informations détaillées) est de réduire le temps de calcul

nécessaire par l'utilisation d'images sous échantillonnées. Malheureusement, cette technique occasionne de nombreuses fausses détections et un taux faible de détection.

2.2 Approches basées sur le « Template-matching »

Les templates peuvent être définis soit "manuellement", soit paramétrés à l'aide de fonctions. L'idée est de calculer la corrélation entre l'image candidate et le template. Ces méthodes rencontrent aussi quelques problèmes de robustesse liés aux variations de lumière, d'échelle, etc. Sinha [19] [20] utilise un ensemble d'invariants décrivant le modèle du visage. Afin de déterminer les invariants aux changements de luminosité permettant de caractériser les différentes parties du visage (telles que les yeux, les joues, et le front); cet algorithme calcule ainsi les rapports de luminance entre les régions du visage retient les directions de ces rapports (par exemple, la région 1 est celle la plus claire ou la plus sombre que la région 2). La figure (9) montre un modèle prédéfini correspondant à 23 relations. Ces relations prédéfinies sont classifiées en 11 relations essentielles (flèches) et 12 relations de confirmations (gris). Chaque flèche représente une relation entre deux régions.

Une relation est vérifiée si le rapport entre les deux régions qui lui correspond dépasse un seuil. visage est localisé si le nombre de relations essentielles et de confirmation dépasse un seuil.

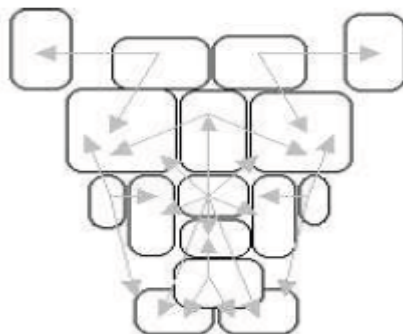


Figure 9:Modèle de visage composé de 16 régions (les rectangles) associées à 23 relations (flèches).

Yuille et al. [21] ont utilisé un template déformable pour modéliser les caractéristiques faciales. Ce template adapte un modèle élastique, connu a priori, aux caractéristiques faciales (e.g., yeux). Dans cette approche les caractéristiques faciales sont décrites par des templates paramétriques. Une fonction d'énergie est définie pour lier les contours, les sommets et les vallées dans l'image d'entrée aux paramètres correspondants dans le template. Le meilleur ajustement du modèle élastique est trouvé en minimisant la fonction d'énergie des paramètres. Bien que leurs résultats expérimentaux démontrent une bonne performance pour le suivi de caractéristiques non rigides, un inconvénient de cette approche est que le template déformable doit être initialisé dans la proximité de l'objet d'intérêt.

Pour détecter les caractéristiques d'un visage, Brunelli et Poggio[22] ont utilisé, pour chaque région extraite, un détecteur approprié. Ils se sont aussi inspirés de la méthode de Kanade[17]. Pour les régions yeux, nez et bouche, ils utilisent la direction du gradient vertical et horizontal. La bouche et le nez sont localisés en utilisant des stratégies similaires. La position verticale est déterminée grâce aux anthropométries standards. D'abord, une estimation fine de leur position réelle est obtenue en recherchant les pics de la projection horizontale du gradient vertical pour le nez, et les vallées de la projection horizontale de l'intensité pour la bouche.

La position des sourcils et leur épaisseur peuvent être trouvées par une analyse similaire. Cette recherche est limitée à la fenêtre d'intérêt, juste au-dessus des yeux, et les sourcils sont trouvés en utilisant la carte du gradient vertical. Le détecteur du sourcil cherche les paires de pics du gradient ayant des directions opposées.

2.3 Approches basées sur l'apparence

Ces approches appliquent généralement des techniques d'apprentissage automatique. Ainsi, les modèles sont appris à partir d'un ensemble d'images représentatives de la variabilité de l'aspect facial. Ces modèles sont alors employés pour la détection. L'idée principale de ces méthodes est de considérer le problème de la détection de visage comme problème de classification (visage, non-visage). L'une de ces approches les plus célèbres pour la détection de visage est l'*Eigenface*[23], elle consiste à projeter l'image dans un espace et à calculer la distance euclidienne entre l'image et sa projection. En effet, en codant l'image dans un espace, on dégrade l'information contenue dans l'image, puis on calcule la perte d'information entre l'image et sa projection. Si cette perte d'information est grande (évaluée à partir de la distance, que l'on compare à un seuil fixé a priori), l'image n'est pas correctement représentée dans l'espace : elle ne contient pas de visage. Cette méthode donne des résultats assez encourageants, mais le temps de calcul est très important.

2.4 Approches basées sur des caractéristiques invariantes

Ces approches sont utilisées principalement pour la localisation de visage. Les algorithmes développés visent à trouver les caractéristiques structurelles existantes même si la pose, le point de vue, ou les conditions d'éclairage changent. Puis ils emploient ces caractéristiques invariables pour localiser les visages. Nous pouvons citer deux familles de méthodes appartenant à cette approche : Les méthodes basées sur la couleur de la peau qu'on détaillera dans la section suivante, et les méthodes basées sur les caractéristiques de visage; elles consistent à localiser les cinq caractéristiques (deux yeux, deux narines, et la jonction nez /lèvre) pour décrire un visage typique.

2.4.1 Détection de visages basée sur l'analyse de la couleur de la peau :

Les méthodes de détection basées sur l'analyse de la couleur de la peau sont des méthodes efficaces et rapides. Elles réduisent l'espace de recherche de la région visage dans l'image.

De plus, la couleur de la peau est une information robuste face aux rotations, aux changements d'échelle, et aux occultations partielles. Plusieurs espaces couleur peuvent être utilisés pour détecter, de l'image, les pixels ayant la couleur de la peau.

3 Techniques de reconnaissance de visage

Plusieurs méthodes d'identification de visages ont été proposées durant les vingt dernières années. L'identification de visage est un axe de recherche ouvert attirant des chercheurs venants de différentes disciplines: psychologie, reconnaissance de formes, réseaux de neurones, vision artificielle et infographie. Les connaissances de ces études sont importantes car elles permettent le développement de nouvelles approches. Le but ultime de la reconnaissance faciale est de rivaliser, voir même dépasser, les capacités humaines de reconnaissance faciale. Les résultats fondamentaux de ces études ont été synthétisés dans un papier publié en 2006 par [24], et peuvent être résumés comme suit :

- Les humains peuvent reconnaître des visages familiers dans des images de faible résolution.
- La capacité de tolérer les dégradations d'images augmente avec la familiarité
- Les informations à haute fréquence telles que les contours, sont insuffisantes pour obtenir une reconnaissance faciale performante
- Les caractéristiques faciales sont traitées de manière holistique
- Parmi les différentes caractéristiques faciales, les sourcils sont les plus importants pour la reconnaissance
- La forme du visage est généralement codée de manière caricaturale
- La pigmentation du visage est aussi importante que sa forme
- La couleur joue un rôle important précisément lorsque la forme est dégradée
- Les changements d'illumination influencent la capacité de généralisation
- Le mouvement de visage semble faciliter la reconnaissance de manière conséquente
- Identité faciale et expressions sont traitées par différents systèmes

C'est ce qui explique l'abondance et la diversité de la littérature dans ce domaine. L'article de Tana et al. [25] donne un état de l'art récent et complet sur les dernières techniques développées. Dans cette section, nous présenterons les approches 2D de la reconnaissance de visage les plus connues. Ces dernières peuvent être subdivisées en trois méthode : les approches holistiques ou globales, les approches locales et les approches hybrides.

3.1 Approches holistiques ou globales

Ces méthodes identifient un visage en utilisant l'image entière de ce dernier comme entrée du système de reconnaissance. Chaque image de visage de dimension (n,m) est représentée par un vecteur simple de dimension $n \times m$, en concaténant les valeurs du niveau de gris de tous les pixels de l'image du visage. L'espace I contenant tous les vecteurs images de visages est appelé espace images. L'avantage de cette représentation est qu'elle préserve implicitement les informations de texture et de forme nécessaire pour la reconnaissance de visages. De plus, elle permet une meilleure capture de l'aspect global du visage par rapport aux représentations locales [26]. Toutefois, son inconvénient majeur réside dans la dimension très grande de l'espace image qu'elle nécessite [27, 28, 29], ce qui rend la classification très difficile.

Ainsi, une image 100×100 , par exemple, est représentée par un vecteur de dimension 10^4 [30] et comme le nombre d'images d'apprentissage pour chaque personne doit être au moins dix fois plus la dimension du vecteur [31], il faut 10^5 images par personne, un nombre assez exorbitant. En pratique, nous n'avons pas besoin d'un nombre énorme de photos pour développer un modèle robuste. Des techniques de réduction de dimension sont généralement employées pour réduire la taille de la base d'apprentissage. Une de ces techniques les plus utilisées pour l'identification du visage est l'Analyse en Composantes Principales (ACP). Une méthode très populaire, basée sur la technique ACP, est la méthode « eigenface »

Son principe est le suivant : étant donné un ensemble d'images de visages exemples, il s'agit tout d'abord de trouver les composantes principales de ces visages. Ceci revient à déterminer les vecteurs propres de la matrice de covariance formée par l'ensemble des images exemples. Chaque visage exemple peut alors être décrit par une combinaison linéaire de ces vecteurs propres. Pour construire la matrice de covariance, chaque image de visage est transformée en vecteur. Chaque élément du vecteur correspond à l'intensité lumineuse d'un pixel. Dans [33], les auteurs ont démontré que la matrice de covariance C peut s'écrire :

$$C = C_I + C_E \quad (1)$$

C'est-à-dire qu'elle est égale à la somme de la matrice de dispersion intra-personne C_I et la matrice de dispersion inter-personne C_E . Dans le cas d'un seul exemple d'apprentissage par personne, $C_I = 0$, et donc l'équation (1) se réduit à C_E .

L'eigenspace estimé à partir de la matrice C_E seulement n'est pas fiable, parce qu'il ne peut pas différencier, de manière efficace, l'erreur d'identification des autres erreurs dues à la transformation et au bruit [33]. Pour illustrer l'influence du nombre d'exemples d'apprentissage par personne sur les performances de la reconnaissance, les auteurs ont utilisé la base de données ORL comme base de test. La base de données ORL contient des images de 40 individus, chacun étant enregistré sous 10 vues différentes. Dans leur expérimentation, les auteurs ont fixé le

nombre de visages de test. Par contre, ils ont fait varier le nombre de visages d'apprentissage. Ainsi, pour chaque personne, ils ont utilisé la dernière image (Figure 10) pour le test et ont choisi aléatoirement les n premières images ($n \leq 9$) pour l'apprentissage.



Figure 10: Les dix vues d'une personne dans la base de données ORL.

Cette procédure a été répétée vingt fois. Les expérimentations ont montré que la performance de la méthode eigenface baisse avec la diminution du nombre d'exemples d'apprentissage de chaque personne. Dans le cas extrême, si seulement un exemple d'apprentissage par personne est considéré, le taux d'identification moyen de l'eigenface tombe en dessous de 65 %. Ce taux atteint 95 % quand on utilise neuf exemples d'apprentissage par personne.

Pendant la dernière décennie, plusieurs chercheurs ont développé diverses extensions de la méthode « eigenface », comme l'eigenface probabiliste [34], les méthodes basées sur l'Analyse Discriminante Linéaire (ADL) les Machines à Vecteur

Support (SVM), la méthode de la ligne caractéristique et La placion faces[35]. Ces approches possèdent des performances supérieures à celles données par la méthode eigenface de base. Cependant, ceci n'est plus vrai, si un seul exemple d'apprentissage par personne est considéré.

Ces méthodes peuvent être résumées comme suit [25] :

- les méthodes basées sur l'Analyse Discriminante Linéaire (ADL) déterminent les directions de projection les plus discriminantes dans l'eigenspace. Pour cela, elles maximisent les variations inter-visage par rapport aux variations intra-visage.

Cependant, si un seul exemple d'apprentissage par personne est utilisé, c'est-à-dire si les variations intra classe sont nulles, alors les performances de l'ADL deviennent faibles par rapport à celles données par l'eigenface [36]. Afin de remédier à ce problème, Zhao et al ont proposé de remplacer la matrice de dispersion intra-visage par une matrice constante. Ainsi, la méthode basée ADL se réduit alors à la méthode eigenface.

- L'approche Probabiliste [36] transforme le problème d'identification de visage en un problème de classification en deux classes. Elle évalue la probabilité de la différence entre une image de test et une image prototype appartenant aux classes intra-visage et inter-visage. Notons que la distribution intra-visage ne peut pas être évaluée dans le cas d'un exemple par personne, et la méthode se ramène aussi à la méthode eigenface.

- “pur suit Evolution”. et « La placionfaces » dépendent largement du nombre d'exemples d'apprentissage par personne. Dans le cas d'un exemple par personne, ces méthodes se ramènent aussi à la méthode « eigenface ».
- Les méthodes SVM et ligne de caractéristiques sont des méthodes de classification dans l'«eigenspace ». Si uniquement un exemple par personne est employé, alors aucune des deux méthodes ne marche.

Comme illustré ci-dessus, on ne peut pas s'attendre à obtenir de bonnes performances en généralisant la technique eigenface lorsque celle-ci utilise un seul exemple par personne. Cependant, cette méthode s'adapte très bien pour la classification de données de grande dimension.

3.2 Méthodes locales

Les méthodes locales utilisent les caractéristiques faciales locales pour la reconnaissance de visage. Elles sont relativement matures comparées aux méthodes holistiques

. Dans ces méthodes, le visage est représenté par un ensemble de vecteurs caractéristiques de dimensions faibles, plutôt que par un seul vecteur de grande dimension.

