

Chapitre 5

Conception de l'Approche RV Monomodale.

5.1 Introduction

Dans ce chapitre, nous détaillons les principales expériences que nous avons menées durant la thèse ainsi que les résultats monomodaux (sur une seule modalité à la fois) obtenus via les trois principales approches testées :

- l'approche 2D pour laquelle des investigations sur la meilleure orientation du filtre de Gabor et une comparaison des méthodes de réduction sont réalisés. Dans cette approche nous étudions les méthodes de réduction linéaires et non linéaires,
- l'approche 3D dans laquelle la génération de cartes de profondeur est réalisée,
- l'approche 3D expression basée sur la combinaison par concaténation des régions d'intérêts et du visage.
- et enfin l'approche 3D expression basée sur la fusion des caractéristiques du visage neutre et des six expressions (dégout, joie, peur, surprise, colère,

Les deux principaux modes de test d'un système biométrique sont l'authentification (ou vérification) et l'identification. L'utilisation d'un système biométrique en authentification nécessite le calcul d'un score de similarité entre deux modalités biométriques, la décision étant ensuite prise en fonction de ce score et d'un seuil. L'utilisation d'un système d'identification nécessite le calcul de n scores de similarité. Le calcul du rang pour une image test, c'est à dire le moment à partir duquel la bonne identité est retrouvée, permet d'évaluer qualitativement le système biométrique.

Toutes les expériences menées sont en identification, c'est-à-dire que le système doit retrouver l'identité de la personne testée. Celles-ci ont été réalisées sous l'hypothèse forte que la personne testée a préalablement été enrôlée.

Dans la suite, nous utilisons parfois le terme galerie pour désigner les images utilisées lors de l'enrôlement des personnes, et le terme probe pour désigner les images utilisées pour tester la méthode d'identification.

5.2 Système de reconnaissance de visage

- a) Première expérimentation, dans un premier temps, nous avons mis en œuvre un système de reconnaissance monomodale sur le visage 2D neutre.

- b) Deuxième expérimentation, nous essayons de contourner le problème de la première expérimentation, à savoir la limite des images 2D face aux changements d'expressions. Nous tentons de faire la fusion des différentes régions d'intérêts afin de mieux identifier les personnes.
- a) Troisième expérimentation, dans cette dernière application, nous essayons d'améliorer les performances du système de reconnaissance 3D qui a des difficultés à reconnaître les visages avec expressions. Nous essayons de faire la fusion du visage neutre avec les six expressions et des résultats encourageants sont obtenus.

5.2.1 Détection de visage

Une multitude de techniques ont été développées allant de la simple détection du visage, à la localisation précise des régions caractéristiques du visage, tels que les yeux, le nez, les narines, les sourcils, la bouche, les lèvres, les oreilles, etc...



Figure 5.1 Détection de visage

Les performances globales de tout système automatique de reconnaissance dépendent amplement des performances de la détection de visages[320].

5.2.2 Prétraitement

Où les données biométriques sont traitées pour enlever le bruit lié à l'environnement ou au dispositif de capture. Pour cela nous faisons une photo normalisation qui a un double effet : d'une part elle supprime pour tout vecteur image x un éventuel décalage par rapport à l'origine et ensuite elle supprime tout effet d'amplification (multiplication par un scalaire). Pour chaque image on effectue l'opération suivante :

$$photo\ normalisation(x) = \frac{x - mean(x)}{std(x)}$$

$mean(x)$: le moyen de x et $std(x)$: l'écart type de x .

5.2.3 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 de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes. Le choix de 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. Les plus performantes dans le contexte de reconnaissance de visages sont les ondelettes de **Gabor** et les indices **LBP Local Binary Patterns**. Ici on se limitera à la première pour les raisons suivantes :

- Les ondelettes de Gabor sont connues comme le moyen d'analyse espace-fréquence qui minimise l'incertitude de Heisenberg à savoir si on est précis dans l'analyse de l'information dans l'espace du pixel on le sera moins dans l'espace fréquentiel et vice versa.
- Cette caractéristique fait des ondelettes de Gabor un moyen puissant pour l'analyse de la texture [321].
- D'après **Morizet** qui a consacré un chapitre à la **Perception Visuelle** et les **Neurosciences** (**Chapitre 3** [3]) en se basant sur les expériences de **Pollen et Ronner** [322] et **Jones et Palmer** [323]. Le filtre de Gabor est le filtre le plus proche des **champs réceptifs linéaires (RFs)** [324]

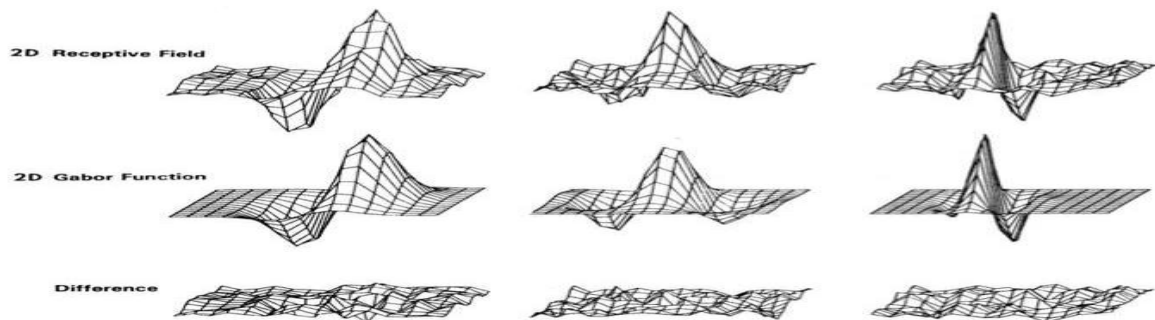


Figure 5.2 Adaptation d'un filtre de Gabor aux données correspondant aux mesures du comportement de la réponse d'une certaine cellule simple. [323]

Sur la première ligne de la **figure 5.2** on peut voir le **RF expérimental** et sur la deuxième ligne, le **RF modélisé par des ondelettes 2D de Gabor**. On peut apprécier la **très grande similitude** entre les deux entités par leur **très faible différence** représentée sur la dernière ligne. Pour chaque type de caractéristique, il y'a plusieurs façons de les utiliser. Nous présentons dans la suite seulement le principe de base. Il convient aussi de noter qu'il existe des variantes de ces caractéristiques originales.

Les ondelettes de Gabor [325]

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 de

Gabor s'apparentent à celles du système visuel humain [323]. Les articles [326,327,328] (le premier est dans Nature) indiquent que la représentation par ondelettes de **Gabor** des images faciales est robuste aux changements causés par des variations d'éclairage ou par des modifications d'expressions faciales.

Les ondelettes en dimension deux de **Gabor** ont été introduites dans le domaine de la recherche biométrique par **Daugman** [329] pour la reconnaissance d'iris. **Ladesetal.** [131] ont utilisé les premières ondelettes de **Gabor** en reconnaissance faciale en utilisant la **Dynamic Link Architecture**. Un noyau de filtre de **Gabor** est le produit d'une onde complexe sinusoidale avec une enveloppe gaussienne. Une ondelette de **Gabor** généralement utilisée dans la reconnaissance faciale est définie comme suit :

$$\Psi_{\theta,\lambda}(z) = \frac{\|k_{\theta,\lambda}\|}{\sigma^2} e^{-\frac{\|k_{\theta,\lambda}\|^2 \|z\|^2}{2\sigma^2}} [e^{ik_{\theta,\lambda}z} - e^{-\frac{\sigma^2}{2}}] \quad (5.1)$$

où $z = (x,y)$ est le point de coordonnées (x,y) dans le plan image. Les paramètres θ et λ définissent l'orientation et la fréquence des noyaux de **Gabor**. $\| \cdot \|$ est l'opérateur norme, et σ l'écart-type de l'enveloppe gaussienne. Le paramètre d'orientation θ (ou θ) fait pivoter l'ondelette autour de son centre. Il détermine l'angle des contours ou des lignes de l'image auxquelles l'ondelette va être sensible. Le paramètre λ (ou r) spécifie la longueur d'onde du cosinus ou inversement la fréquence centrale de l'ondelette. Les ondelettes avec une grande longueur d'onde seront sensibles à des changements progressifs d'intensité dans une image. Les ondelettes avec une petite longueur d'onde seront sensibles à des contours et des lignes abruptes. 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 fréquences et d'orientations différentes. La convolution de l'image I et d'un noyau de **Gabor** $\Psi_{\theta,\lambda}(z)$ est définie par :

$$G_{\theta,\lambda}(z) = I(z) * \Psi_{\theta,\lambda}(z) \quad (5.2)$$

Comme on peut le remarquer dans l'équation 5.1, les filtres de **Gabor** ont une forme complexe, il est important d'utiliser les informations données par la partie réelle et la partie imaginaire des coefficients de **Gabor**. Deux choix triviaux s'offrent à nous l'étude de l'amplitude et l'étude de la phase de **Gabor**. L'amplitude et la phase seront notées respectivement par :

$$M(IG_{r,\theta}(x,y)) = \sqrt{\text{Im}(IG_{r,\theta}(x,y))^2 + \text{Real}(IG_{r,\theta}(x,y))^2}$$

$$P(IG_{r,\theta}(x,y)) = \arctan\left(\frac{\text{Im}(IG_{r,\theta}(x,y))}{\text{Real}(IG_{r,\theta}(x,y))}\right)$$

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 **figure 5.6**. Par conséquent, si on ne considère que la réponse en amplitude, chaque pixel est décrit par un vecteur de dimension 40. Ce vecteur de dimension 40 est également appelé « jet » et il a été utilisé largement dans les systèmes les plus anciens, tels que le **DLA** et les plus récents **EGBM**[330]. Notons que ce sont des méthodes basées sur des points caractéristiques qui doivent être détectés très précisément. Dans les méthodes telles que [279], les auteurs ont utilisé directement une image de caractéristiques contenant les 40 cartes d'amplitude comme entrée de reconnaissance. Dans [57], **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 ou l'**ACP** avec 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.

5.3 Evaluation expérimentale de la reconnaissance monomodale 2D

5.3.1 Base de données

Nos expériences sont exécutées sur des images frontales de visage de la base de données **XM2VTS**. C'est une base multimodale développée au sein du projet européen **ACTS**, elle est utilisée pour la vérification d'identité, elle contient 8 images par visage de 295 personnes. La base de données **XM2VTS**[331] a été réalisée sur une longue période ce qui a permis d'obtenir plusieurs images d'une même personne et donc une grande variabilité d'apparence (changement de coiffure, présence et/ou absence de lunettes, etc...). Par contre, seules les expressions neutres ont été considérées. Pour la tâche de vérification, un protocole standard d'estimation de performances a été développé (appelé «Lausanne Protocol Splits Randomly»). Il existe deux configurations différentes, la configuration I et la configuration II. Nous n'utiliserons que la **configuration I** puisqu'elle est la plus complexe. Elle consiste à séparer les individus en deux classes, Client et Imposteur. Le groupe Client contient 200 sujets, alors que le groupe Imposteur est divisé en 25 imposteurs pour l'évaluation et 70 imposteurs pour le test, huit images des quatre sessions sont utilisées.