Nous pouvons classer les méthodes locales en deux catégories : les méthodes basées sur les caractéristiques locales : extractions et localisation des points caractéristiques, et les méthodes basées sur les apparences locales : partitions des images de visage en région caractéristiques.

3.2.1 Méthodes basées sur les caractéristiques locales

Les approches basées sur l'extraction de points caractéristiques peuvent être subdivisées en deux catégories : les approches géométriques et les approches basées sur les graphes.

3.2.1.1 Approches géométriques

Elles sont basées sur l'extraction de la position relative des éléments qui constituent le visage (tel que le nez, la bouche et les yeux). La plupart des approches géométriques utilisent des points d'intérêt (comme les coins de la bouche et des yeux). Au début des années 1990, Brunelli et Poggio[22] ont décrit un système de reconnaissance faciale qui extrait automatiquement 35 caractéristiques géométriques du visage. La similitude est calculée à l'aide du classificateur de Bayes. Un taux d'identification de 90 % sur une base de données de 47 visages a été rapporté par les auteurs. Le coût de stockage des techniques géométriques est très bas comparé à celui des autres techniques. Toutefois, les approches purement géométriques présentent quelques inconvénients, notamment :

- les caractéristiques géométriques sont généralement difficiles à extraire, sur tout dans des cas complexes : illumination variable, occultations, etc.

- les caractéristiques géométriques seules ne suffisent pas pour représenter un visage ,tandis que d'autres informations utiles comme les niveaux de gris de l'image ne sont pas du tout exploitées.

3.2.1.2 Approches basées sur les graphes

Plutôt que d'utiliser des méthodes purement géométriques, certains chercheurs ont choisi de représenter les caractéristiques locales du visage sous forme de graphes. Manjunath et al. [38] ont proposé une méthode de détection de caractéristiques locales du visage, basée sur la décomposition de Gabor en ondelettes [39]. La reconnaissance de visages est alors formulée comme un problème de mise en correspondance de graphes. L'efficacité de cette méthode a été validée sur un ensemble de données de visage de 86 visages, contenant des variations d'expression et de pose, Un taux de reconnaissance de 90% en moyenne a été rapporté démontrant la robustesse de cette approche.

Cependant, une fois le graphe topologique est construit, il ne peut pas être modifié. Or, les images de visage changent facilement d'apparence en raison des différentes variations (illumination, expression, pose, etc.), ainsi, un schéma de graphe topologique fixe n'est plus adéquat.

3.2.2 Méthodes basées sur l'apparence locale (Local appearance-based methods)

Ces techniques sont utilisées de manière modulaire pour les différentes régions faciales. Un modèle global est alors défini à partir de la combinaison des différents modèles locaux. Ainsi, les différentes régions faciales ne seront plus affectées de la même manière par les différentes sources de variabilité. Par exemple, le port de lunettes de soleil change considérablement l'aspect des yeux, tandis qu'un sourire affecte plus la région de la bouche. Deux paramètres sont utilisés pour définir les régions locales du visage : la forme (rectangle, ellipse) et la taille. Les caractéristiques des régions locales sont déterminées à partir d'une analyse des valeurs de niveau gris [40][41].

3.2.2.1 ACP et ADL modulaire

Dans Pentland et al ont introduit l'approche « Modular Eigenspace ». Les régions faciales utilisées sont : la totalité du visage, les yeux et le nez. L'ACP est appliqué sur chacune de ces régions et les résultats de classifications obtenus sont réunis. La bouche étant trop sensible à des changements d'expression faciale, sa prise en compte engendre une baisse du taux de reconnaissance. Cette approche peut être qualifiée d'hybride, puisqu'elle utilise à la fois des

caractéristiques globales et locales. Les auteurs ont montré qu'elle est plus efficace que les techniques globales ou strictement locales appliquées séparément.

Comme expliqué ci-dessus, les méthodes basées ADL échouent si un seul échantillon par classe est utilisé. Chen et al. ont proposé d'appliquer la méthode ADL sur un exemple de petite taille. Ils ont partitionné chaque image de visage en un ensemble d'images de même dimension (figure 11).

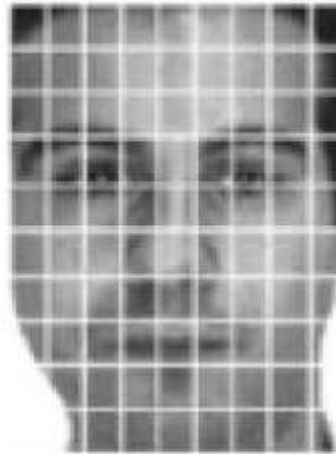


Figure 11: Image de visage partitionnée en image

Pour chaque classe, des échantillons composés de toutes les images partitionnées sont produits pour l'apprentissage. Enfin, la méthode ADLF (Analyse Discriminante Linéaire Fisher) est appliquée aux échantillons nouvellement produits. Cette méthode a été évaluée sur un sous-ensemble de données FERET contenant 200 personnes avec une image d'apprentissage par personne : un taux d'identification de 86.5 % a été atteint.

Dans Price et Gee ont introduit une technique modulaire basée sur une variante de l'ADL. Les régions sélectionnées sont : la région faciale dans son ensemble, une bande faciale (de même largeur que la région faciale) s'étalant du front jusqu'au-dessous du nez, et une bande faciale contenant les yeux. Les résultats expérimentaux montrent que cette approche est plus performante que les techniques *eigenfaces* et *fisherfaces*, elle est notamment robuste aux changements d'illumination, d'expression et d'occultation du visage.

3.2.2.2 Les approches probabilistes Locales modulaires

Martinez et al [42] présentent une approche probabiliste locale pour la reconnaissance de visages avec des occultations partielles et des variations d'expression. Ils utilisent un seul exemple d'apprentissage par classe. Chaque visage est divisé en six régions locales sous forme d'ellipses. Les régions ayant les mêmes positions pour chaque visage sont groupées séparément dans un sous-espace. Pour une représentation plus compacte et efficace, chaque sous-espace de visage est à nouveau transformé en un *eigenspace*, où la distribution est évaluée par la moyenne d'un

modèle Gaussien utilisant l'algorithme EM. Dans l'étape d'identification, les images de test sont aussi divisées en six régions locales et sont projetées respectivement dans l'*eigenspace* calculé.

Une approche probabiliste est utilisée pour mesurer la similitude d'un appariement donné. Les tests sur un jeu de 2600 images ont montré que l'approche probabiliste locale ne réduit pas la précision même pour 16 des visages partiellement occultés.

Bien qu'un certain nombre d'exemples soit synthétiquement produit comme décrit ci-dessus, les coûts de calcul et de stockage avec la procédure de production des exemples virtuels peuvent être très élevés (par exemple 6615 exemples par individu dans [42]) lorsque la base de visages est énorme.

Tan et al. [41] ont étendu l'approche probabiliste locale en proposant une méthode alternative pour représenter le sous-espace de visages en utilisant la carte SOM (Self-Organizing Map)

Plus spécifiquement, chaque image de visage I est d'abord divisée en M différents sous blocs locaux R_i , $i=1, \dots, M$ puis un réseau SOM est entraîné en utilisant tous les sous-blocs obtenus à partir de toutes les images d'apprentissage disponibles, sans tenir compte des classes. Chaque sous-bloc R_i de la même image de visage I est ensuite mappé avec les unités qui correspondent le mieux (BMU : best matching units) selon la stratégie suivante : le plus proche voisin, dans la SOM 2D de l'espace topologique, est défini comme un vecteur de position

$li = \{xi, yi\}$. Tous les vecteurs de position du même visage sont groupés en un ensemble, c'est-à-dire $I = \{U_i\}_{i=1}^M = \{x_i, y_i\}_{i=1}^M$, appelé la représentation "SOM-FACE" du visage (figure 12). Cette représentation possède plusieurs avantages. Des défauts possibles, comme le bruit dans l'image de visage originale, peuvent être éliminés durant le processus d'entraînement de la SOM. Cette méthode est une façon compacte de représenter le visage.

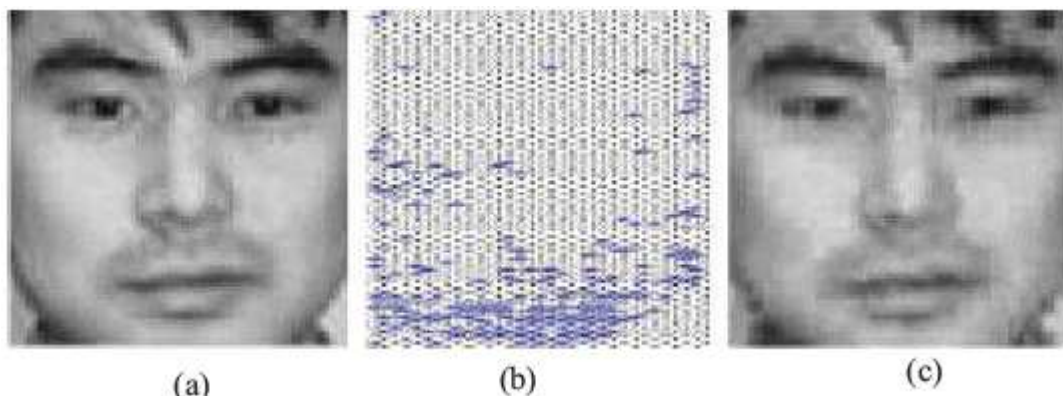


Figure 12: (a) une image du visage originale, (b) sa projection (SOM-Face) et (c) image reconstruite

À la différence d'autres méthodes comme l'ACP, cette représentation est intuitivement compréhensible. En effet, chaque élément d'un SOM-VISAGE a une signification physique donnée par le vecteur de poids stocké dans le nœud correspondant de la carte SOM, qui peut être interprété comme une pièce (patch) faciale locale dans l'espace d'entrée.

3.2.2.3 Les Modèles de Markov Caché modulaire

Les méthodes mentionnées ci-dessus ne considèrent pas explicitement le rapport entre les caractéristiques locales. Une autre façon intéressante d'incorporer des informations globales est la méthode basée sur les Modèles de Markov Caché (MMC). Plutôt que de considérer l'image de visage comme un graphe topologique statique où les caractéristiques locales représentent les nœuds, la méthode basée sur les MMC caractérise le modèle de visage comme un processus aléatoire dynamique avec un ensemble de paramètres.

Samaria et al. [43] ont illustré l'utilité des techniques MMC dans l'identification de visages. Dans leur méthode, un modèle de visage est divisé en cinq régions qui se chevauchent, incluant le front, les yeux, le nez, la bouche et le menton. La technique MMC est introduite en considérant chaque région comme un état caché d'un modèle MMC. Un modèle de visage est alors considéré comme une séquence d'observation constituée de cinq états, dont chacun peut être modélisé par une distribution gaussienne multi-variables, et les transitions probabilistes entre les états peuvent être apprises à partir des frontières entre les régions. Après que le MMC soit entraîné, un visage peut être reconnu en calculant la probabilité de sortie de son ordre d'observation. L'inconvénient de cette méthode est qu'elle est trop exigeante en termes d'échantillons d'apprentissage afin d'assurer la fiabilité de l'estimation des paramètres.

Le et al. [44] ont présenté une méthode pour rendre la technique MMC applicable dans le cas d'un exemple d'apprentissage. Deux facteurs ont contribué à la faisabilité et à l'efficacité de leur méthode. D'abord, ils ont produit une grande collection de vecteurs d'observation de chaque image, dans les deux directions verticale et horizontale, agrandissant ainsi l'ensemble d'apprentissage. Deuxièmement, l'Ondelette de Haar a été appliquée à l'image pour diminuer la dimension des vecteurs d'observation et améliorer la robustesse. Leurs résultats expérimentaux évalués sur la base de données de visage frontal.AR. Montrent que la méthode proposée surpasse les approches PCA, LDA et l'analyse de caractéristiques locales (LFA local feature analysis [45]).

2.3.3 Méthodes Hybrides

Les méthodes hybrides sont des approches qui combinent les caractéristiques holistiques et locales afin d'améliorer les performances de la reconnaissance de visages. En effet, les caractéristiques locales et les caractéristiques globales ont des propriétés tout à fait différentes.

On peut espérer pouvoir exploiter leur complémentarité pour améliorer la classification. Le tableau 1 récapitule qualitativement la différence entre les deux types de caractéristiques

Facteurs de variations	Caractéristiques locales	Caractéristiques globales
Illuminations [Hal99]	Très sensible	Sensible
Expressions [Mar02][Tan05]	Pas sensible	Sensible
Pose [Hei01]	Sensible	Très sensible
Bruit [Cos02]	Très sensible	Sensible
Occlusion [Mar02][Tan05]	Pas sensible	Très sensible

Tableau 1: Comparaison des méthodes basées sur les caractéristiques locales ou globales.

Nous pouvons voir que les caractéristiques locales et globales réagissant différemment sensibles aux facteurs de variation. Par exemple, les changements d'illumination peuvent avoir plus d'influence sur les caractéristiques locales, tandis que les changements d'expression ont plus d'impact sur les caractéristiques holistiques. Ainsi, les méthodes hybrides peuvent constituer une approche efficace pour réduire la complexité des classificateurs et améliorer leur capacité de généralisation. Malgré ces avantages, les travaux de recherche dans ce domaine restent relativement peu nombreux, probablement en raison des difficultés mentionnées ci-dessus, tandis que les méthodes hybrides classiques.