Figure 5.3 Exemples des images de la base de données XM2VTS
 (a) personnes différents, (b) même personne.

La répartition des photos dans les différents ensembles est présentée dans le **tableau 5.1**.

Tableau 5.1 Répartition des photos dans les différents ensembles.

Ensemble	Clients	Imposteurs
Apprentissage	600 images (3 par personne)	0 images
Evaluation	600 images (3 par personne)	200 images (8 par personne)
Test	200 images (2 par personne)	560 images (8 par personne)

Notre travail est exécuté avec un ordinateur : Intel Pentium Dual CPU 2.2 GHz , 1.49 Go de RAM et le langage de programmation utilisé est le **MATLAB R2009b**.

5.3.2 Architecture de l'approche monomodale 2D

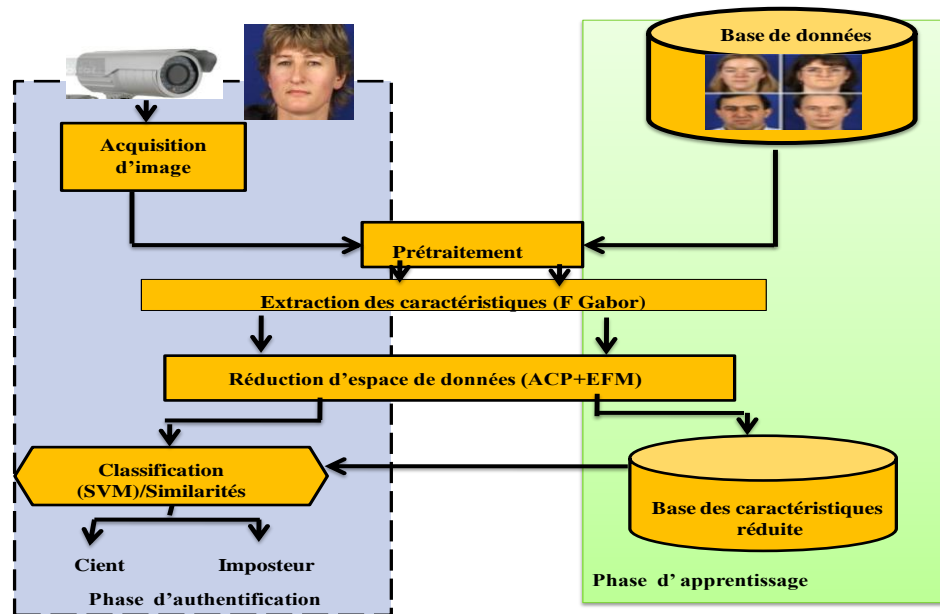


Figure 5.4 Schéma de conception de notre approche monomodale.

5.4 Représentation du visage et choix des filtres de Gabor

5.4.1 Représentation du visage : Chaque image est constituée de plusieurs informations comme la couleur, l'arrière plan, les cheveux, les cols de chemise, les oreilles etc... Pour cela, la première étape nécessaire est de découper l'image par une fenêtre rectangulaire de taille **161x121** centrée autour des caractéristiques les plus stables liées aux yeux, aux sourcils, au nez et à la bouche. Une décimation de facteur 2 pour réduire par 4 la taille de l'image

découpée puis nous avons utilisé l'espace de couleur **HSV** (Hue, Saturation, Value). Notre choix est justifié par les résultats performants dans la littérature [57,332,333].



Figure 5.5(a) Image originale, (b) Image découpée, (c) Image décimée, (d) Image en système HSV, (e) La composante S de HSV.

Nous considérons la composante *S* "Saturation" comme caractéristique selon [17] puis nous faisons la comparaison des performances des méthodes de réduction d'espace par calcul de distance.

5.4.2 Choix des filtres de Gabor : La méthode de reconnaissance que nous étudions dans cette partie repose principalement sur l'extraction des caractéristiques du visage par filtrage de **Gabor 2D**. La robustesse des filtres de Gabor a fait d'eux un moyen puissant d'analyse de textures et de classification. Les filtres de Gabor analysent la texture d'un objet suivant différentes résolutions et différents angles. Dans le domaine spatial, un filtre de Gabor 2D est une fonction à noyau gaussien modulée par une onde sinusoïdale plane complexe. L'équation générale d'une ondelette de Gabor en 2D est :

$$W(x, y, \theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} e^{i(\frac{x'}{\lambda} + \varphi)} \quad (5.1),$$

Où : $x' = x \cos \theta + y \sin \theta$ et $y' = -x \sin \theta + y \cos \theta$.

Il y a donc cinq paramètres qui contrôlent l'analyse par ondelettes [320,330]. Nous choisissons les cinq paramètres comme suit :

- θ spécifie l'orientation du filtre. Nous utilisons dans ce cas huit orientations : $\theta = \{0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$.
- λ spécifie la longueur d'onde et donc la fréquence de la sinusoïde. Cette variable peut prendre cinq valeurs comprises entre 4 et 16 : $\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$.
- φ spécifie la phase de la sinusoïde. Elle vaut 0 ou $\pi/2$ selon que l'on veut la partie réelle ou imaginaire.
- σ spécifie la variance de la gaussienne (sa taille ou résolution). Elle est proportionnelle à la longueur d'onde de la sinusoïde. Dans notre cas $\sigma = \lambda$.
- γ spécifie l'aspect de la gaussienne. Ici les gaussiennes sont de forme circulaire $\gamma = 1$.

Rappelons que la représentation de **Gabor** d'une image de visage est le résultat de la convolution de l'image avec la famille des filtres de **Gabor**, définie par $IG(r, o) = I * G(r, o)$ où $IG(r, o)$ à une certaine résolution r et à une orientation o .

Nous commençons donc par étudier l'influence des caractéristiques des familles des filtres de **Gabor** sur les performances de l'authentification pour en déduire le choix optimal. Pour cela nous utilisons huit orientations $\{0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$ et cinq longueurs d'onde $\{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$. La représentation de l'amplitude l'image considérée est comme le montre **figure 5.6**.

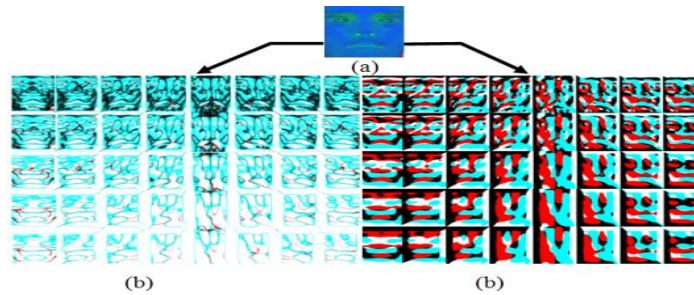


Figure 5.6 Résultats de la convolution d'une image de visage avec une famille de 40 filtres de Gabor (8 orientations (horizontales) et 5 résolutions (verticales)). (a) Image en espace de couleur HSV, l'ensemble (b) les amplitudes et (c) les phases de cette convolution.

- Influence des caractéristiques de la famille des filtres de Gabor sur les performances de la reconnaissance* : Nous nous intéressons d'abord à l'étude de l'amplitude des réponses des filtres de Gabor de l'image considérée **figure 5.6** (b). L'image d'entrée est une image de visage découpée et décimée. Nous considérons la composante **S** de l'image couleur dans l'espace **HSV**. L'algorithme de reconnaissance utilisé dans ce cas est l'*EFM* et la mesure de distance utilisée est la *corrélation*. Le **tableau 5.2** présente les résultats de reconnaissance en terme TEE dans l'ensemble d'évaluation pour différent niveau de résolution et d'orientations des filtres.

Tableau 5.2 TEE pour différents niveaux de résolution et d'orientations des amplitudes des filtres.

Résolution ($\sigma = \lambda$)	Orientations des filtres (θ)							
	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$	$5\pi/8$	$3\pi/4$	$7\pi/8$
4	9.28	10.1	8.13	8.02	8.01	8.04	8.63	7.3
$4\sqrt{2}$	8.33	9.01	7.54	9.61	5.35	7.95	7.2	8.5
8	9.31	7.34	8.7	5.7	7.85	5.13	8.02	8.17
$8\sqrt{2}$	9.54	8.64	7.31	9.36	10.19	8.3	7.07	7.54
16	9.17	8.48	8.65	9.18	9.18	8.64	7.84	7.62

Dans ce tableau nous constatons que le meilleur résultat est obtenu avec un **TEE= 5.13%** ce qui n'est pas satisfaisant.

- Problème de l'utilisation de la phase de Gabor pour les visages* : Lorsque nous considérons une image du visage, certaines parties du visage n'ont aucune texture informative qui pourrait être analysée par les basses résolutions des filtres de Gabor. Pour ces régions, l'analyse par

filtrage de Gabor donne $Real(IG_{s,o}) \cong 0$ et $Im(IG_{s,o}) \cong 0$. Même si ces valeurs sont très proches de 0, l'amplitude de la convolution n'est pas affectée par ce problème, alors que la phase devient une forme indéterminée pour ces régions spécifiques. Pour éviter les formes indéterminées, nous sélectionnons les régions informatives par seuillage de l'amplitude à chaque point d'analyse.

$$P(IG_{s,o}(x,y)) = \begin{cases} \arctan\left(\frac{Im(IG_{s,o}(x,y))}{Real(IG_{s,o}(x,y))}\right) & \text{if } M(IG_{s,o})(x,y) > Th \\ 0 & \text{if } M(IG_{s,o})(x,y) < Th \end{cases} \quad (5.3)$$

Où $(x; y)$ sont les coordonnées du point analysé et Th est le seuil de sélection de la phase.

- *Optimisation du seuil pour la sélection de la phase*

Afin d'étudier l'influence du seuillage de la phase en fonction des performances. La **figure 5.7** montre l'évolution des taux d'égale d'erreur *TEE* en fonction du seuil Th par un filtre de Gabor avec résolution $\sigma = \lambda = 4$ et Orientation $\theta = \pi/2$.

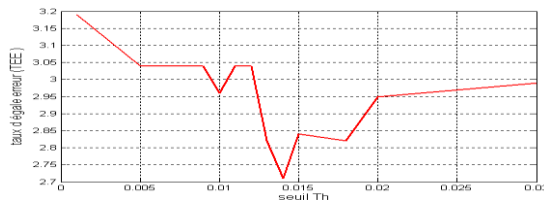


Figure 5.7 Taux d'égale erreur dans l'ensemble d'évaluation en fonction du seuil Th .

D'après la courbe de la **figure 5.7** nous constatons que la variation de *TEE* en utilisant la phase de Gabor est bien liée aux seuils du filtrage. Donc, notre choix est porté sur le seuil du filtrage $Th = 0.014$.

Dans cette deuxième série de mesures, ce seuil est maintenu et nous nous intéressons donc aux phases de Gabor. Nous cherchons dans ce cas les filtres de Gabor optimaux ce que nous présentons dans le **tableau 5.3**.