4. L'extraction de caractéristiques

Dans toutes les méthodes de reconnaissance faciale, le point le plus délicat concerne la définition et l'extraction des caractéristiques faciales les plus pertinentes, à savoir les caractéristiques qui représentent le mieux les informations portées par un visage. Le choix des caractéristiques locales présente plusieurs avantages par rapport à des caractéristiques globales. C'est pour cette raison que les systèmes les plus récents s'appuient sur des caractéristiques faciales locales. Ici on présentera deux types de caractéristiques locales les plus performantes dans le contexte de reconnaissance de visages, à savoir les ondelettes de Gabor et les indices Local Binary Patterns. Pour chaque type de caractéristique, il y a plusieurs façons de les utiliser. Nous présentons dans la suite leur principe de base.

4.1 Les ondelettes de Gabor

Le filtre de Gabor, défini par Dennis Gabor, est largement utilisé en traitement d'images car les ondelettes de Gabor présentent deux propriétés intéressantes : la localisation fréquentielle et la sélectivité en orientation. Les représentations en fréquence et en orientation du filtre Gabor s'apparentent à celles du système visuel humain [46] .

La représentation par ondelettes de Gabor dans la reconnaissance faciale est robuste aux changements causés par des variations d'éclairage ou par changements d'expressions. Lades et al. [47] sont les premiers qui ont employé les ondelettes de Gabor en reconnaissance faciale en utilisant Dynamic Link Architecture.

Un noyau du filtre de Gabor est le produit d'une onde complexe sinusoïdale avec une enveloppe gaussienne. Une ondelette de Gabor généralement utilisée dans la reconnaissance faciale est définie comme suit :

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{u,v}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (2)$$

Où $z = (x, y)$ est le point des coordonnées (x, y) dans le plan image. Les paramètres u et v définissent l'orientation et la fréquence des noyaux de Gabor. $\| \cdot \|$ est l'opérateur norme, et δ l'écart-type de l'enveloppe gaussienne. La représentation en ondelettes de Gabor d'une image résulte du produit de convolution de l'image avec une famille de noyaux de Gabor de fréquence et d'orientation différente comme définis par l'équation (2)[48].

La convolution de l'image I et d'un noyau de Gabor $\psi_{u,v}(z)$ est définie par :

$$G_{u,v}(z) = I(z) * \psi_{u,v}(z) \quad (3)$$

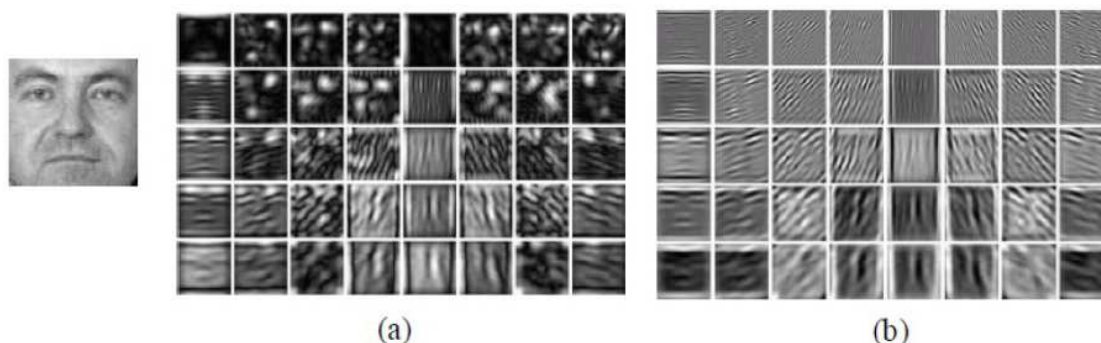


Figure 13: Exemple de représentation faciale en ondelettes de Gabor

Les réponses en l'amplitude (a) et en phase (b) d'un visage avec 40 noyaux de Gabor (5 échelles, 8 orientations). L'intérêt d'utiliser les filtres de Gabor pour extraire des caractéristiques faciales est qu'ils permettent de capturer les informations de visage dans des orientations et des résolutions différentes. De plus, ils sont robustes aux changements d'illumination, aux distorsions et aux variations d'échelle.

En effet, la convolution d'une image avec une banque de 40 noyaux de Gabor (5 échelles et 8 orientations) conduit à 40 cartes d'amplitude et 40 cartes de phase qui sont de même taille que l'image originale, comme illustrée sur la figure13 . Par conséquent, si on ne considère que la réponse en amplitude, chaque pixel est décrit par un vecteur de dimension 40.

Dans les méthodes les plus récentes telles que [49], les auteurs ont utilisé directement une image de caractéristiques contenant les 40 cartes d'amplitude comme entrée de reconnaissance. Dans [50] , Mellakh utilise les réponses en phase comme indice facial. Bien entendu, des méthodes de réduction de dimension ont été utilisées, telles que le sous-échantillonnage avec l'ACP et ses variantes. Plusieurs métriques ont été testées pour les caractéristiques basées sur les filtres de Gabor et celle qui est la plus souvent utilisée est la distance cosinus.

4.2 Local Binary Pattern

L'opérateur LBP a été proposé initialement par Ojala et al [51] dans le but de caractériser la texture d'une image. Le calcul de la valeur LBP consiste à seuiller les huit voisins directs de chaque pixel avec un seuil dont la valeur est le niveau de gris du pixel courant. Tous les voisins prendront alors une valeur 1 si leur valeur est supérieure ou égale au pixel courant et 0 si leur valeur est inférieure (figure14). Le code LBP du pixel courant est alors produit en concaténant ces 8 valeurs pour former un code binaire. On obtient donc, pour une image en niveaux de gris, une image contenant des valeurs LBP.

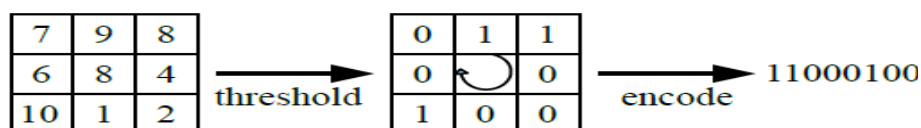


Figure 14:Opérateur LBP.

Le LBP a été étendu ultérieurement en utilisant des voisinages de tailles différentes. Dans ce cas, un cercle de rayon R autour du pixel central est considéré. Les valeurs des P points échantillonnés sur le bord de ce cercle sont prises et comparées avec la valeur du pixel central. Pour obtenir les valeurs des P points échantillonnés dans le voisinage pour tout rayon R, une

interpolation est nécessaire. On adopte la notation (P;R) pour définir le voisinage de P points de rayon R d'un pixel [52].

La figure 15(a) illustre trois voisinages pour des valeurs de R et P différentes.

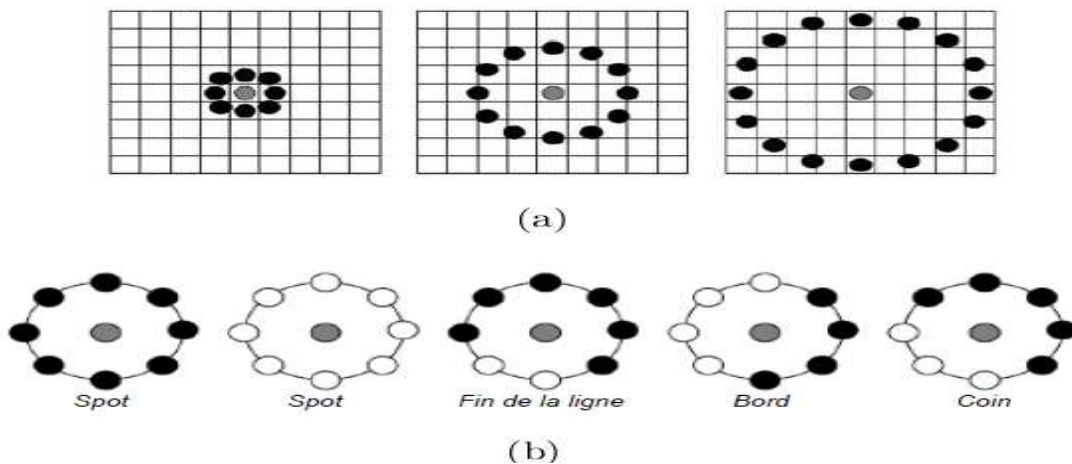


Figure 15: Trois voisinages pour des R et P différents ; (b) : Textures particulières détectées par LBPU2

Soient g_c le niveau de gris du pixel central, $g_p(1...P)$ les niveaux de gris de ses voisins, l'indice LBP du pixel courant est calculé comme suit:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=1}^P s(g_p - g_c) 2^{p-1} \quad (4)$$

où

$$s(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (5)$$

Où (x_c, y_c) sont les coordonnées du pixel courant, $LBP_{P,R}$ est le code LBP pour le rayon R et le nombre de voisins P.

LBPU pour la reconnaissance de visages :

Une fois le code LBP calculé pour tous les pixels de l'image, on calcule l'histogramme de cette image LBP pour former un vecteur de caractéristiques représentant l'image faciale. En général, afin d'incorporer plus informations spatiales au vecteur représentant le visage, l'image codée par l'opérateur LBP est tout d'abord divisée en petites régions et l'histogramme est construit pour

chaque région. Finalement, tous les histogrammes des régions seront concaténés afin de former un grand histogramme représentant l'image des caractéristiques faciales (figure 52).

L'efficacité du code LBP comme indice facial s'explique par le fait que le LBP permet de caractériser les détails fins d'un visage. Lorsque, seules les LBPs uniformes sont utilisés, tous les codes LBPs non-uniformes sont étiquetés avec une étiquette unique, alors que chacun des codes uniformes est regroupé dans un histogramme unique. Par exemple, lorsque $P = 8$, nous avons 58 codes uniformes mais l'histogramme est de dimension 59. De même manière $P = 6$ produit un histogramme de dimension 33 [52].

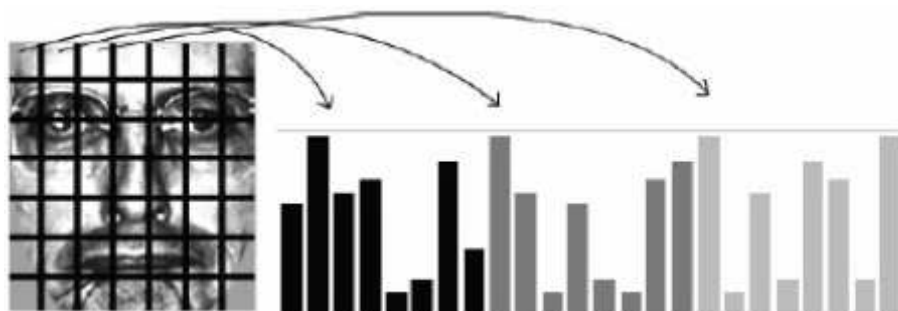


Figure 16: Représentation d'un visage par les histogrammes du code LBP [52].

Etant donnés deux histogrammes de LBP (H^1, H^2) de deux visages, l'étape suivante est d'utiliser une métrique pour calculer la similarité entre ces deux histogrammes. En testant les trois métriques X^2 , Histogramme intersection et Log-likelihood statistic, Ahonen et al [53] ont observé que la première métrique permet d'obtenir les meilleurs résultats :

$$\mathbf{x}_2(H^1, H^2) = \sum_i \frac{(H_i^1 - H_i^2)^2}{H_i^1 + H_i^2} \quad (6)$$

5. Conclusion

Dans ce chapitre, nous avons vu le principe de la détection de visages humains dans une image et celui de l'extraction de caractéristiques, ensuite nous avons présenté quelques approches à propos de ce sujet : Les méthodes basées sur les caractéristiques du visage (la géométrie du visage, la couleur de la peau, la connaissance généralisée,...) et Les méthodes globales (PCA, Approches probabilistes, réseaux de neurones...), ainsi que quelques méthodes hybrides.

La liste des méthodes présentées dans ce chapitre n'est pas exhaustive, mais constitue un aperçu de la diversité des approches qui existent pour faire la détection de visages.

Chapitre 3 :

Les Machines à Vecteur Support (SVM),les modèles de Markov Caché (MMC) et leur combinaison

1. Résumé

Dans ce chapitre, nous allons décrire en détail le principe de base du classificateur SVM et celui des modèles de markov caché. Ces deux classificateurs jugés efficaces ont été adoptés et ensuite combinés afin de profiter de leurs avantages pour réaliser un système efficace de la reconnaissance faciale. Nous présentons par la suite une brève description du principe des méthodes de la combinaison de classificateurs.

2 Présentation

Depuis quelques années, de nouvelles méthodes d'apprentissage ont été développées basées sur la théorie de l'apprentissage statistique (Statistical Learning Theory) de Vapnik et Chervonenkis. L'une de ces méthodes, appelée machines à vecteur de support ou SVM (Support Vector Machines) qui fut introduite par Vapnik en 1995, a permis de réaliser la classification (à deux classes ou plus) et aussi la régression. Le grand succès de cette méthode dans beaucoup de domaines se justifie essentiellement par les bases théoriques solides qui la soutiennent [27].

De telles méthodes permettent généralement de s'affranchir des contraintes statistiques sur les données étudiées comme la normalité de la distribution. De plus, elles sont non linéaires ce qui leur donne un pouvoir de généralisation supérieur par rapport aux méthodes de la régression classiques

3. Les Machines à Vecteur Support (SVMs)

Les SVMs sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de la classification binaire et de la régression. Les SVMs reposent sur deux idées principales : la notion de *la marge maximale* et la notion de *la fonction noyau*.