Tableau 5.3 TEE pour différents niveaux de résolutions et d'orientations des phases des filtres.

λ	Orientations des filtres (θ)							
	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$	$5\pi/8$	$3\pi/4$	$7\pi/8$
4	4.79	5.14	4.12	4.96	2.69	3.3	3.79	4.64
4	4.8	5.29	6	5.28	4.15	4.88	4.87	5.3
$\sqrt{2}$								
8	6.03	6.53	7.16	6.79	6.04	6.85	6.29	7.04
$8\sqrt{2}$	6.64	7.47	7.29	8.21	8.52	8.14	7.66	7.8
16	6.5	7.01	8.16	8.45	9.01	8.61	7.99	7.84

Nous remarquons que la première résolution et les orientations : $\theta = \pi/2, 5\pi/8, 3\pi/4$ donne le meilleur TEE. Les résultats obtenus par les phases de **Gabor** sont satisfaisants et encourageants. Ce qui confirme les travaux d'**Oppenheim** ainsi que ceux de

Mellakh dans [334,57], qui ont démontré sur d'autres bases de données que la phase des analyses multi-résolutions est plus informative que les amplitudes qui peuvent souffrir des problèmes d'illumination.

Nos meilleurs systèmes d'authentification de visages sont présentés dans le **tableau 5.4**.

Tableau 5.4 Résultats du système d'authentification de visages pour les quatre meilleurs systèmes retenus dans l'ensemble d'évaluation et test (intervalle de confiance paramétrique à 95%).

Méthodes	Ensemble évaluation	Ensemble de test		
	TEE %	TFR %	TFA %	TR %
Système 1	2.66 ± 0.72	2 ± 1.37	3.66 ± 0.11	94.33 ± 1.48
Système 2	2.69 ± 0.72	0.5 ± 0.69	4.07 ± 0.12	95.43 ± 0.81
Système 3	3.3 ± 0.8	2 ± 1.37	4.41 ± 0.12	93.59 ± 1.49
Système 4	3.79 ± 0.85	0.5 ± 0.69	4.47 ± 0.12	95.03 ± 0.82

- Le **système 1** qui utilise l'image couleur originale prétraitée.

Les trois autres systèmes utilisent la phase filtrée de la convolution de la **composante S** de l'espace de couleur TSV avec le filtre de **Gabor** pour la première résolution ($\sigma = \lambda = 4$) pour diverses orientations.

- Le **système 2** utilise l'orientation ($\theta = \pi/2$).
- Le **système 3** utilise l'orientation ($\theta = 5\pi/8$).
- Le **système 4** est l'orientation ($\theta = 3\pi/4$).

Nous utiliserons donc ces trois phases des filtres de **Gabor** pour la conception de notre Système Multi-Classifieurs (MCS en **chapitre 6**).

5.5 Choix des méthodes de réduction

Première expérience : Cas de méthodes linéaires

5.5.1 Comparaison des méthodes de réduction associée au calcul de distance [336]

Dans cette partie, nous étudions les algorithmes de réduction d'espace : l'ACP, la LDA, la DLDA, LDA2D et EFM et surtout leur application au domaine du visage. Une présentation théorique de ces approches, est faite dans la première partie du **chapitre 3**. Une étude comparative entre ces algorithmes est présentée dans ce chapitre. Nous nous focaliserons ensuite sur plusieurs particularités que rencontrent ce type d'algorithmes comme le choix des conditions adéquates d'apprentissage. L'objectif de cette partie du chapitre est de comparer les différentes méthodes afin de pouvoir choisir le meilleur algorithme pour la suite de nos études.

Mesure de distance

Cette étape consiste à calculer la distance euclidienne minimale entre le visage test et l'ensemble des visages A_i représentant soit la classe de personnes à vérifier (2D) ou la totalité des cartes de profondeur (3D) en mode identification. Lorsqu'on souhaite comparer deux

vecteurs de caractéristiques issus du module d'extraction de caractéristiques d'un système biométrique, on peut soit effectuer une mesure de similarité (ressemblance), soit une mesure de distance (divergence). La première catégorie de distances est constituée de distances Euclidiennes et sont définies à partir de la *distance de Minkowski d'ordre p* dans un espace euclidien R^N (N déterminant la dimension de l'espace euclidien). [24,69,103]

$$d_{Euclidienne} = \|A - B\|^2 \quad (5.4)$$

La distance employée dans notre expérience pour la mesure similarité est l'angle ou (corrélation normalisée) qui est définis comme suit: $Angle(A, B) = -\frac{A^T B}{\|A\| \cdot \|B\|}$ (5.5)

Cette fonction calcule simplement le cosinus de l'angle entre les deux vecteurs caractéristiques A et B. Une valeur élevée de corrélation normalisée correspond à une bonne similarité entre les deux vecteurs.

$$d_k = \min(d_{euclidienne_i}) \text{ pour } i = 1 \dots k. \quad (5.6)$$

Dans les systèmes de reconnaissance nous calculons deux types de distances :

- **La distance_intra** : la distance entre deux images de visage de même individu.
- **La distance_extra** : la distance entre deux images de visage de deux individus différents.

Seuillage : Cette étape consiste à évaluer le résultat de l'étape précédente avec un ensemble de seuil et produit un résultat selon le mode de vérification ou le mode d'identification.

1 Mode de vérification

Lors de la reconnaissance en mode vérification on compare la distance minimale d_k à un ensemble de seuils, l'opération se déroule comme suit :

- Si d_k est inférieure à un seuil s_1 ; alors l'image du visage est considérée comme un **visage connu**.
- Si d_k est supérieure au seuil s_1 , mais inférieure au deuxième seuil s_2 , l'image peut être déterminée comme un **visage inconnu**.
- Si le visage d'entrée est supérieur aux deux seuils, le visage déterminé **n'est pas un visage**.

Dans le problème de vérification d'identité, nous cherchons à définir, pour chaque personne, ou de manière globale un seuil. Ce seuil va déterminer le minimum de ressemblance entre deux images pour admettre qu'il s'agit de la même personne. Ce **minimum de ressemblance** va s'exprimer comme un **distance maximale** entre les caractéristiques des deux images. Pour fixer ce seuil nous allons utiliser l'**ensemble d'évaluation**. Théoriquement $\max(\text{distance intra}) < \min(\text{distance extra})$ dans cette condition, il est possible de déterminer un seuil S qui détermine une classification parfaite.

Il suffit de prendre :
$$S = \frac{\max(\text{distance intra}) + \min(\text{distance extra})}{2} \quad (5.7)$$

Malheureusement cette situation ne se rencontre que très rarement en pratique. Par conséquent, un compromis doit être fait dans le choix du seuil, et ce dernier est très dépendant de l'application: **TFA** doit avoir une valeur très faible dans les applications bancaires afin de garantir plus de sécurité. Généralement, un système d'authentification opère entre les deux extrêmes pour la majorité des applications: on ajuste le seuil de décision pour atteindre une valeur désirée et prédéfinie de **TFR** ou de **TFA**.

Donc nous devons choisir un critère à minimiser pour fixer le seuil de manière à atteindre le **TEE** (**TFA=TFR =TEE**, Taux d'erreur égale). Dans notre expérience nous choisissons le seuil de manière à atteindre **TEE** tel que :

$$(TFA - TFR) < 10^{-2} \quad (5.8)$$

Le programme ainsi conçu se présente sous plusieurs fonctions, l'essentiel est présenté sous la forme suivante :

- *V_{caractéristique}()* pour préparer les images ainsi lues à la phase de traitement.
- *PCA ()* cette fonction renvoie une base de l'ACP normalisée dans laquelle on projette les deux matrices moyennes centrées de chaque classe et de chaque images
- *Fisherface ()* cette fonction renvoie une base LDA dans laquelle on projette toutes les images d'apprentissage et du test.
- *Dis₁ ()* cette fonction renvoie les valeurs des distances entre les images des clients évaluation ou des imposteurs et les images d'apprentissage

2 Mode d'identification

A chaque visage de référence I_k est associée une identité, sous la forme d'un numéro $id(k)$. Nous chercherons à cette étape l'identification d'un visage test à partir des visages de référence. La méthode consiste à comparer le visage test J avec chaque image de référence I_k . La dissimilitude entre les deux est quantifiée par la **distance Euclidienne**: En évaluant cette distance pour chaque visage de référence, on peut déterminer le visage de référence I_k le plus proche du **visage test** J . Si l'erreur d'identification, (distance correspondante) est inférieure à un seuil θ , on peut estimer qu'il s'agit d'un visage reconnu d'identité $id(J) = id(k)$, sinon le visage est classifié comme inconnu :
$$id(J) = \begin{cases} \text{connu} & \text{si } \min d_k(J) \leq \theta \\ \text{inconnu} & \text{sinon} \end{cases} \quad (5.9)$$

Finalement, si le visage est un visage inconnu, on peut décider si on souhaite ajouter à l'avenir le visage pour la reconnaissance. Nous utilisons dans nos expériences plusieurs algorithmes de réduction d'espace, on peut légitimement se poser la question :

Quel algorithme choisir?

Pour répondre à cette question nous effectuons nos mesures sur plusieurs algorithmes et nous retenons le meilleur au sens performance du système. Nous utilisons entre autre les deux méthodes d'extraction des caractéristiques l'Analyse Discriminante Linéaire Bidimensionnelle Orientée en Lignes (ADL2DOL) et en Colonnes (ADL2DOC). Pour ce type d'analyse nous utilisons la distance proposée dans [103] l'AssembledMatrixDistance: AMD (équation (5.10) pour la mesure de similarité.

$$d(Y_i, Y_j) = \left(\sum_{k=1}^{d_1} \left(\sum_{h=1}^n (y_{hk}^{(i)} - y_{hk}^{(j)})^2 \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \text{ avec } p > 0 \quad (5.10)$$

Avec : $y_k^{(i)} = [y_{1k}^{(i)} y_{2k}^{(i)} \dots y_{d_1k}^{(i)}]$ et $y_k^{(j)} = [y_{1k}^{(j)} y_{2k}^{(j)} \dots y_{d_1k}^{(j)}]$ les deux matrices caractéristiques. Les figures ci-dessous illustrent les taux d'égale erreur TEE obtenus sur l'ensemble d'évaluation pour les méthodes (ADL2DOL, ADL2DOC) en appliquant la AMD pour la mesure similarité avec $p=0.125, 0.5, 1, 2$. Le meilleur résultat est obtenu par $p=0.125$ pour les deux méthodes d'extraction des caractéristiques par un nombre restreint des vecteurs propres.

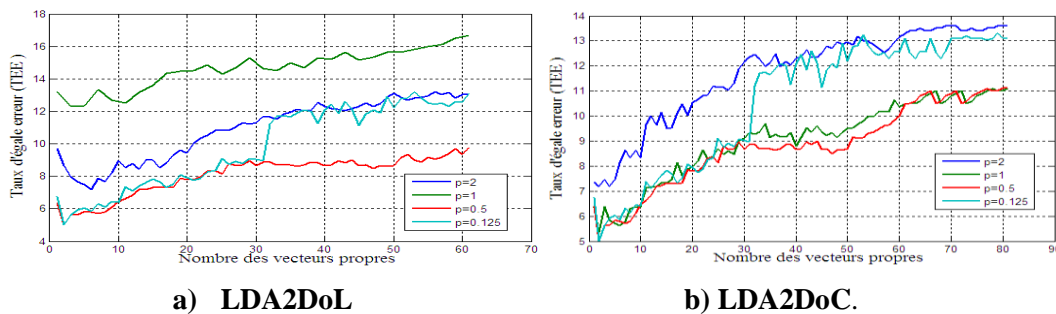


Figure 5.8 TEE dans l'ensemble d'évaluation par différentes fonctions de mesure de distance

Tableau 5.5 Comparaison des performances des méthodes LDA2DOL et LDA2DOC dans l'ensemble d'évaluation et test.

Méthodes	Ensemble évaluation	Ensemble de test			N _p
	TEE%	TFR%	TFA%	TR%	
LDA2DOL	5.1 ±0.98	2±1.37	6.3±0.15	91.7±1.52	2*81
LDA2DOC	5.01±0.97	2±1.37	6.19±0.15	91.81±1.5	2*61

Nous remarquons que le meilleur résultat est obtenu par $p=0.125$ pour les deux méthodes d'extraction des caractéristiques par un nombre restreint des vecteurs propres. Le **tableau 5.5** montre les résultats obtenus dans l'ensemble d'évaluation et l'ensemble de test pour la meilleure distance avec $p=0.125$. La méthode LDA2DOC nous donne le même résultat avec un nombre de caractéristiques $2*61=122$ pour un **TR=91.81±1.5%**.

Le **tableau 5.6** montre les résultats obtenus dans l'ensemble d'évaluation et l'ensemble de test des différentes méthodes de réduction d'espace.

Tableau 5.6 Comparaison des performances des méthodes ACP, ADL, EFM et ADLD dans l'ensemble d'évaluation et de test (intervalle de confiance paramétrique à 95%).

Méthodes	Ensemble évaluation		Ensemble de test		Np
	TEE %	TFR %	TFA %	TR %	
	ACP	5.7 ±1.04	2.7 ± 1.58	8.1 ±0.17	
ACP+LDA	2.83 ±0.74	2 ±0.08	4.12 ±1.94	93.88 ±2.03	90
ACP+EFM	2.66 ±0.13	2 ± 1.37	3.66 ±0.12	94.33 ±1.49	80
ACP+DLDA	3 ±0.72	2.1 ± 1.4	4.3 ±0.12	93.6 ±1.53	195
LDA2DOL	5.1 ±0.98	2 ± 1.37	6.3 ± 0.15	91.7 ±1.52	2*81
LDA2DOC	5.01 ±0.97	2 ± 1.37	6.19 ± 0.15	91.81 ±1.52	2*61

Les taux d'égale erreur TEE obtenus sur l'ensemble d'évaluation des quatre méthodes (ACP, ADL, EFM et ADLD) pour l'authentification des visages en appliquant corrélation normalisée pour la mesure similarité sont représentés sur la **figure 5.9**.

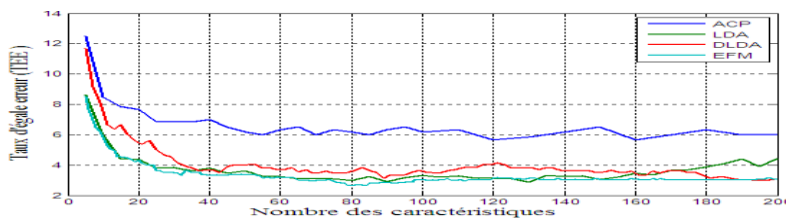


Figure 5.9 Comparaison des performances des méthodes ACP, ADL, EFM et ADLD dans l'ensemble d'évaluation.

La méthode EFM est mieux performante que l'ADL suivie de l'ADLD suivie de l'ACP en employant un nombre restreint de caractéristiques (vecteurs propres) dans notre base de données XM2VTS. Nous constatons que le résultat obtenu avec ACP+EFM est le meilleur avec un TEE=2.66±0.13% et un TR=94.33±1.49% en utilisant seulement un nombre de paramètres caractéristiques Np = 80. Nous remarquons aussi que la DLDA ne dépasse pas de trop les performances de la LDA et le mauvais résultat est présenté par la PCA. Les deux méthodes LDA2DOL et LDA2DOC donnent presque le même résultat avec un nombre de paramètres caractéristiques restreint. Nous remarquons aussi que ces méthodes à deux dimensions n'améliorent pas la performance d'authentification de visage par rapport aux ADL, EFM et ADLD car la base de données XM2VTS a été réalisée sur une longue période

ce qui a permis d'obtenir plusieurs images d'une même personne et donc une grande variabilité d'apparence (changement de coiffure, présence et/ou absence de lunettes, etc...). Par contre, seules les expressions neutres ont été considérées, donc *nous proposons d'utiliser ces méthodes (LDA2DOL et LDA2DOC) pour les bases de données qui contiennent plusieurs expressions.*

L'algorithme de réduction d'espace retenu est donc l'ACP+EFM et la mesure de similarité utilisée est la corrélation.

5.5.2 Méthode retenue ACP+EFM associée au SVM [335]

Dans cette partie de nos travaux nous retenons que la méthode ACP+EFM est la meilleure méthode de réduction d'espace. Nous l'utiliserons donc dans ce qui suit.

La méthode à vaste marge SVM est utilisé comme méthode de classification à deux classes. Pour cela, nous proposons d'utiliser un nombre de SVMs égale au nombre des Clients (200 Clients dans le cas de la XM2VTS). La figure 5.10 montre l'apprentissage des SVMs. On voit bien que les deux entrées du premier SVM₁ sont composées par :

- trois vecteurs caractéristiques réduits associés aux trois poses du Client₁,
- deux cents vecteurs caractéristiques réduits des huit poses des vingt cinq Imposteurs.

Pour les deux cent SVMs, il n'y a qu'une seule entrée qui varie c'est celle correspondant aux clients : SVM₁ est associée au Client₁, SVM₂ au Client₂,..... SVM₂₀₀ au Client₂₀₀. Les 200 vecteurs caractéristiques réduits des 8 poses des 25 imposteurs sont maintenus pour les 200 SVMs. Les vecteurs étiquettes sont les mêmes pour chaque SVM avec, -1 correspondant à la classe des Imposteurs et +1 à la classe des Clients.

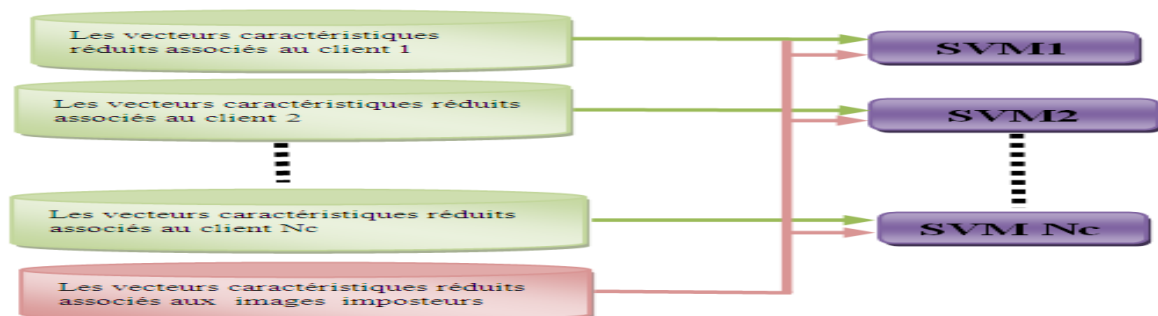


Figure 5.10 Apprentissage des SVM

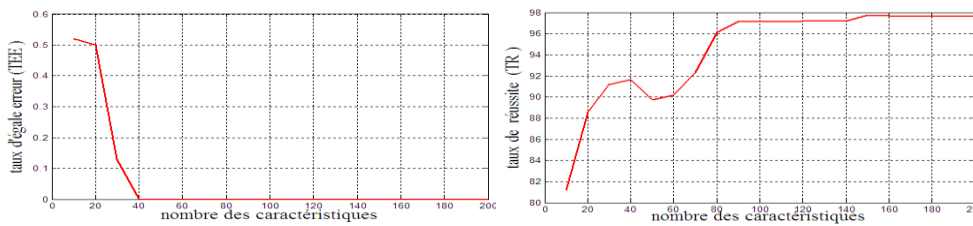
Chaque client est caractérisé par trois descripteurs correspondants aux trois poses (600 vecteurs). Alors que les imposteurs sont représentés par 25x8 descripteurs (25x8 images).

Nous choisissons d'utiliser une machine à vecteurs de support (SVM) avec un noyau RBF (Radial Basis Function ou fonction à base radiale) (voir équation 5.11).

Le SVM a été implémenté en utilisant la librairie **LibSVM** disponible gratuitement (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>). Le noyau **RBF** utilisé est de la forme :

$$K_{RBF}(u, v) = e^{-\gamma \|u-v\|^2} \quad (5.11)$$

Où γ est un paramètre qui règle la largeur de la marge. Le SVM nécessite également le réglage du paramètre C qui règle le compromis entre la largeur de la marge et le nombre d'erreurs admissibles. Les figures ci-dessous présentent les taux d'égale erreur (**TEE**) dans l'ensemble d'évaluation et taux de réussite (**TR**) dans l'ensemble de test en fonction du nombre des caractéristiques de la méthode de réduction d'espace **ACP+EFM**. La méthode de classification est celle qui utilise un nombre des **SVMs** égale au nombre des Clients.



a) **TEE** dans l'ensemble d'évaluation b) **TR** dans l'ensemble de test.

Figure 5.11 Taux en fonction des caractéristiques de l'ACP+EFM associée au SVM.

Nous remarquons que le **TEE** s'annule à partir d'un un nombre de caractéristiques égale à 40 c'est un **TEE** parfait et cela dans l'ensemble d'évaluation. Mais d'après la figure **5.11** nous remarquons que dans l'ensemble de test le **TR** prend une valeur intéressante et stable depuis un nombre de caractéristiques égale 90.

Nous présentons dans le **tableau 5.7** les résultats dans l'ensemble d'évaluation et l'ensemble de test, le nombre des caractéristiques et le temps de calcul pour la méthode proposée. Notre travail est exécuté avec un ordinateur : Intel Pentium Dual CPU 2.2 GHz, 1.49 Go de RAM et le langage de programmation utilisé est le **MATLAB R2009b**.

Tableau 5.7 Performance de notre système d'authentification de visages.

Ensemble d'évaluation	Ensemble de test			Np	Tps de calcul
	TFA%	TFR %	TR %		
0±0	1.5±0.09	0.78±0.01	97.72±1.46	150	0.32 s

Le descripteur donne un TEE= 0±0 % dans l'ensembled'évaluation. La méthode proposée est parfaite dans la phase d'évaluation. L'erreur augmente dans l'ensemble de test.