La marge maximale est employée pour les problèmes de la classification linéaire. Elle représente la distance entre la frontière de séparation et les échantillons d'apprentissage les plus proches. Ces derniers sont les *vecteurs supports*. *Les fonctions noyau* sont employées dans le cas des problèmes de la classification non-linéaire pour transformer l'espace de représentation des données d'entrées en un espace de plus grande dimension dans lequel il est probable qu'il existe de séparateurs linéaires [54].

Le classificateur SVM est un algorithme qui maximise la marge entre les classes du problème à résoudre et réduit au minimum l'erreur de classification. L'objectif de la marge maximale est de faire séparer deux classes par un hyperplan de telle sorte la distance par rapport aux vecteurs supports soit maximale [55].

Dans la tâche de classification, un SVM construit l'hyperplan optimal de séparation des attributs caractéristiques dans un espace de haute dimension. Le calcul de cet hyperplan est fondé sur la maximisation de la marge entre les exemples d'apprentissage les plus proches qui appartiennent à différentes classes.

2.1 Le cas linéairement séparable

L'hyperplan optimal de séparation est un classificateur de marge dont la sortie est donnée par [56], [55]:

$$f(x) = \text{sign}(x^T w + b) \quad (7)$$

Avec x un modèle d'entrée, le vecteur des poids W et le biais b sont ajustés par apprentissage en maximisant la marge $1 / \|w\|$ sous la contrainte de la bonne classification des N modèles d'apprentissage à l'extérieur de la marge [55]:

$$\min \frac{1}{2} \|w\|^2$$

$$\text{Sous la contrainte} \quad y_i(x_i^T w + b) \geq 1, \quad i = 1, \dots, N \quad (8)$$

Avec $y_i \in \{-1, 1\}$ représente l'étiquette d'un modèle d'apprentissage x_i . ($y_i=1$ pour la classe 1, et -1 pour la classe 2).

La solution de ce problème correspond au point extrême du lagrangien primal [55]:

$$L_P = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i [y_i(x_i^T w + b) - 1] \quad (9)$$

où α_i sont les multiplicateurs de Lagrange.

Ce problème mène à la maximisation du lagrangien dual par rapport aux α_i [56]:

$$\max L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i^T x_j) \quad (10)$$

$$\text{Sous les contraintes :} \quad \alpha_i \geq 0, \quad i = 1, \dots, N \quad \text{et} \quad \sum_{i=1}^N \alpha_i y_i = 0$$

La règle de séparation est donc:

$$f(x) = \text{sign} \left(\sum_{\text{vecteurs supports}} y_i \alpha_i (x_i^T x) + b \right) \quad (11)$$

Où " x_i " sont les vecteurs supports (SVs) lorsque les multiplicateurs de Lagrange (α_i) sont différent de zéro. Les SVs sont les modèles d'apprentissage qui se trouvent sur les frontières de la marge [56]. α_i et b sont déterminés en résolvant un problème quadratique [57].

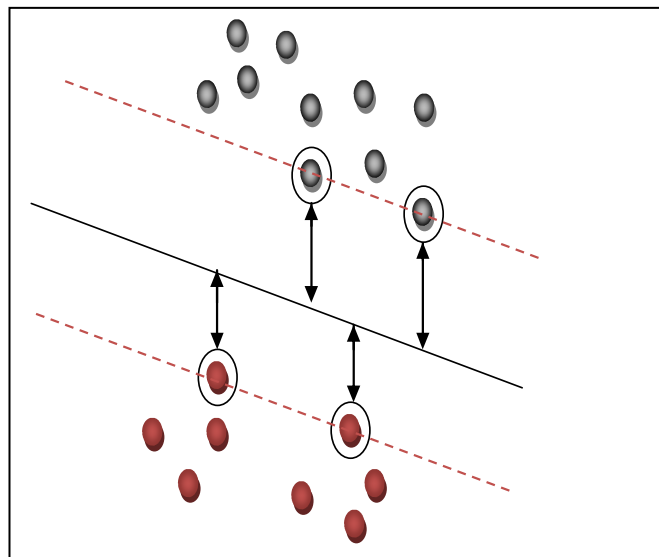


Figure 17: Discrimination linéaire

2.2 Le cas non-linéairement séparable

Dans le cas d'un espace non-linéaire. Chaque point " x " de l'espace d'entrée est transformé à un point $z = \Phi(x)$ dans un espace de plus grande dimension, appelé espace des attributs. La propriété principale de la fonction de transformation $\Phi(\cdot)$ est que le produit scalaire de deux points dans l'espace d'attributs $\Phi(x)$. $\Phi(x)$ peut-être réécrit par une fonction noyau : $K(x, y)$ [58].

Ainsi, le classificateur SVM linéaire peut être étendu au cas non-linéaire en remplaçant le produit scalaire entre les vecteurs d'entrée " x " et les vecteurs supports " x_i " par une fonction noyau $K(x_i, x)$ [57].

La règle de séparation devient donc sous la forme suivante [58]:

$$f(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b \right) \quad (12)$$

Les coefficients α_i et b sont les solutions d'un problème de programmation quadratique [57].

Les fonctions noyau les plus utilisées sont [59]:

Polynomial:

$$K(x, x') = \langle x, x' \rangle^d \quad (13)$$

Gaussienne:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (14)$$

Sigmoïde :

$$K(x, x') = \tanh(k\langle x, x' \rangle + \vartheta) \quad (15)$$

Où $k > 0$ et $\vartheta < 0$ sont des paramètres

2.3 Classification Multi-classes

Bien que la méthode SVM soit adaptée aux problèmes de deux classes, elle peut être facilement transformée en un outil très efficace pour la classification N-classes. Il existe deux méthodes pour résoudre un problème multi-classe avec des classificateurs SVMs [56] :

1. **Stratégie un contre tout :** Dans cette stratégie, un classificateur est construit pour chaque classe et chargé de la séparation de cette classe des autres classes restantes. L'idée consiste tout simplement à transformer le problème du k classes en k classificateurs binaires. La classification finale est donnée par le classificateur qui répond le mieux.
2. **Stratégie un contre un :** Dans cette stratégie, un classificateur est construit pour chaque paire des classes pour séparer les classes deux à deux. Le problème est remplacé par $N*(N-1)/2$ tâches de classification entre deux classes. La classification finale est donnée par le vote majoritaire ou un graphe acyclique de décision.

Une autre stratégie, pour la reconnaissance de n chiffres, est d'employer n-classe SVM au lieu des classificateurs binaires avec la stratégie un contre tout résolvant de ce fait un problème d'optimisation contraint [56]. Mais cette méthode n'est pas très populaire dans les applications de la reconnaissance de chiffres et n'apporte pas de meilleures performances que les autres classificateurs [60]. Une description détaillée des SVMs peut être trouvée dans [61].

3 Pourquoi combiner SVM avec une technique d'extraction de paramètres

Les performances d'un système de détection de visages basés sur les machines à vecteurs support (le temps de détection et la précision) sont en relation direct avec les données utilisées dans les images d'apprentissage.

La dimension des données à un effet direct sur le temps de calcul des support-vecteur et sur le temps de détection. De même que la corrélation de ces dernières.

Théoriquement si les images d'apprentissage sont indépendantes entre elles, la tâche du classificateur sera plus robuste et plus simple. Ce qui laisse à croire que le classificateur est plus efficace. Dans ce concept nous définissons la notion d'extraction de paramètres qui vise à extraire l'information utile des images d'apprentissage. L'information utile est de moindre dimension que l'information originale et ne contient pas de redondances. On sait que les images naturelles sont fortement corrélées (l'existence de régions entières où des pixels ayant la même valeur), alors on constate tous l'intérêt de l'extraction de paramètres. Plusieurs méthodes entrent dans le concept d'extraction de paramètres, elles se basent souvent sur le concept de transformation orthogonale telle que lePCA ,la DCT ou la transformée de Fourier et bien d'autres techniques.

4 L'analyse en composantes principales (PCA)

L'analyse en composantes principales fut introduite en 1991 par MA. Turk et MP.

Pentland[62]dans le domaine de la biométrie. Ils ont conçus le premier système de reconnaissance de visages basé sur PCA, mais l'utilisation du PCA comme moyen de réduction de la dimension des données et pour la dé corrélacion datait déjà des années 70.

4.1 Principe générale

Le but du PCA aussi connue sous le nom de transformé de Karhunen-Loève(KLT) est de construire un sous espace plus approprié pour représenter les images d'apprentissage. La méthode cherche les axes orthogonaux pour lesquels la variance des images exemples projetées soit maximale. Ces axes ont la particularité de bien représenter les images d'apprentissage, mais ils ne permettent pas forcement de les séparer.

La figure suivante illustre une projection suivant le principe du PCA :

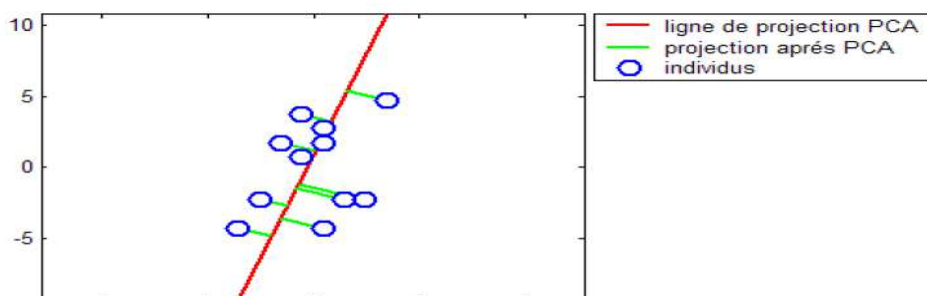


Figure 18:projection suivant le principe de la PCA.

La PCA cherche donc la projection la plus précise dans un sous espace de dimension inférieur.

4.2 Étapes à suivre pour effectuer une analyse en composantes principales

Nous supposons que nous avons un échantillon de données $D=[x_1, x_2, \dots, x_n]$. Où chaque x_i est un vecteur dimensionnel, formant ainsi une matrice X .

$$X = \begin{pmatrix} x_{11} & \dots & \dots & x_{1d} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & \dots & \dots & x_{nd} \end{pmatrix}$$

Notre but est d'avoir une matrice Y qui est une représentation de la matrice X en projetant l'ensemble des x_i par de nouvelles composantes y_i de dimension $p \leq d$ de tel façon que l'information soit préservée de manière optimale.

$$Y = \begin{pmatrix} y_{11} & \dots & \dots & y_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ y_{n1} & \dots & \dots & y_{np} \end{pmatrix}$$

Pour effectuer une analyse en composantes principales nous devrions procéder comme suit :

- Calculer le vecteur moyen :

$$\mu = \frac{1}{n} \sum_{i=1}^n X_i \quad (16)$$

- Soustraire aux données le vecteur moyen :

$$V_i = X_i - \mu \quad (17)$$

- Calculer la matrice variance covariance :

$$\text{COV} = \sum_{i=1}^n V_i V_i^T \quad (18)$$

- Calculer les valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ ainsi que les vecteurs propres qui lui sont associés $E = (e_1, e_2, \dots, e_n)$ est la matrice des vecteurs propres. Les vecteurs propres doivent être mis sous ordre décroissant ($\lambda_1 \geq \lambda_2 \geq \lambda_3, \dots, \lambda_n$) ; E est construite tel que e_i est le vecteur propre associé à la valeur propre λ_i .

- Calculer les vecteurs (y_1, y_2, \dots, y_p) qui représentent une projection des x_i mais dans un nouvel espace, l'espace propre de dimension $p \leq d$ par la formule suivante :

$$V_i = E^T V_i \quad (19)$$

4.3 Caractéristiques de PCA

- **Compactage de l'énergie** : Le PCA fournit le meilleur compactage d'énergie de toutes les transformations orthogonales, par rapport à celui offert par le DCT.
- **Décorrélation** : Le PCA fournit les meilleurs coefficients de décorrélation de toutes les transformations orthogonales.
- **Dépendance aux données** : Les performances obtenues par le PCA en matière de décorrélation et de compactage de l'énergie sont bien meilleures que celles du DCT. Cependant, le PCA dépend fortement des données. Il n'a pas de paramètres pré-calculés comme ceux du DCT ce qui démontre le temps considérable pour calculer le PCA.

4.4 Classification basée sur PCA-SVM

Le schéma d'un classificateur basé SVM-PCA diffère du celui basé sur SVM. Contrairement à la méthode DCT-SVM, le DCT est appliqué sur chaque image. Dans le cas présent, le PCA est appliquée sur la matrice modélisant l'ensemble des données d'apprentissage ou chaque donnée est représentée sous sa forme vectorielle. Le but est d'obtenir une indépendance entre les images de l'apprentissage et non entre les pixels d'une même image. Concernant les performances du système basé sur la méthode PCA-SVM, nous constatons le temps énorme du calcul que prend l'apprentissage, aussi le test est moins rapide par rapport à celui effectué par le SVM-DCT.