Nous pouvons affirmer à ce stade que dans ce cas de **classification par SVM**, nous obtenons un parfait taux d'égale erreur avec **TEE=0±0%** dans l'ensemble d'évaluation et un bon taux de réussite **TR=97.72±1.46%** dans l'ensemble de test avec un temps de calcul Tc satisfaisant **t=0.32s**. Cette méthode est plus performante que celle qui utilise la **classification** basée sur la **distance** de corrélation normalisée avec un : **TEE=2.66 ±0.13% et TR=94.33 ±1.49%**.

1) Tableau comparatif avec notre approche

D'après cette série de résultats, un tableau de synthèse (**tableau 5.8**) permet de comparer la performance de notre système de reconnaissance de visages par rapport à la performance obtenue dans d'autres travaux sur la base de données **XM2VTS** selon le protocole de **Lausanne I**.

Tableau 5.8 Tableau de synthèse de performances, obtenues dans l'authentification de visages.

Dates	Auteurs	Méthodes	Évaluation(%)	Test (%)
2006	Saigaa[337-1]	couleur ACP+EFM+RN	TEE=1.51	TFA=1.49,TFR=1.5,TR=97.01
2007	Yan[337-2]	Weighted Variant for Effectiveness	TEE=2.37	/
2007	Kotsia[337-3]	Discriminant Non-negative Matrix Factorization Algorithms	TEE=0.7	/
2009	Ejarque[337-4]	Fisherfaces +SVM	TEE=1.25	/
2009	Salimi[337-5]	KPCA+LDA	TFA=1.08, TFR=1.17	TFA=1.23,TFR=1,TR=97.77
2011	Nos systèmes	ACP+EFM+corrélation PCA+EFM+200 SVMs	TEE=2.66±0.13 TEE=0	TR=94.33±1.49 TFA=1.5,TFR=0.78,TR=97.72

Nous pouvons dire que la performance de notre système d'authentification de visages est presque équivalente à celle obtenue en **2009** par **Salimi** dans l'ensemble de test. En plus notre système est parfait dans l'ensemble d'évaluation. Notre objectif était l'amélioration du système de reconnaissance de visages nous affirmons en nous basant sur les résultats obtenus que notre objectif est bien atteint. De cette étude comparative de six méthodes de réduction d'espace, il en ressort que les méthodes d'extraction des caractéristiques : **ADL**, **EFM** et **ADLD** sont plus performantes que l'**ADL2DO** et la **meilleure méthode de réduction d'espace** reste l'**ACP+EFM** [336]. En utilisant un nombre de **SVMs** égale au nombre des Clients nous obtenons une bonne efficacité du système justifiée par un **TEE** parfait dans l'ensemble d'évaluation, un bon taux de réussite avec **TR=97.72±1.46%** dans l'ensemble de test et un temps de calcul égal **0.32**seconde.

Deuxième expérience : Cas de méthode de réduction non linéaire KPCA

Notre système ainsi conçu est associée à la **KPCA**. Nous utilisons pour ce cas deux classifieurs. Le premier basé sur une mesure métrique et le deuxième est assuré par **SVM**.

1 Architecture de l'approche à base de KPCA

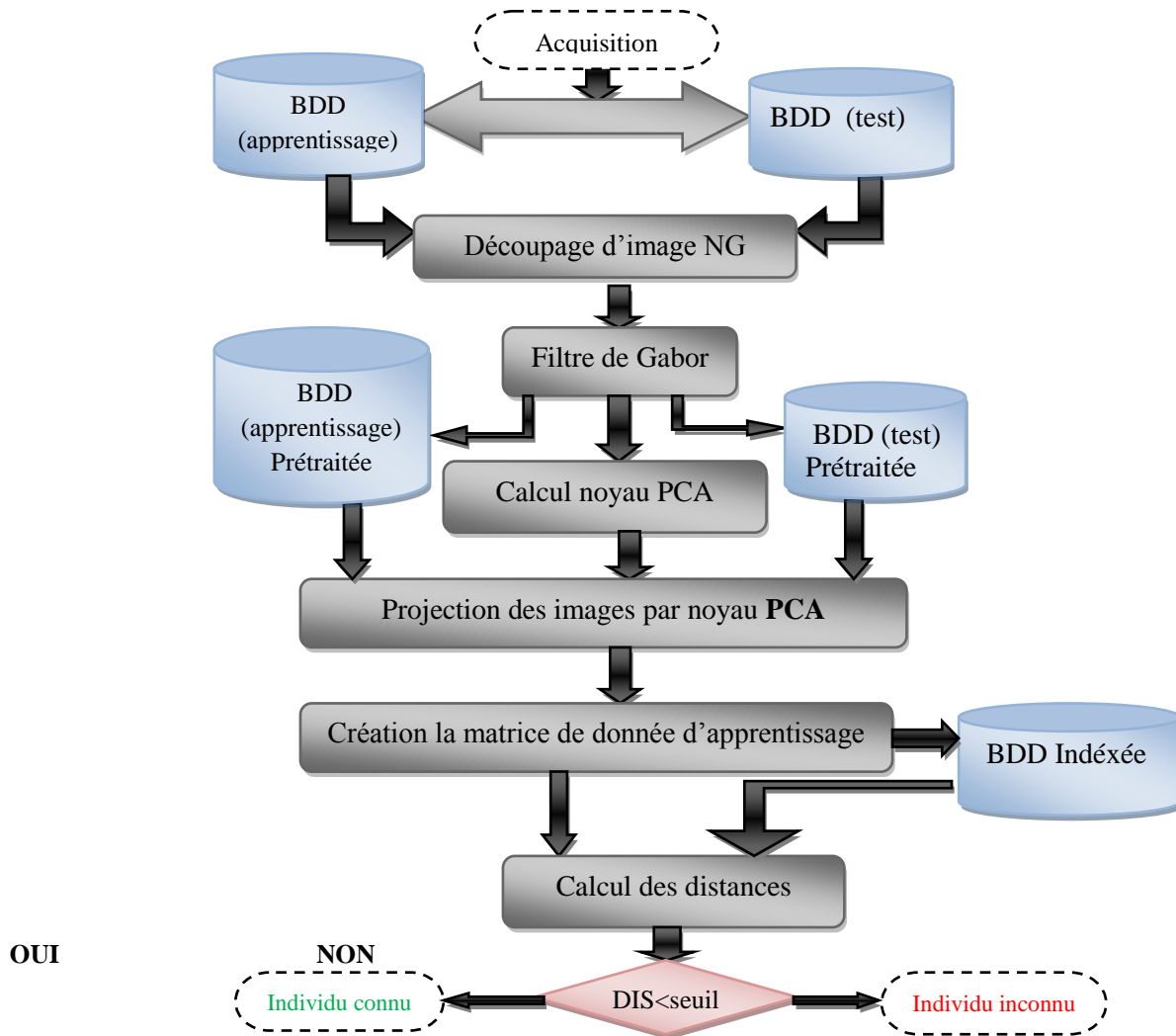


Figure 5.12 Organigramme du système de vérification d'identité de visage

1.1 L'algorithme de KPCA

```

Procédure de calcul KPCA /* Algorithme de KPCA*/
Fonction noyau Gaussien
  x,y : deux coord.points
  Param=0.5/sigma*sigma /*paramètre de la fonction kernel, sigma: la largeur du noyau Gaussien*/
Function k=kernel (x,y,param)
  Diff=x-y;
  K = exp(-(dif*diff)*param);
functionKPCA (data,sigma,numev,outlier)*
  [n,d]=size (data); /*matrice de n lignes et m colonnes
  *centrer les données de la matrice noyau *
  K=zeros(n,n);
  *Calcul de la matrice de covariance*
  Calculer la matrice de covariance c'est le produit de la matrice centrée et sa transposée.
  *Calcul des valeurs et vecteurs propres*
  Calculer les valeurs propres et les vecteurs propres par la fonction eig();
  *Tri des valeurs propres par ordre décroissant *
  Trier les valeurs propres λ dans l'ordre décroissant avec la fonction
  sort();
  fin
  
```

Mesures de similarités

Une fois que les images sont projetées dans un sous espace il reste à déterminer quelles sont les images, semblables...Il y a beaucoup de mesures possibles de distance et desimilitude. Dans ce cas la distance L_2 est retenue. Connue aussi sous le nom norme euclidienne, elle est définie comme suit :c'est la somme de la différence carré entre les composantes des deux vecteurs A et B. Elle est donnée par l'équation suivante:

$$L_2 = \sum_{i=1}^N (A_i - B_i)^2 \quad (5.12)$$

1.2 Etude comparative

a) Comparaison entre PCA et KPCA(avec normalisation et classification par mesure de métrique L_2)

Tableau 5.9 La comparaison entre PCA et KPCA classification métrique L_2 .

N_p		10	100	400	500	600
TR(%)	PCA	85.9	86.2	86.01	86	86
	KPCA	89.92	91.7	91.51	92.33	93.51

D'après les résultats illustrés dans le **tableau 5.9** le **TR** obtenu pour la **KPCA** est nettement mieux que le celui donné par la **PCA**. La courbe suivante nous montre cette différence qui est loin d'être négligeable. Elle est accentuée par le nombre de paramètres caractéristiques N_p .

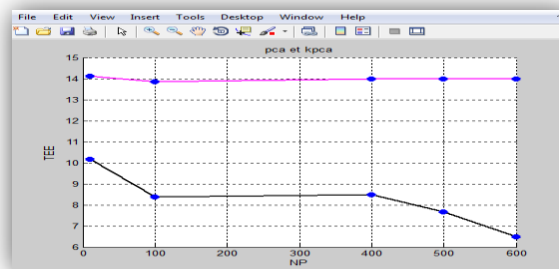


Figure 5.13 Les courbes du **TEE** pour la **PCA** et la **KPCA** en fonction N_p . (avec normalisation et mesure de métrique L_2).

Les courbes de la figure **figure 5.13** montre le grand écart entre les deux méthodes du taux d'égale erreur le **TEE**. Cet écart atteint **8.51%** ce qui est considérable dans le domaine de reconnaissance en biométrie.

b) Classification par SVM

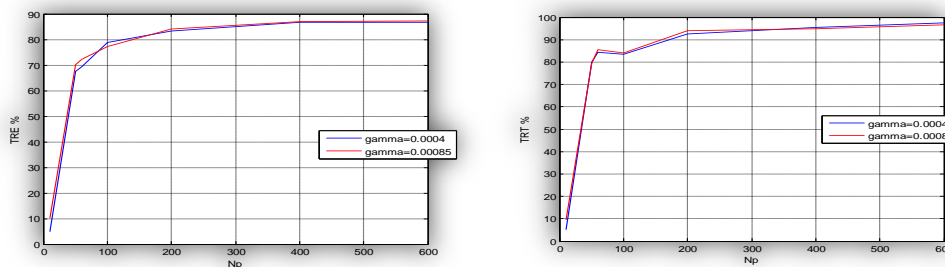
Pour $N_p = 600$ et σ variable dans les deux ensembles :

Sigma (σ) : Largeur du noyau **Gausséen**.