5. Les Modèles de Markov Cachés (MMC)

Les modèles de Markov sont des modèles mathématiques de processus stochastiques qui produisent des séquences de résultats aléatoires selon certaines probabilités. Un modèle de Markov caché est un modèle dans lequel on observe une séquence des émissions sans savoir la séquence des états qui a produit ces émissions

5.1 Les éléments d'un MMC

Un MMC est défini par un quintuplet $A = \{\Sigma, \Omega, \Pi, A, B\}$ où [64]:

L'alphabet $\Sigma = \{s_1, \dots, s_m\}$ décrit les états de la chaîne de Markov

L'alphabet $\Omega = \{o_1, \dots, o_k\}$ sont des symboles émis par les s_i pour un MMC discret

Les probabilités initiales des états $\Pi = \{\pi_i = P(s_i)\}$

La matrice des probabilités de transitions entre les états $A = \{a_{ij} = P(s_j/s_i)\}$

Les probabilités d'émission $B = \{b_i(o_k) = P(o_k/s_i)\}$

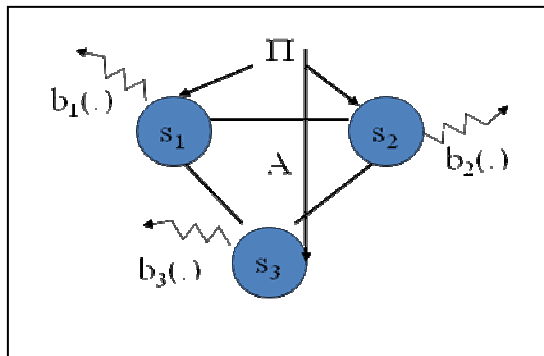


Figure 19: Une chaîne de Markov à 3 états avec les processus associés

Les MMCs permettent de répondre aux trois questions suivantes [63], [64] :

- On peut calculer la probabilité d'une séquence d'observation à l'aide de l'algorithme Forward-Backward
- Partant d'un ensemble d'observation O , on peut trouver la séquence la plus probable des états qui ont conduit à la génération de cette séquence à l'aide de l'algorithme Viterbi
- Partant d'un ensemble d'observation O , on peut régler les paramètres d'un model MMC afin de maximiser la vraisemblance de $P(O/A)$ (c-à-d l'apprentissage d'un MMC) en utilisant l'algorithme de Baum-Welch ou l'algorithme de Viterbi.

5.2 Apprentissage d'un MMC

Les méthodes d'apprentissage d'un MMC les plus utilisées sont : l'algorithme de Viterbi, et l'algorithme Baum-Welch [64].

5.2.1 L'algorithme de Viterbi

Idée générale :

L'algorithme de Viterbi est un algorithme de la programmation dynamique basé treillis qui exécute un MMC pour trouver la séquence la plus susceptible des états cachés, appelé le chemin de Viterbi [65]. On détermine, pour chaque sommet du treillis, le meilleur chemin (le chemin le plus probable)[64] menant à une séquence particulière.

Une présentation détaillée des MMCs peut être trouvée dans R. Rabiner,

5.2.2 L'algorithme Rabiner:

a. Initialisation :

Initialisation des probabilités d'état qui sont le produit des probabilités d'états cachés initiales avec les probabilités d'observation associées.

Pour $t=1$ et $1 \leq i \leq$ nombre d'états m ,

$$\alpha_1(i) = \pi_i * b_i(O_1) \quad (20)$$

$$\psi_1(i) = 0 \quad (21)$$

b. Récurrence :

Détermination du chemin le plus probable au prochain état en enregistrant, comment y arriver. Ceci est réalisé à travers le calcul des produits des probabilités de transition avec les probabilités maximales dérivées dans l'étape précédente.

Pour $t = 2, \dots, n$, et $1 \leq i \leq m$,

$$\alpha_{t+1}(i) = \max_{j=1 \dots m} [\alpha_t(j) * a_{ji}] * b_i(O_{t+1}) \quad (22)$$

$$\psi_t(i) = \operatorname{argmax}_{j \in \{1 \dots m\}} (\alpha_{t-1}(j) * a_{ji}) \quad (23)$$

c. Terminaison :

Ceci détermine l'état le plus probable du système ($t=T$).

$$s(n) = \operatorname{argmax}_i \alpha_T(i) \quad (24)$$

d. Retour en arrière :

Retour en arrière par le treillis en suivant le chemin le plus probable.

La séquence $(I) \dots s(T)$ est la séquence la plus probable des états cachés pour la séquence des observations traitée.

Pour $t = n-1, \dots, 1$,

$$s(t) = \psi_{t+1}(s(t+1)) \quad (25)$$

6. Extraction de caractéristiques par SVD

La décomposition en valeur singulière (SVD, *Singular Value Decomposition*) permet d'exposer la structure géométrique d'une matrice, afin de trouver la dimension qui possède la plus grande variance d'après les vecteurs de cette matrice. Pour la matrice A de dimensions $m \times n$, nous obtenons les matrices $U=[u_1, \dots, u_n]$ de dimensions $m \times m$ et $V=[v_1, \dots, v_n]$ (de transposée V^t) de dimensions $n \times n$ deux matrices orthogonales et D une matrice diagonale dont les valeurs $[d_1, \dots, d_n]$, organisées tel quel que $d_1 \geq d_2 \geq \dots \geq d_n$ sont les valeurs singulières de la matrice A

SVD de la matrice A : $A=U D V^t$

SVD de la matrice $A^t A$: $A^t A = V D^2 V^t$

Ces valeurs singulières pondèrent les vecteurs de V tel que V_1 soit plus important que V_2 et ainsi de suite. V_1 représente donc l'axe sur lequel les données de A ont la plus grande dispersion (projection sur l'axe ayant la plus grande p-norme).

7. Combinaison de classificateurs

Un classificateur est un outil de reconnaissance qui reçoit un vecteur des attributs d'entrées et produit un vecteur des classes de ces entrées. La combinaison de plusieurs classificateurs permet le plus souvent d'améliorer la qualité de la classification en termes d'exactitude. Cette combinaison est d'autant plus importante lorsque les classificateurs sont complémentaires. Lorsque les classificateurs sont identiques, il n'y aura aucune amélioration de l'exactitude ; il faut, donc créer des experts différents.

Le problème de la combinaison de classificateurs peut être considéré comme problème d'optimisation d'une fonction (de combinaison) qui accepte des vecteurs N -dimensionnels représentant les classes issues de M classificateurs et produit N points d'une classification finale en réduisant par exemple au minimum l'erreur de classification [66].

8. Types de sorties d'un classificateur

On peut distinguer trois types de sorties d'un classificateur [67]:

Soit un classificateur (e) qui reçoit une entrée (x) et fournit une sortie (C_i) tel que ($i=1..m$), et m représente le nombre de classes.

• Sortie de type classe

C'est le type le plus simple qui apporte le moins d'informations. Le classificateur donne une étiquette (C_i) indiquant la classe de l'entrée traitée.

$e_j(x) = C_i \ (i \in \{1..m\})$.

• Sortie de type rang

Le classificateur donne un ensemble de classes en précisant leurs faveurs par un rang.

$e_j(x) = [r_{j1}, r_{j2}, \dots, r_{jm}]$ où r_{ji} est le rang accordé à la classe (i) par le classificateur (j). Les rangs fournis par chaque classificateur peuvent être considérés comme des votes pondérés pour chaque classe.

• Sortie de type mesure

Cette sortie est la plus riche en information. Dans ce cas, pour chaque classe est associée une mesure qui indique sa préférence par une probabilité, une vraisemblance, une distance, un score ou une confiance,

$e_j(x)=[M_{j1}, M_{j2}, \dots, M_{jm}]$ où M_{ji} est la mesure fournie à la classe (i) par le classificateur (j).

Les mesures de performances d'un classificateur

Pour une certaine entrée, un classificateur peut produire l'une des réponses suivantes :

- **Un rejet** : si le classificateur n'a pas pu classifier l'entrée.
- **Une reconnaissance** : si le classificateur accorde la classe exacte à l'entrée traitée.
- **Une substitution** : si le classificateur associe une classe incorrecte à l'entrée.

$$\text{tauxderejet} = \frac{\text{nombredeformesrejetées}}{\text{nombretotaldeformes}}$$

$$\text{tauxdereconnaissance} = \frac{\text{nombredeformesreconnues}}{\text{nombretotaldeformes}}$$

$$\text{tauxdesubstitution} = \frac{\text{nombredeformesmalesreconnues}}{\text{nombretotaldeformes}}$$

9 .Les méthodes de combinaison des classificateurs

Il existe trois architectures de la combinaison de classificateurs que sont [37] :

- Séquentielle ou série
- Parallèle
- Hybride (mixte séquentielle et parallèle)

9.1 Méthodes de vote

Dans cette méthode, la sortie de chaque classificateur est considérée comme un vote (acceptation/rejet). La plupart des méthodes de vote utilisent un seuil qui représente la proportion des classificateurs devant voter la même classe pour qu'elle soit la décision finale. La classe qui a reçu un nombre de vote (acceptation) supérieur à ce seuil est considérée la décision finale sinon

rejet. Par exemple, dans la méthode de vote le plus simple (*le vote à la pluralité*), la classe qui a reçu le plus grand nombre de vote est considérée la classe finale sinon rejet [31].

Si on suppose que les sorties de L classificateurs sont des vecteurs c - dimensionnels d'étiquettes de classes (le nombre de classes =c) :

$$[e_{i,1}, \dots, e_{i,c}]^T \in \{0,1\}^c,$$

$i=1, \dots, L$ où $e_{i,j}= 1$ si le classificateur « e_i » classe l'entrée « x » dans sa classe C_i et 0 autrement.

$e_{i,j}$ est la $j^{\text{ème}}$ sortie du classificateur i pour l'entrée x ($j=1, \dots, c$).

Les méthodes de vote peuvent être dérivées de la règle, avec seuil, suivante [31]:

$$e(x) = \begin{cases} C_i \text{ si } \sum_{j=1}^L e_{i,j} = \max_{i=1}^M \sum_{j=1}^L e_{i,j} \geq \lambda.L \\ \text{sinon rejet} \end{cases} \quad (9)$$

L est le nombre de classificateurs, M le nombre de classes et λ un seuil de décision.

Majorité unanime :

Si $\lambda = 1$, tous les classificateurs doivent voter la même classe pour qu'elle soit la décision finale sinon la décision finale est le rejet. Cette méthode est appelée *majorité unanime* présente l'inconvénient majeur de produire un taux de reconnaissance assez faible [31], [37].

Majorité simple :

Si $\lambda \geq 0,5$, plus de la moitié des classificateurs doivent voter la même classe pour qu'elle soit la décision finale. Il s'agit de la méthode de vote par majorité. Dans un cas de conflit, des informations supplémentaires comme les performances individuelles des classificateurs sont utilisées.

Vote à la pluralité :

Si $\lambda = 0$, la classe qui a reçu le plus grand nombre de vote est considérée la classe finale sinon rejet. Cette règle s'appelle le vote à la pluralité [31], [37].

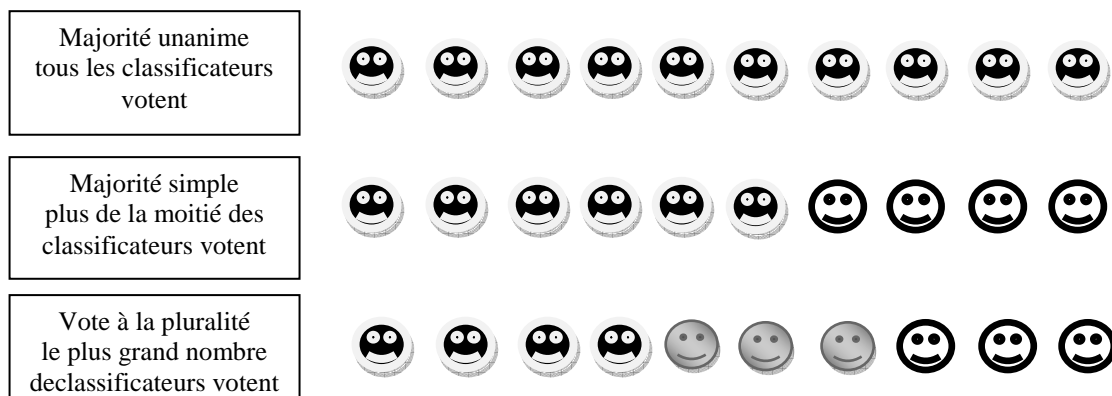


Figure 20: Les méthodes de vote : un anisme, majorité simple et vote à la pluralité, la décision finale est celle du groupe noir

Le vote pondéré :

Cette technique associe des poids w_j aux classificateurs de base e_j . Ces poids varient au fur et à mesure que les classificateurs s'entraînent suivant leur performance de prédiction [68].

$$g(x) = \begin{cases} c_i \text{ si } \sum_{j=1}^L w_j e_{i,j} = \max_{t=1}^M \sum_{j=1}^L w_j e_{t,j} \\ \text{sinon rejeter} \end{cases}$$

(10)

Par exemple, dans la règle de combinaison, les classificateurs ayant les taux les plus élevés de la reconnaissance sont beaucoup plus importants que les autres classificateurs ayant une faible performance, même si un certain exemple est souvent bien classifié par ces derniers [11].

Chapitre 4 :

Mise en œuvre et résultats expérimentaux

1. Résumé

Dans ce chapitre, il est question d'analyser le fonctionnement du système pour extraire les structures à implémenter et la façon dont ils devront communiquer pour réaliser les tâches. On commence par donner les cas d'utilisations du système, et puis les résultats des tests effectués avec les deux classificateurs SVM et MMC aussi bien qu'avec leur combinaison à travers la méthode de vote.

2. Conception:

La conception consiste à représenter un système par un ensemble de diagrammes d'utilisation expliquant le fonctionnement du système par l'illustration des différents cas d'utilisation.

2.1. Identification des acteurs du Système:

- L'Administrateur:

C'est la personne qui gère le système et qui choisit la configuration la plus efficace pour le rendre robuste et performant.

- Il prend à sa charge la gestion de la base de données ; il enregistre les nouvelles personnes et choisit les poses qui assurent un apprentissage efficace pour le système.

- Il met à jour les informations des personnes et supprime ceux qui quittent le système.

- Il teste également les paramètres de chaque méthode et observe les avantages et les inconvénients de chacune, pour choisir la meilleure configuration.

- La Personne teste:

C'est un individu qu'on veut identifier pour avoir des informations sur lui ; ou dans un autre cas une personne qui demande une autorisation d'accès (vérification).

2.2. Identification des cas d'utilisation du Système:

2.2.1. Enregistrement :

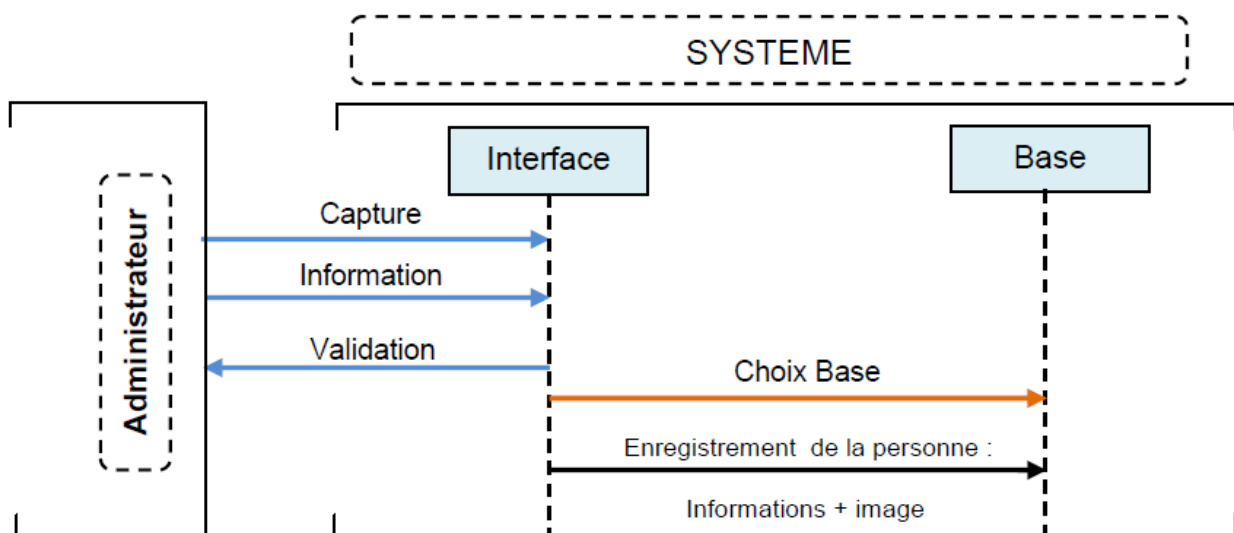


Figure 21: diagramme de séquence pour l'enregistrement d'une personne dans la base.

2.2.2. Vérification :

C'est le cas d'une personne qui demande une autorisation d'accès (par exemple : accès à un laboratoire). Pour ce type d'utilisation la personne en question doit proclamer son identité et doit également coopérer pour que le système prend une image de son visage pour vérifier : est ce que c'est la bonne personne ou pas.

Le diagramme suivant illustre ce cas d'utilisation :

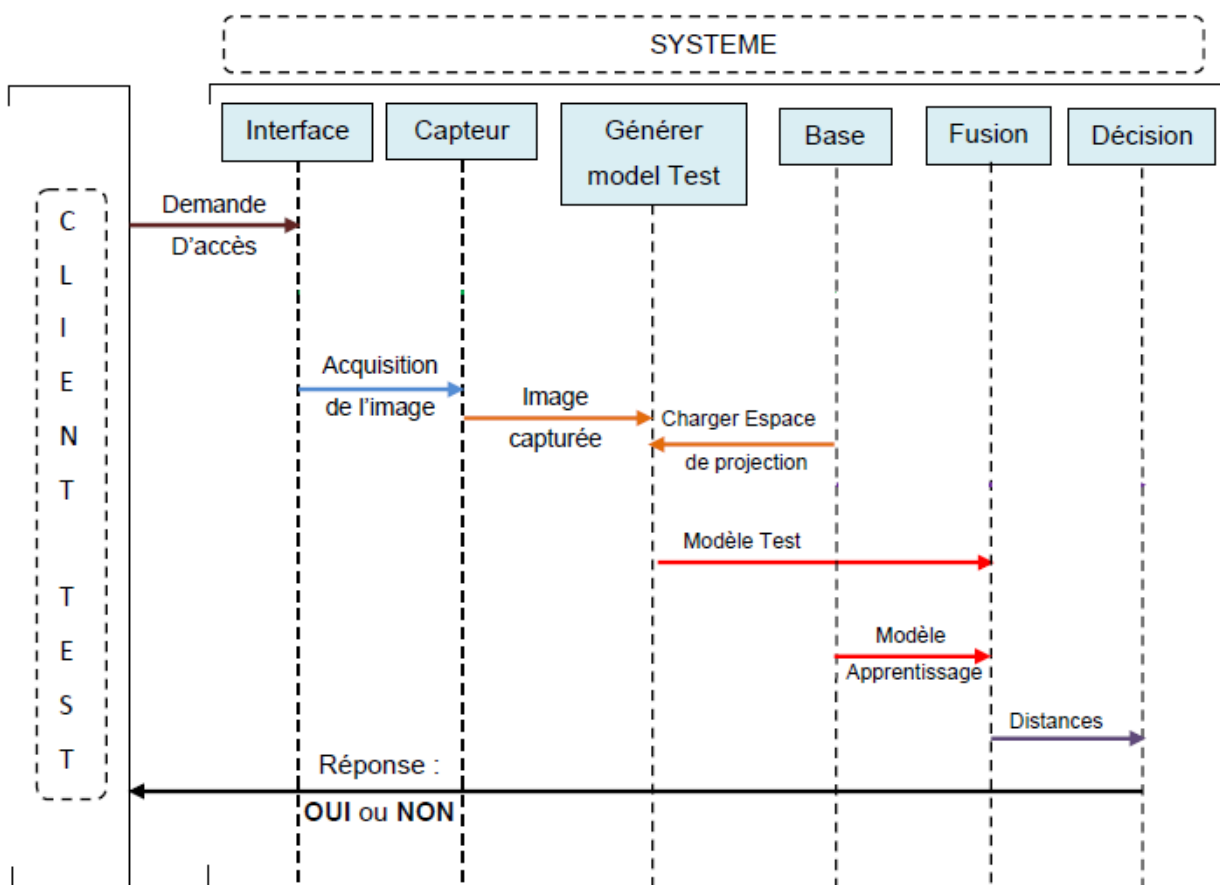


Figure 22:diagramme de séquence pour la vérification.

Un client envoie sa demande d'accès à une banque par exemple, le système capture son visage par caméra, calcule ses attributs caractéristiques par projection d'espace en utilisant PCA et SVD, le modèle de test (le vecteur d'attribut caractéristiques du visage) est ensuite envoyé à l'un des modèles d'apprentissage suivants : SVM, MMC ou leur combinaison par vote majoritaire, après l'étape d'apprentissage, le système décide s'il autorise l'accès ou non de la personne concernée suivant si elle est reconnue ou pas et ceci basé sur un calcul de similarité à travers le classificateur adopté.

2.2.3. Identification :

C'est le cas où on veut avoir plus d'information sur les personnes (par exemple : la Télésurveillance). Pour ce type d'utilisation le système peut être contrôlé ou pas, mais dans les deux cas le système n'exige pas une coopération de la personne en question, car il se charge de capter la personne et prendre une image du visage pour l'identifier.

Le diagramme suivant illustre ce cas d'utilisation :

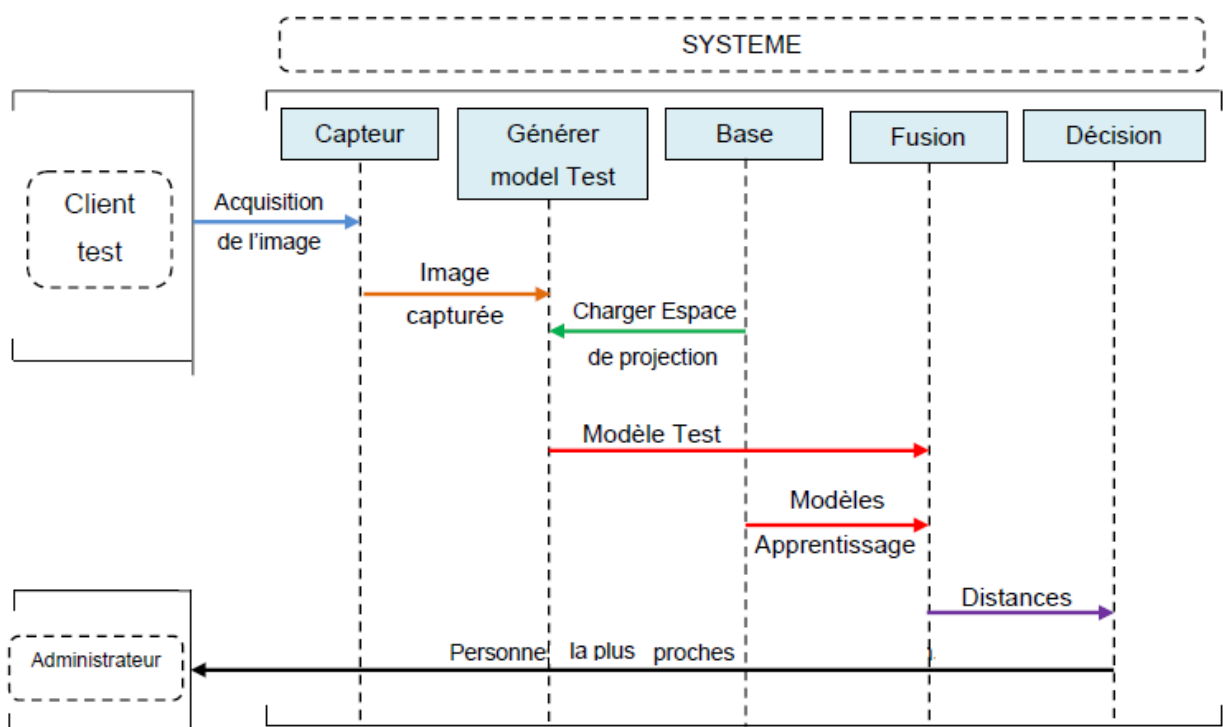


Figure 23:diagramme de séquence pour identification.

Pour l'identification on procède presque de la même manière que celle de la vérification, le système capture le visage de la personne à identifier par camera, calcul ses attributs caractéristiques par projection de l'espace de la base en utilisant PCA et SVD, le vecteur d'attribut caractéristiques du visage à identifier est ensuite envoyé à l'un des classificateurs: SVM, MMC ou leur combinaison, après le système décide ,basé sur un calcul de similarité, la classe du visage à identifier et renvoi ses informations stockées précédemment dans la base telles que : le nom, prénom, adresse

2.2.4. Calculs de performances :

La performance du système, elle est mesurable en termes d'erreurs et de vitesse d'exécution.

Ici, on présentera les indicateurs de performances que l'administrateur doit observera fin de juger l'efficacité du système.

Pour évaluer le système en mode d'identification et en mode vérification, on calcule son taux de reconnaissance et son temps de réponse, il consiste à mesurer le temps d'exécution de l'apprentissage, de l'identification et de la vérification, pour chaque type de classificateur, afin de préciser le degré de rapidité d'un classificateur par rapport à un autre.

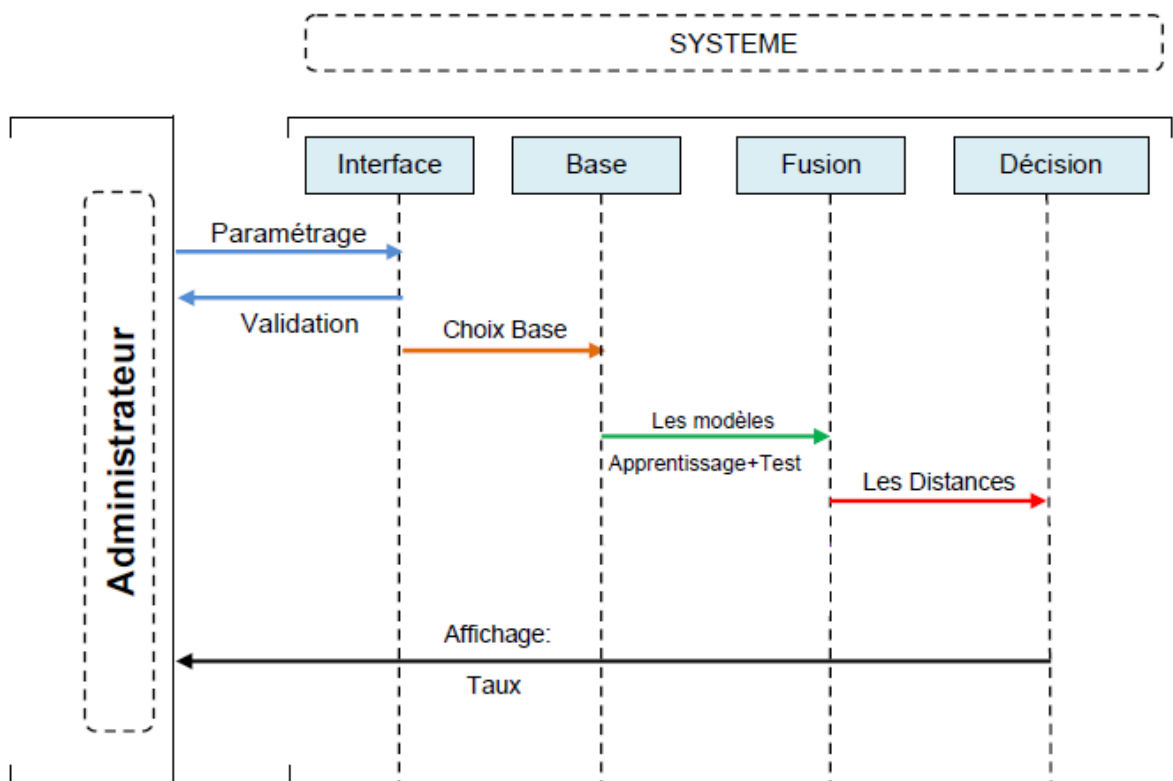


Figure 24:diagramme de séquence pour le calcul des performances du système.