Tableau 5.10 Taux de reconnaissance classification SVM

Résolution(σ)	0.0004	0.0007	0.00085	0.0009	0.003	0.01
N_p	600	600	600	600	400	400
TR évaluation%	86.83	87	87.33	87.33	86.83	82.83
TR test%	97.5	97	96.5	96.5	95.5	94

Dans ce tableau nous constatons que le meilleur $TR_{test} = 97.5\%$ avec résolution $\sigma = 0.0004$ et $N_p = 600$ c'est un bon résultat.



a) Ensemble d'évaluation

b) Ensemble de test

Figure 5.14 Taux de reconnaissance en fonction de N_p pour diverses σ . (Classification par SVM)

Tableau de KPCA et SVM

Comparaison entre le KPCA et la PCA avec différentes classifications

Tableau 5.11 Comparaison entre le **KPCA** et **PCA** pour différentes classifications

		Classifieurs	
		Distances	SVM
TR(%)	PCA	86.2	90.18
	KPCA	93.51	97.5

Nous remarquons que le TEEs'annule à partir d'un un nombre de caractéristiques égale à $N_p = 40$ c'est un TEEparfait et cela dans l'ensemble d'évaluation. Mais d'après la **figure 5.14** nous remarquons que dans l'ensemble de test le **TR** prend une valeur intéressante et stable depuis un nombre de caractéristiques égale $N_p = 90$. En se basant sur les résultats de Salimi en 2009 voir **tableau 5.8** qui arrive à $TR=97.77\%$ en utilisant la KPCA +LDA, nous pouvons affirmer que la KPCA est une approche stable et prometteuse. Notre approche est plus simple et performante. Nous arrivons à un $TR=97.55\%$ en utilisant seulement la KPCA, en plus l'outil SVM est très efficace et facilement implémentable.

5.6 Approche monomodale 3D

5.6.1 Description de la BDD CASIA 3D expression

Entre Août 2004 et Septembre 2004, on a recueilli une base de données de visages 3D constituée de 4624 scans de 123 personnes qui utilisent numériseur 3D, Minolta Vivid 910, comme indiqué dans la figure 5.15 a). Au cours de la construction de la base de données, on considère non seulement les variations individuelles de poses, les expressions et les illuminations, mais aussi les variations combinées des expressions sous illumination et poses sous des expressions, comme montré sur la figure 5.15 b) et 5.15 c). Pour les sujets avec des lunettes, on a recueilli un scan supplémentaire avec des lunettes. Ainsi, chaque personne dispose de 37 ou 38 balayages. Et à partir de chaque balayage, une image couleur 2D et une surface 3D du visage triangulaire sont également générés. On vise à construire une base de données complète du visage en 3D, qui est en outre amenés à être une plate-forme publique pour tester les algorithmes de reconnaissance faciale 3D ou autres[9].

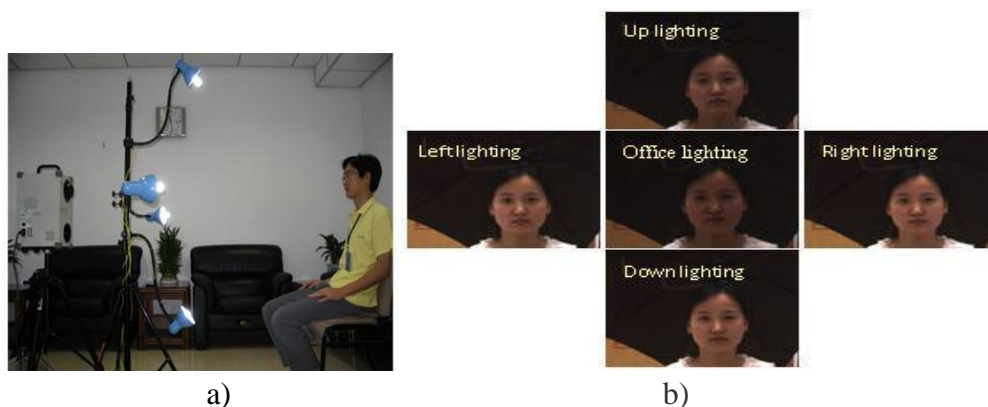
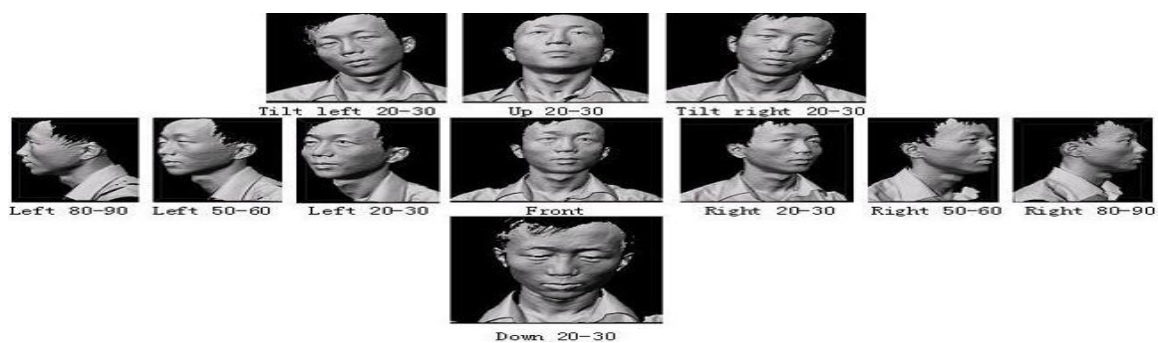


Figure 5.15 a) Capture des scènes de la base de données CASIA.

b) Variations d'éclairage de la base de données CASIA visage 3D.



c) Variation des poses de la base de données CASIA visage 3D.

- **Base d'apprentissage** : contient l'information concernant les personnes connues par le système (seulement des clients). Pour cela on prend les M premières images de 1

jusqu'à M de chaque personne. La **BDD CASIA 3D** expressions contient des images de visages présentant des expressions faciales dans l'ensemble d'apprentissage. Ce sous-ensemble contient des variations de pose des visages, des variations d'éclairage ainsi que des expressions faciales.

- **Base de test** : permet de tester le système en lui présentant des images de personnes imposteurs et clients les N dernières images de x à y de chaque individu nous ont servies pour la réalisation des différents tests.
- **Stockage et nomenclature** : Toutes les images sont stockées sous la nomenclature « $xx-yy.wrl$ ».

' xx ' : représente le numéro de la personne (de « 00 » à « 123 » pour les deux sexes).

' yy ' : représente numéro de l'image de « 01 » à « 37 ».

5.6.2 Système de reconnaissance de visage 3D

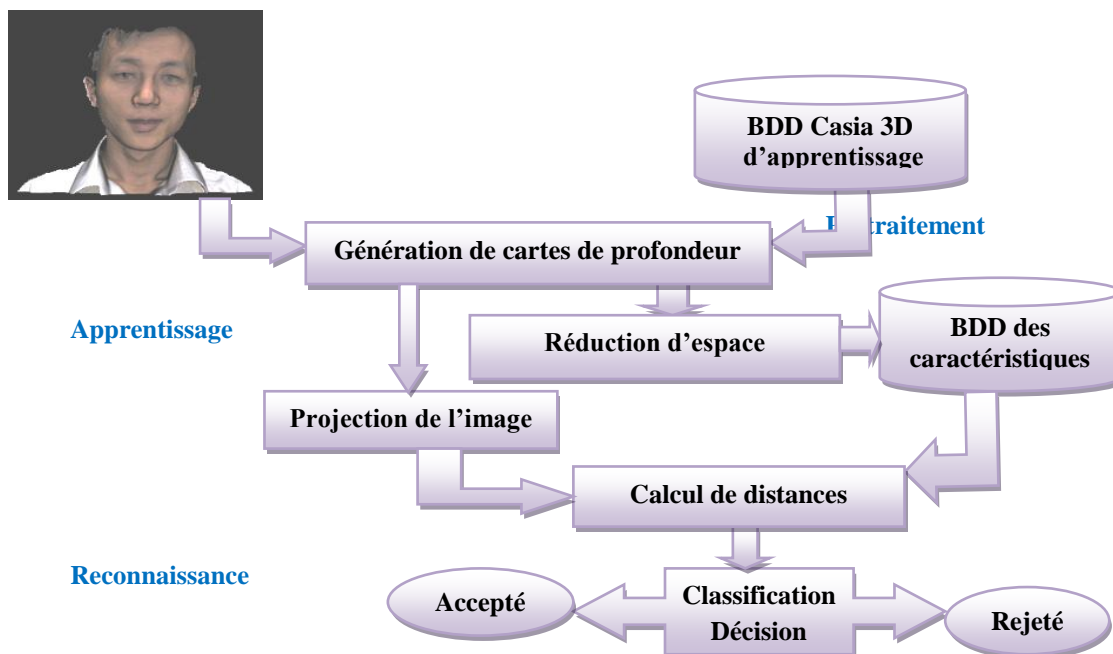


Figure 5.16 Organigramme du SRV 3D (basé sur la génération de cartes de profondeur).

1 Phase de prétraitement

1.1 Génération des cartes de profondeur

Comme nous l'avons déjà vu, la base de données utilisée est un ensemble de fichiers VRML contenant des maillages triangulaires. La façon d'exploiter l'information 3D, des caractéristiques géométriques d'un visage humain, est la construction de la carte de profondeur. La carte de profondeur est la fonction donnant pour chaque pixel de l'image la

profondeur du point 3D correspondant. Les étapes de génération de cartes de profondeur sont les suivantes :

- Préparation de la carte de profondeur ;
- Normalisation des Z-coordonnées ;
- Chargement de la carte par les coordonnées Z normalisées ;
- Rotation ;
- Découpage ;
- Affichage de la carte de profondeur.



Figure 5.17 Cartes de profondeur avant le découpage

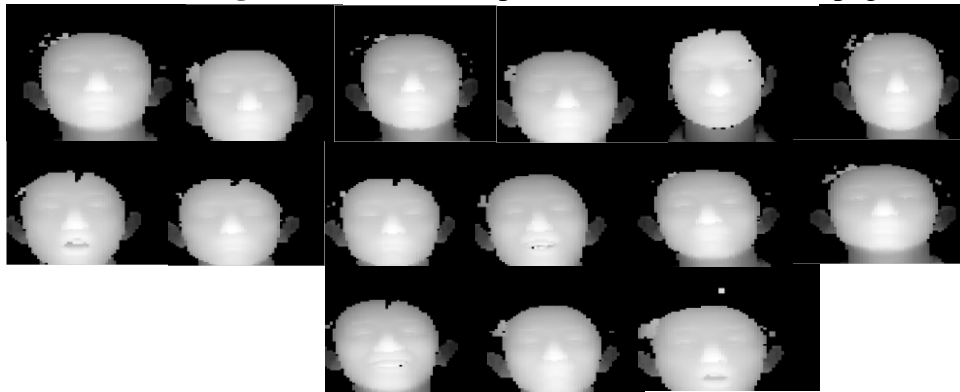


Figure 5.18 Cartes de profondeur après le découpage

2 Phase d'apprentissage

Le processus d'apprentissage a pour but de recueillir des informations biométriques sur les personnes à identifier. Il prend comme entrée les cartes de profondeur filtrées générées à partir de la base de données d'apprentissage et produit comme sortie les k premiers grands vecteurs propres en se basant sur l'ACP.