Taux de Reconnaissance:

Le taux de reconnaissance représentera le pourcentage de la bonne identification sur ensemble de visages de tests.

Généralement la base de données visages est décomposée en deux ensembles : ensemble utilisé pour l'apprentissage et le reste pour le test.

Pour s'approcher de la réalité ce découpage se fait d'une façon aléatoire, c'est-à-dire on répète 10 expériences où à chaque fois on sélectionne aléatoirement l'ensemble d'apprentissage et les images restantes sont utilisées pour le test, le taux final sera la moyenne des taux obtenus en 10 expériences (plus le nombre d'expérience est grand plus l'estimation du taux est réelle).

Algorithme de calcul du taux de reconnaissance :

Posant N_{app} nombre d'image d'apprentissage pour un individu de la base, N_{test} le nombre d'image test pour la meme personne, et $N_{personne}$ le nombre de personne dans la base.

On dispose d'une matrice des modèles d'apprentissage W de taille : $p*(N_{app}*N_{personne})$

Avec p : la taille du vecteur modèle et une matrice des modèles de test T de taille :

$p*(N_{test}*N_{personne})$.

1- Calcul des distances entre les vecteurs colonnes des deux matrices T et W , on obtient la matrice D des distances de taille $(N_{test}*N_{personne}) * (N_{app}*N_{personne})$. Chaque ligne représente un vecteur de distances entre un modèle de test et les modèles de la base.

2- $Taux0$;

Pour i allant de 1 à $(N_{test}*N_{personne})$

Pos $\rightarrow \min(D(i, :))$

Si $(pos/N_{app} = i/N_{test})$

$Taux++$

Fsi

Fpour

$Taux \rightarrow Taux / (N_{test}*N_{personne})$.

3. Structure de Données et Implémentation :

3.1 Environnement du travail

Dans cette section, nous présenterons les environnements matériel et logiciel de notre travail.

3.1. Environnements matériel

Afin de mener à bien ce projet, il a été mis à notre disposition un ensemble de matériels dont les caractéristiques sont les suivantes :

Un ordinateur DELL avec les caractéristiques ci dessous:

-Processeur : Intel ® Core(TM) i3-3110M CPU @ 2.40GHZ2.40GHZ

- Mémoire installée (RAM) : 4 .00 Go (3.38 Go utilisable)
- Disque dur :500 Go
- OS : Microsoft Windows Seven

3.2 Outils de développement :

Nous avons eu recours lors de l'élaboration de notre system à matlab 11 que nous présenterons ci-dessous.

3.2.1 Matlab 2011

Matlab et son environnement interactif et un langage de haut niveau qui permet l'exécution de tâches nécessitant une grande puissance de calcul et dont la mise en œuvre sera plus simple et rapide qu'avec des langage de programmation traditionnels tels que le C ,C++. Il dispose de plusieurs boites à outils en particulier celle du traitement d'image « *ImageProcessing Toolbox* » qui propose un ensemble d'algorithmes et d'outils graphique de référence pour le traitement, l'analyse, la visualisation et le développement d'algorithmes de traitement d'image .

Dans cette application nous avons employé les programmes matlab disponibles sur internet de : [69], [70].

4 Présentation de l'application :

On présente dans cette section les différents aspects de reconnaissance.

4.1 Interface Authentification

C'est une interface destinée à l'utilisateur, un utilisateur doit s'authentifier avant d'utiliser le logiciel qui lui demande un nom d'utilisateur et un mot de passe afin d'éviter des erreurs de sauvegarde de faux visages dans la base.



Figure 25:Interface graphique d'authentification

4.2 Interface Présentation

Elle est simple et permet d'illustrer les principaux processus du système de reconnaissance (identification, calculer le taux.....)

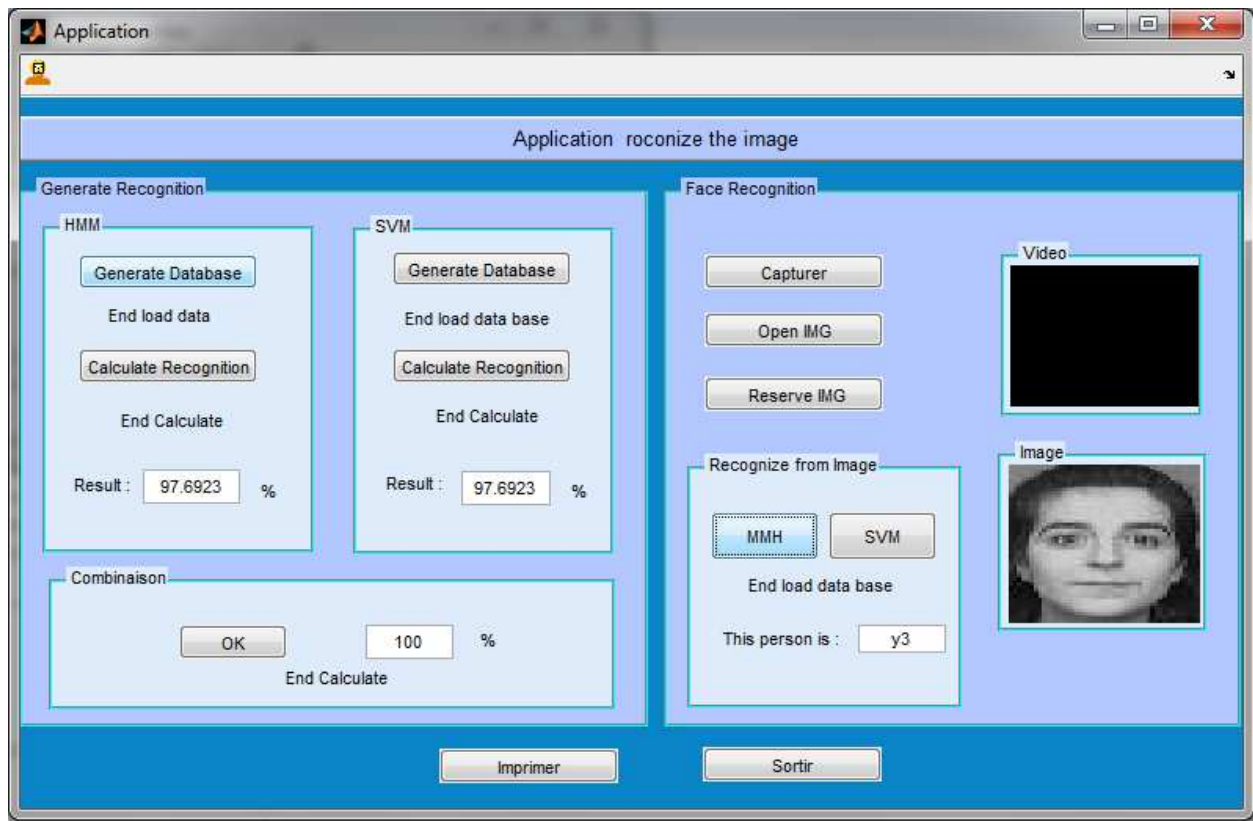


Figure 26:Interface graphique de reconnaissance faciale

Application : regroupant 2 panneaux (voir figure 26)

Generate recognition : destiné à montrer les performances du système sur un ensemble de grande taille d'images.

Face recognition : destiné à montrer les performances du système sur une image de la base ou sur une image capturée par camera (figure 27).

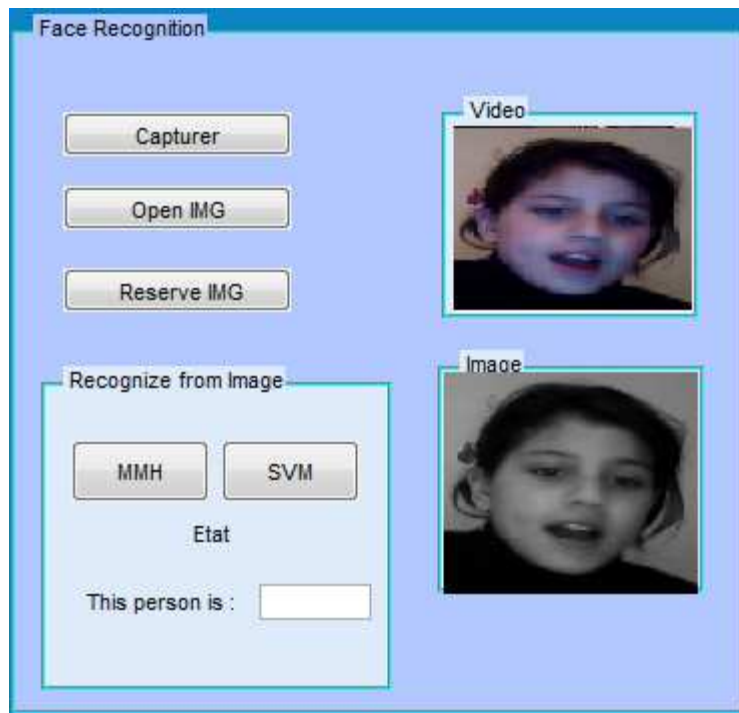


Figure 27:Le panneau face recognition

Ce panneau contient les trois boutons suivant :

Capturer : permet de capturer l'image d'un visage par caméra

Open IMG : pour afficher une image de test de la base du système

Reserve IMG : pour sauvegarder une nouvelle image capturée dans la base du système

Les autres boutons HMM et SVM permettent d'identifier le visage de test utilisant respectivement MMC et SVM

Le panneau **Generate recognition** regroupe trois sous panneaux (voir figure 28)

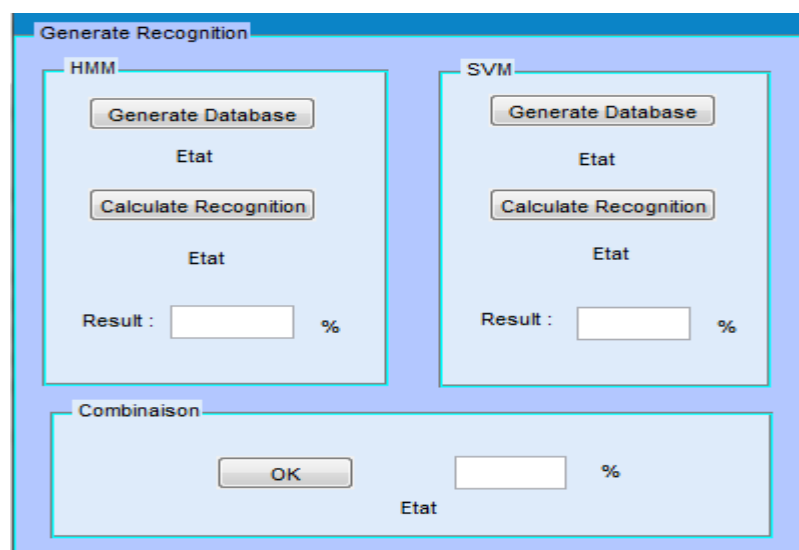


Figure 28:8Interface graphique de reconnaissance faciale générale

Un panneau HMM destiné pour montrer les performances du classificateur MMC (voir figure 29)

Ce panneau contient les boutons

- Generate database : permet de générer la base projetée par SVD
- calculate Recognition : calcule le taux de reconnaissance du classificateur MMC

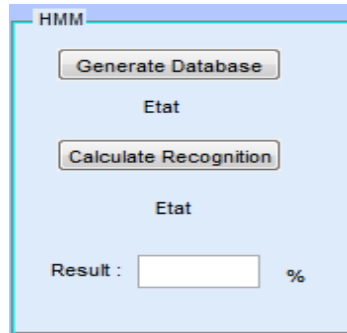


Figure 29:Panel HMM

Un panneau SVM destiné pour montrer les performances du classificateur SVM (voir figure 30)

Ce panneau contient les boutons

- Generate database : permet de générer la base projetée par PCA
- Calculate Recognition : calcule le taux de reconnaissance du classificateur SVM

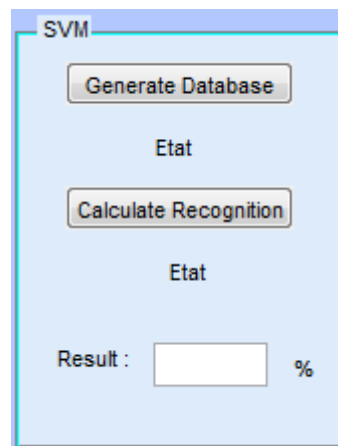


Figure 30:Panel SVM

Un panneau combinaison : pour montrer le résultat de la combinaison des deux classificateurs SVM et MMC utilisant la méthode de combinaison par vote majoritaire (voir figure 31).