3 Phase de reconnaissance

Le processus de reconnaissance se subdivise en trois étapes : l'image d'entrée $I(N \times N)$ contenant le visage à identifier est transformé en un vecteur Γ de dimension $(N^2 \times 1)$ puis comparé avec le vecteur moyen Ψ . La différence entre le vecteur Γ et le vecteur moyen est ensuite multipliée par l'ensemble des vecteurs propres e_k pour générer le vecteur de poids Ω^T ,

$$\omega_k = e_k^T (\Gamma_i - \Psi), \quad k = 1, \dots, M \quad (5.13)$$

$$\Omega^T = [\omega_1, \omega_2, \dots, \omega_M]^T \quad (5.14)$$

5.6.3 Résultats et discussions

Afin de mieux étudier l'apport de notre méthode, nous allons valider les résultats obtenus avec une méthode de base et ces résultats serviront par la suite de comparaison.

Pour cela nous faisons varier les paramètres suivants:

- Coefficients : les coefficients de projection des vecteurs propres (triés suivant les valeurs propres décroissantes).
- Les nombres des paramètres de l'ACP et de LDA : pour l'ACP le nombre de paramètre maximum est égal à $N_{eapp} * N_c$ et pour la LDA la dimension maximale est égale à N_c (N_c est le nombre des personnes dans la base apprentissage et N_{eapp} c'est le nombre des images (poses) de chaque personne dans la base d'apprentissage).
- Mesure de score (similarité) : corrélation normalisée (l'angle).
- Classification globale : en utilisant l'algorithme de calcul de TFA et TFR.

Nous testons notre approche sur différents aspects de la BDD CASIA 3D. Nous menons trois expériences essentielles sur une personne sous trois variantes: changement d'illumination, changement d'expressions et changement d'illumination et expression.

1 Changement d'illumination

Pour un nombre de personnes égal à **50** et un nombre d'images pour chacune égal à **5** ; **3** images pour la base d'apprentissage et **2** images pour la base du test.

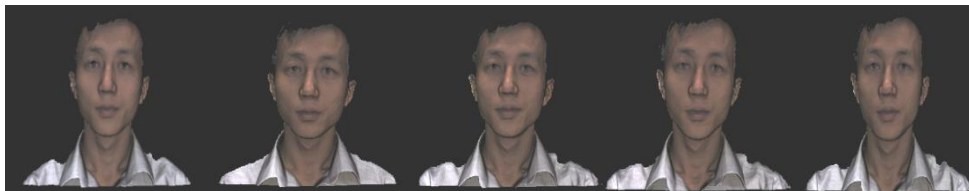


Figure 5.19 Les cinq visages de changement d'illumination.

Meilleur résultat d'identification : TR= 88 % pour $N_p(ACP)=20$ et $N_p(LDA)=45$

Les meilleurs résultats du taux de reconnaissance sont obtenus dans le cas de l'intervalle N_p de l'ACP [25,50] quelque soit $N_p(LDA)$. Pour $N_p(ACP) > 50$ le taux chute de réussite considérablement. Alors que le système reste stable concernant le nombre de paramètres $N_p(LDA)$. Nous pouvons affirmer à ce stade que *les paramètres de l'ACP ont une grande influence sur la performance du système et le grand nombre de paramètres caractéristiques est non sollicité.*

2 Changement d'expression faciale

Pour un nombre de personnes égal à **50** et un nombre d'images de chacun d'elles égal à **10** ; **6** images dans la base d'apprentissage et **4** images pour la base du **test**. Nous notons que:

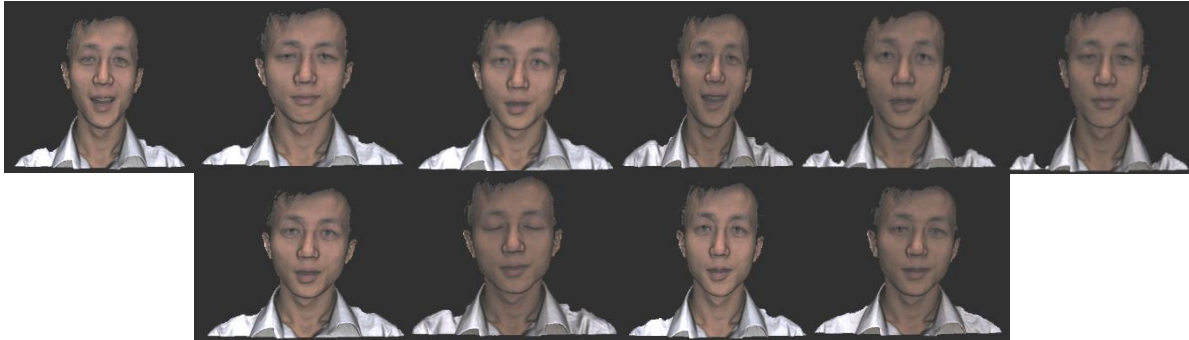


Figure 5.20 Les dix visages de changement d'expressions faciales

Résultats d'identification : $TR = 84,5\%$ pour $N_p(ACP) = 45$ et $N_p(LDA) = 40$. Notre approche est **performante** pour l'intervalle $N_p(ACP)[25,50]$. Nous remarquons une **stabilité** du système pour $N_p(LDA) > 30$ et un $N_p(ACP) < 100$. Pour $N_p = 100$ le taux de réussite (**TR**) est **trop faible** et le **système** devient **instable** et ceci **quelque soit** $N_p(LDA)$. *L'expression faciale influe beaucoup sur le taux de réussite relativement au changement d'illumination.* Nous constatons que **notre approche est vulnérable pour ce cas précis**. Cette différence réside à notre avis dans le fait de normalisation dans l'étape génération de cartes de profondeur. Et ceci malgré un échantillon d'images apprentissage et test important. Pour cette raison nous jugeons bon de faire des tests sur la présence d'illumination et expressions faciales en augmentant la taille de l'échantillon images soumis dans les deux phases.

3 Changement d'illumination avec changement d'expressions faciales

Pour un nombre de personnes égal à **50** et un nombre d'images pour chacune d'elles égal à **15** ; **10** images dans la base d'apprentissage et **5** images pour la base du **test**.

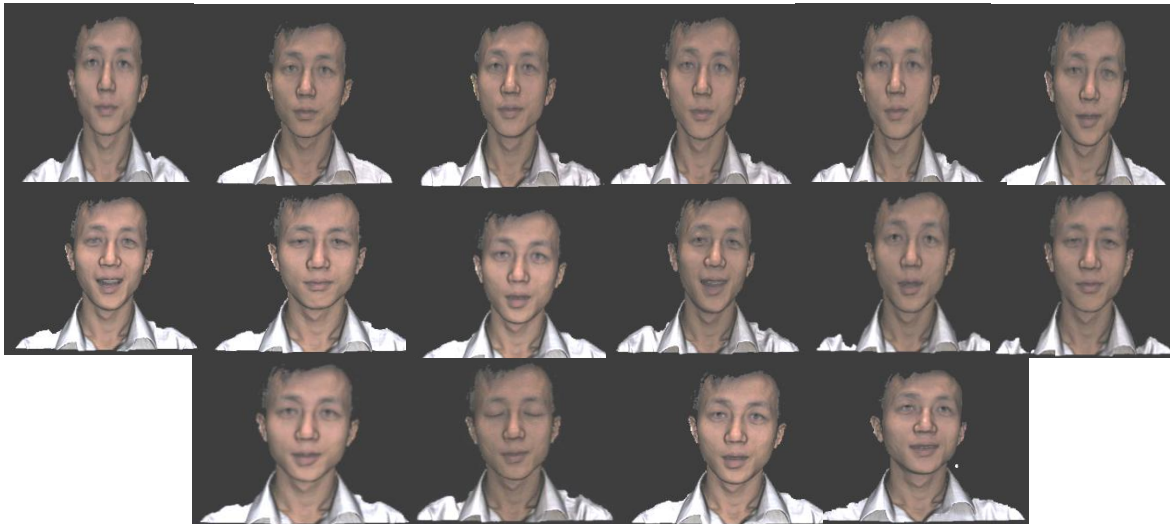


Figure 5.21 Visages avec changement d'illumination et d'expressions faciales.

Meilleur résultat d'identification : TR = 91,5% pour $N_p(\text{ACP}) = 30$ et $N_p(\text{LDA}) = 20$

Notre approche répond mieux à cette dernière expérience. Nous pensons que ceci peut être dû à l'échantillon plus important des images soumis à l'apprentissage et au test. Nous remarquons que dans l'intervalle de nombre de paramètres $N_p(\text{ACP})$ [25,50], le système est performant alors que pour $N_p(\text{ACP})$ supérieur à 50 il devient instable avec des taux de réussite très faible. Le taux de réussite obtenu reste non satisfaisant pour cela nous tentons dans ce qui suit étendre notre recherche et nous intéresser aux visages avec expressions. Nous prenons en compte les régions d'intérêts (yeux, nez et bouche) et ceci dans le souci d'améliorer la performance du système de reconnaissance.

5.7 Approche monomodale 3D expressions (Fusion des régions d'intérêts)

Un échantillon de visages de la **BDD CASIA 3D expressions** et le protocole utilisé dans notre approche est présenté dans la **figure 5.23**.

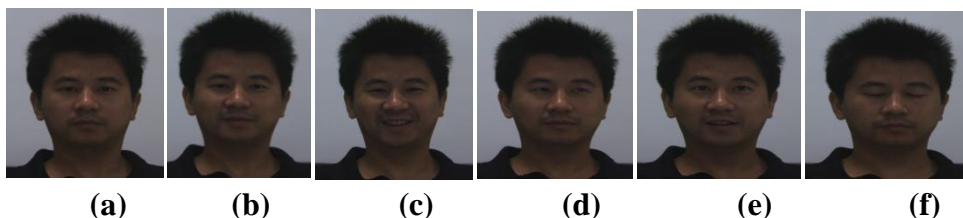


Figure 5.22 Variations d'expression de la base de données CASIA:

(a) visage neutre, (b) sourire, (c) joie, (d) colère, (e) surprise, (f) près des yeux.

5.7.1 Protocole de la BDD CASIA 3D expressions

La **CASIA 3D expressions** est partagée en trois sous ensembles selon le protocole suivant.

- *Apprentissage* de l'image (a),(c),(d) dans la base CASIA l'image 1,7,8.

- *Evaluation* de l'image (b),(e) dans la base CASIA l'image 6,9.
- *Test l'image*(f) dans la base CASIA l'image 10.