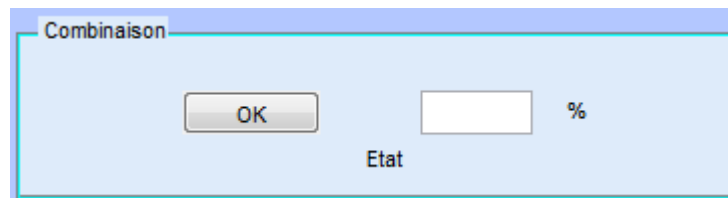


Figure 31:Panel Combinaison

Classificateur	Taux de reconnaissance
MMC	96.9231%
SVM	96.5385%
Combinaison SVM & MMC par vote majoritaire	100%

Tableau 2:Table récapitulative des résultats expérimentaux

Ces résultats sont tous très prometteurs et très significatifs dans le cas d'un classificateur issu de la combinaison du SVM et MMC dû à la complémentarité des résultats de ces deux classificateurs jugés efficaces dans la littérature. Il nous reste à valider ce programme sur d'autres bases de visages.

Conclusion Générale:

Dans ce travail de master, nous avons traité un problème principal de la biométrie, celui de la reconnaissance faciale. Ce problème a été toujours ouvert du à l'incapacité des anciennes méthodes de classification et d'extraction des attributs caractéristiques les plus pertinents.

Pour résoudre ce problème de la reconnaissance faciale, nous avons proposé une approche par combinaison de classificateurs notamment la combinaison d'un SVM et un MMC. La règle de combinaison que nous avons employée est le vote majoritaire. Les résultats obtenus sont très significatifs sur la base ORL ainsi que sur des images réelles capturées par Caméra. Durant cette étude, nous avons remarqué l'efficacité de ces deux classificateurs SVM et MMC en résolvant le problème de la reconnaissance faciale, ces deux classificateurs ont atteint un taux supérieur à 96% alors que leur combinaison a fourni un taux de 100%, ce qui prouve l'intérêt de ce travail dans le domaine de la reconnaissance faciale.

Bibliographie :

- [1] <http://www.developpez.net/>
- [2] THÈSE Walid Hizem : Capteur Intelligent pour la Reconnaissance de Visage 2009.
- [3] S.Z. LI and A.K. Jain. Handbook of Face Recognition. Springer, 2005.
- [4] R. Gross and V. Brajovic. An image preprocessing algorithm for illumination invariant facerecognition. In 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA). Springer, June 2003.
- [5] THÈSE SOUHILA GUERFI ABABSA Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D.2008
- [6] R. Gross and V. Brajovic. An image preprocessing algorithm for illumination invariant face recognition. In 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA). Springer, June 2003.
- [7] R. Brunelli and T. Poggio. Face recognition : Features versus templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10) :1042_1052, 1993.
- [8] A.J. Goldstein, L.D. Harmon, and A.B. Lesk. Identifcation of human faces. Proceedings of the IEEE, 59(5) :748_760, May 1971.
- [9] I.J. Cox, J. Ghosn, and P.N. Yianilos. Feature-based face recognition using mixture distance. Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on, pages 209_216, Jun 1996.
- [10] Y. Adini, Y. Moses, S. Ullman, Face recognition: The problem of compensating for changes in illumination direction. IEEE Trans. Patt. Anal. Mach. Intell. 19, 721–732, 1997.
- [11] N. Azizi, N. Farah, M. Sellami, “Off-line handwritten word recognition using ensemble of classifier selection and features fusion”, Journal of Theoretical and Applied Information Technology, Vol 14. No. 2 – 2010.
- [12] P. J. Phillips, P. J. Grother, R. J. Micheals, D. M. Blackbrun, E. Tabassi, and J. M. Bone. Face recognition vendor test 2002: Evaluation report. NISTIR 6965. Available online at <http://www.frvt.org>, 2003.
- [13] R. Gross, J. Shi, J. Cohn. Quo Vadis Face Recognition?. Third Workshop on Empirical Evaluation Methods in Computer Vision, December, 2001.
- [14] A.M Martínez, R. Banavente. The AR face database. Tech. Report 24 CVC Barcelone, Espagne, June 1998.
- [15] Ming-Hsuan Yang, David J. Kriegman et Narendra Ahuja. Detecting faces in images : A survey. Dans IEEE Transactions on Pattern Analysis and Machine Intelligence,

- volume 24(1), pages 34–58, 2002.
- [16] C. Kotropoulos and I. Pitas. Rule-Based Face Detection in Frontal Views. Proc. Int'l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997.
- [17] T. Kanade. "Picture Processing by Computer Complex and Recognition of Human Faces," PhD thesis, Kyoto Univ., 1973.
- [18] G. Yang and T. S. Huang. "Human Face Detection in Complex Background," Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.
- [19] P. Sinha, "Object Recognition via Image Invariants: A Case Study," Investigative Ophthalmology and Visual Science, vol. 35, no. 4, pp. 1735-1740, 1994.
- [20] P. Sinha, "Processing and Recognizing 3D Forms," PhD thesis, Massachusetts Inst. of Technology, 1995.
- [21] A. Yuille, P. Hallinan, and D. Cohen, "Feature Extraction from Faces Using Deformable Templates," Int'l J. Computer Vision, vol. 8, no. 2, pp. 99-111, 1992.
- [22] R. Brunelli, T. Poggio. Face recognition: features versus templates. IEEE Trans. Pattern vol. 3, no. 1, pp. 71-86, 1991.
- [23] M. Turk and A. Pentland. "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
- [24] P. Sinha, B. Balas, Y. Ostrovsky, R. Russell. Face Recognition by Humans: 19 Results all Computer Vision Researchers Should Know About. Proceedings of the IEEE, 1948-1962, 2006.
- [25] X. Tana, C. Songcan. Face recognition from a single image per person: A survey. Pattern Recognition, 2006.
- [26] A.J. O'Toole, H. Abdi, Low-dimensional representation of faces in higher dimensions of the face space, Opt. Soc. Am. 10 (3), 405– 411, 1993.
- [27] A.K. Jain, B. Chandrasekaran. Dimensionality and sample size considerations in pattern recognition practice, in: P.R. Krishnaiah, L.N. Kanal (Eds.), Handbook of Statistics vol. 2, pp. 835–855, 1982.
- [28] A.K. Jain, B. Chandrasekaran. Dimensionality and sample size considerations in pattern recognition practice, in: P.R. Krishnaiah, L.N. Kanal (Eds.), Handbook of Statistics, vol. 2, North-Holland, Amsterdam, pp. 835–855, 1987.
- [29] S.J. Raudys, A.K. Jain, Small sample size effects in statistical pattern recognition: recommendations for practitioners, IEEE Trans. Pattern Anal. Mach. Intell. 13 (3) 252–264, 1991
- [30] R.P.W. Duin. Small sample size generalization, in: G. Borgefors (Eds.), SCIA'95,

- Proceeding of the Ninth Scandinavian Conference on Image Analysis, vol. 2, Uppsala, Sweden, 6–9, pp. 957–964, 1995.
- [31] L.I. Kuncheva, “Clustering-and-selection model for classifier combination”. In Proc. Knowledge-Based Intelligent Engineering Systems and Allied Technologies, Brighton, UK, pp. 185–188. 2000.
- [32] L. Sirovich, M. Kirby, Low-dimensional procedure for the characterization of human faces, *J. Opt. Soc. Am. A* 4 (3) (1987) 519–524.
- [33] X. Wang, X. Tang, Unified subspace analysis for face recognition, Proceedings of the Ninth IEEE International Conference on Computer Vision, 2003, pp. 679–686.
- [34] B. Moghaddam, A. Pentland, Probabilistic visual learning for object representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (7) 696–710, 1997
- [35] X. He, X. Yan, Y. Hu, p. Niyogi, H. Zhang. Face recognition using Laplacian faces, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (3) 328–340, 2005.
- [36] A. Martinez, A.C. Kak. PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2) 228–233, 2001.
- [37] H. Zouari, L. Heutte, Y. Lecourtier, A. Alimi. "Un panorama des méthodes de combinaison de classificateurs en reconnaissance de formes". *RFIA'2002*, Angers, France, vol. 2, pp. 499-508, 2002.
- [38] B.S. Manjunath, R. Chellappa, C.V.D. Malsburg, A feature based approach to face recognition, in: Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 373–378, 1992.
- [39] T.S. Lee, Image representation using 2-d Gabor wavelets, *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (10) 959–971, 1996.
- [40] S.C. Chen, J. Liu, Z.-H. Zhou, Making FLDA applicable to face recognition with one sample per person, *Pattern Recognition* 37 (7), 1553–1555, 2004.
- [41] X. Tan, S.C. Chen, Z.-H. Zhou, F. Zhang, Recognizing partially occluded, expression variant faces from single training image per and soft kNN ensemble, *IEEE Trans. Neural Networks* 16 (4) 875–886, 2005.
- [42] A.M. Martinez, Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (6) (2002) 748–763.
- [43] F. Samaria, Face segmentation for identification using hidden Markov models, in: British Machine Vision Conference, BMVA Press, pp. 399–408, 1993.
- [44] H.S. Le, H. Li. Recognizing frontal face images using hidden Markov models with one

- training image per person, Proceedings of the 17th International Conference on Pattern Recognition (ICPR04), vol. 1, pp. 318– 321, 2004.
- [45] P. Penev, J. Atick, Local feature analysis: a general statistical theory for object representation, *Netw.: Comput. Neural Syst.* 7 477– 500, 1996.
- [46] W. T. Hubel, D. Functional architecture of macaque monkey visual cortex. In *Royal Society on Biology* 198, 1978.
- [47] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. on Computers*, 42(3) :300_311, 1993.
- [48] Contributions à la reconnaissance de visages à partir d'une seule image et dans un contexte non-contrôlé, l'École Doctorale Électronique, Électrotechnique, Automatique, Traitement du Signal, GRENOBLE, 19 novembre 2010
- [49] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans. Image Processing*, 11 :467_476, 2002.
- [50] A. Mellakh. Reconnaissance des visages en conditions dégradées. PhDthesis, l'Institut National des Télécommunications, 2009.
- [51] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. PAMI*, 24(7) :971_987, 2002.
- [52] Contributions à la reconnaissance de visages à partir d'une seule image et dans un contexte non-contrôlé, l'École Doctorale Électronique, Électrotechnique, Automatique, Traitement du Signal, GRENOBLE, 19 novembre 2010.
- [53] T. Ahonen, A. Hadid, and M. Pietikainen. Face recognition with local Binary patterns. In *ECCV*, pages 469_481, 2004.
- [54] L. Pagie and P. Hogeweg “Evolutionary consequences of coevolving targets”. *Evolutionary Computation* 5, pp.401-418. 1997.
- [55] S. B. Kotsiantis, “Supervised Machine Learning: A Review of Classification Techniques”, *Informatica* 31 (2007) 249-268. 2007.
- [56] F. Lauer, C. Y. Suen and G. Bloch, "A trainable feature extractor for handwritten digit recognition," *Pattern Recognition*, vol. 40, 1816-1824, 2007.
- [57] G. G. Rajput, S. M. Mali, “Handwritten Marathi Numeral Recognition with Fourier Computer Vision and Information Technology: Advances and Applications, I.K. Int. Pub. India (2009), pp: 528-534. 2009.

- [58] A.Borji, et M. Hamidi. Support Vector Machines for Persian Font recognition. In: International Conference on Computer, Electrical, Systems Science, and Engineering (CESSE 2007), Prague, Czech Republic July 27-29, 2007.
- [59] M.Antkowiak, "Artificial Neural Networks vs. Support Vector Machines for Skin Diseases Recognition". Master's Thesis, Ume University. 2006.
- [60] E. J. Bredensteiner, K. P. Bennett, Multicategory classification by support vector machines, Computational Optimization and Applications, 12 (1999). pp 53-79. 1999.
- [61] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 1-47, 1998.
- [62] Turk, M.A., Pentland, A.P., 1991. "Eigenfaces for recognition". J. Cognit. Neurosci. Vol .3, p.71-86.
- [63] B. H. Juang, L. R. Rabiner, "Hidden Markov Models for Speech Recognition", Technometrics, Vol. 33, No. 3. pp. 251-272. 1991.
- [64] L.R. Rabiner "A tutorial on hidden Markov models and select applications in speech recognition", Proceeding of IEEE, Vol 77, N° 2, pp: 257-286. 1989.
- [65] A. Churbanov, S. Winters-Hilt, Implementing EM and Viterbi algorithms for Hidden Markov Model in linear memory, BMC Bioinformatics.:9:224. 2008.
- [66] S.Tulyakov, S.Jaeger, V.Govindaraju, D.Doermann, "Review of classifier combination methods". In: Simone Marinai, H.F. (Ed.), Studies in Computational Intelligence: Machine Learning in Document Analysis and Recognition, pp. 361-386. Springer, 2008
- [67] I. Gasmi, H. Merouani, L. Souici-Meslati, "Combinaison de classificateurs". SETIT 2005. 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, TUNISIA, March 27-31. 2005.
- [68] G. Claude, Indexation multimédia pour la reconnaissance de gestes et le contrôle de la production sonore, Master informatique, 2011.
- [69] H. Miar-Naimi and P. Davari A New Fast and Efficient HMM-Based Face Recognition System Using a 7-State HMM Along With SVD Coefficients, IJEEE. 2013.
- [70] michael.song
https://www.google.fr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CDIQFjAA&url=http%3A%2F%2Fread.pudn.com%2Fdownloads94%2Fdoc%2F374475%2Fsvm_pca%2Fclassifier.m__.htm&ei=CnB2U9bbINCX0QX2jICwCw&usg=AFQjCNGkgPeSLhLOLOIhnjOX6Mn2FG69fA