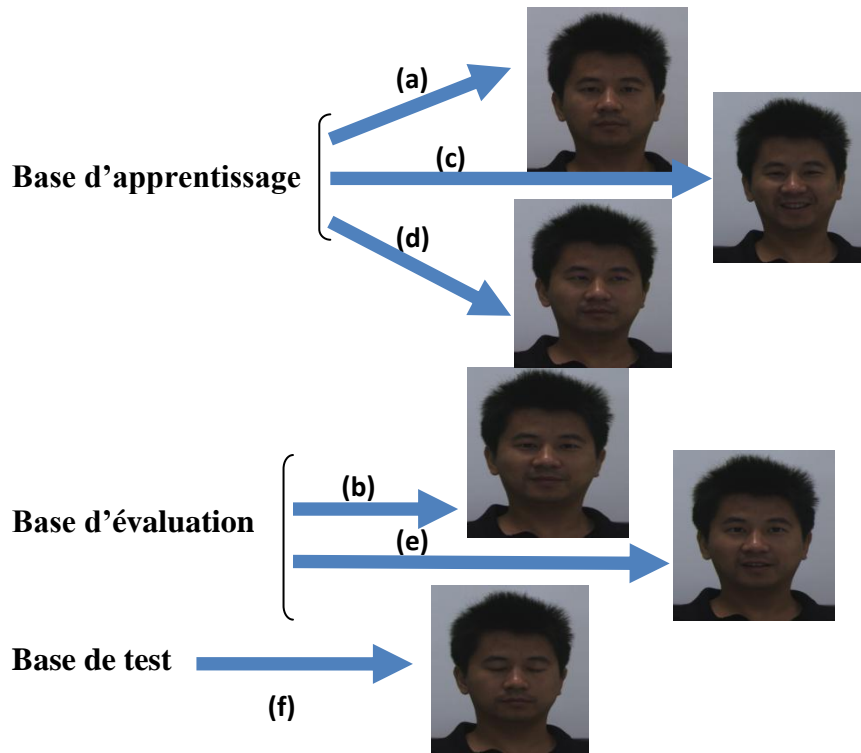


Figure 5.23 Protocole choisi pour notre SRV

5.7.2 Architecture structurelle du système de conception

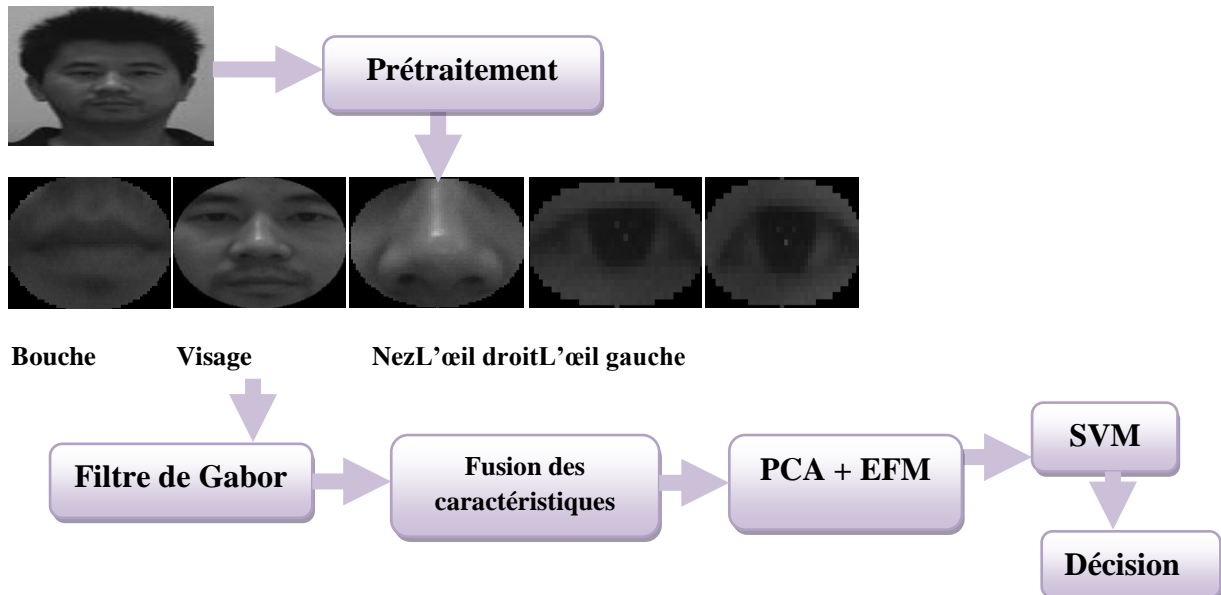


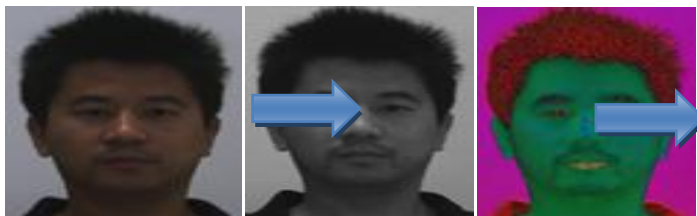
Figure 5.24 Schéma de l'architecture structurelle du système de conception SRV 3D expressions (fusion des régions d'intérêts).

1 Prétraitement d'image

Pour notre travail nous avons besoin de trois phases de prétraitement : la décimation, la conversion de la couleur en l'espace de couleur TSV puis la détection de visage et des régions d'intérêts.

1.1 La composante V de l'espace de couleur TSV de l'image

Les systèmes d'identification de visages utilisent souvent la représentation en niveaux de gris de l'image de visage comme caractéristique d'entrée. Mais lorsque les images sont en couleurs, il est avantageux d'utiliser cette information supplémentaire pour améliorer les performances de ce système d'identification. Nous rappelons que nous utilisons en général l'espace de couleur TSV (Teinte, Saturation, Valeur) ou HSV (Hue, Saturation, Value) ce sont les deux systèmes les plus utilisés dans la littérature [Mel09]. La quantification de la couleur de notre base de données CASIA est faite dans le système RVB (Rouge, Vert et Bleu), le passage du système de couleurs RVB au système TSV se fait par une fonction de transfert non linéaire. Pour plus de détails sur cette transformation, voir [Raf02]. Dans ce cas de base de données CASIA nous utilisons les espaces de couleurs TSV et nous trouvons que la composante V a donné le meilleur résultat. Partant de ce résultat nous utilisons donc la composante V de l'espace de couleur TSV comme caractéristique de l'image. La figure 5.25 présente la conversion de couleur et l'extraction de la composante V de visage.



(a) Image couleur (b) Image en système TSV (c) Composante V

Figure 5.25 Extraction de la composante V du système couleur TSV

1.2 Détection de visage et de points et régions d'intérêts

1.2.1 Détecteurs de visages et points d'intérêts

Avant d'entamer toute opération d'extraction de paramètres caractéristiques de notre image. Nous procédons tout d'abord à la détection de ce qui est significatif dans l'image du visage captée c'est-à-dire nous détectons le visage proprement dit. Car les cheveux, fond, la peau présentent peu d'intérêt pour l'analyse et l'identification. La détection du visage est la première phase de prétraitement et la figure nous montre le résultat de notre détection sur une image de visage neutre de la base CASIA 3D expression utilisée.

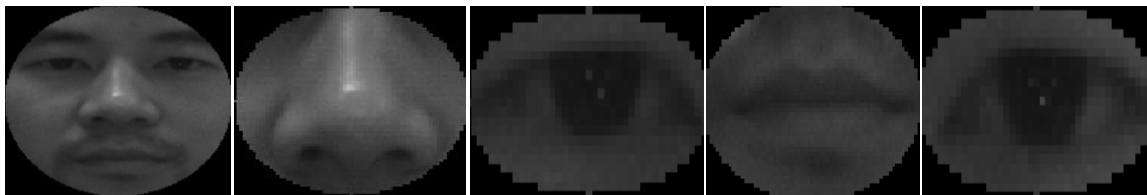


(a) Composante V de TSV (b) Détection manuelle des yeux (c) Détection de visage

Figure 5.26 Détection du visage.

1.2.2 Détecteurs des régions d'intérêts

Les régions d'intérêts sont illustrées par la **figure 5.27**.



(a) Détection de visage (b) Détection de nez (c) Détection de bouche (d) Détection d'œil droit (e) Détection d'œil gauche

Figure 5.27 Détection des régions d'intérêts du visage.

1.3 Extraction des caractéristiques par filtre de Gabor

Cette partie repose principalement sur l'extraction des caractéristiques du visage par filtrage de Gabor 2D.

1.4 Réduction d'espace

La réduction d'espace est réalisée par l'ACP+EFM. Finalement nous réalisons notre base de données indexée dans la phase enrôlement selon la **figure 5.28**.



Figure 5.28 Schéma de la phase enrôlement de notre modèle RV.

1.5 Classification de scores avec les machines à vecteurs de support (SVM)

Les images originales des personnes de la base de données **CASIA 3D expressions** subissent la phase de prétraitement, puis des descripteurs sont extraits et ensuite projetés sur des espaces réduits et stockés dans une nouvelle base de données indexée. Celle-ci est répartie sous le protocole décrit précédemment. La base d'apprentissage nous servira dans le calcul de la méthode de réduction et à l'entraînement de nos **SVMs** dans leur phase d'apprentissage. Une fois cette phase est effectuée nous procédons à l'étape classification par **SVM** en utilisant la base d'évaluation et test de la **BDD CASIA 3D expressions**. Le schéma de ce processus est présenté dans la **figure 5.29**.

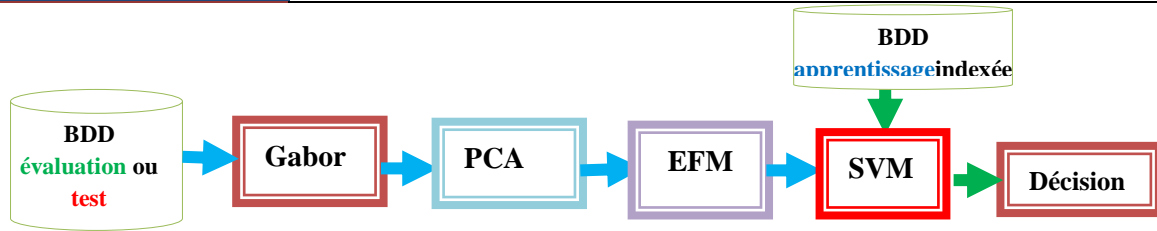


Figure 5.29 Schéma de principe de la classification par SVM.

Ainsi notre étude sur la classification des visages par SVM est menée sur deux phases :

1. **Phase d'apprentissage** : Pour notre protocole de la base de données CASIA 3D expressions, nous avons 123 images de personnes.

D'où : $N_c=123$ classes $\Rightarrow 123$ SVMs puisque le Nombre de classes = Nombre de SVMs. Chaque personne représente 3 échantillons $\Rightarrow N_e=3$. Nous choisissons le numéro d'échantillons d'apprentissage égal 3. Alors $N_{e_{app}}=3$.

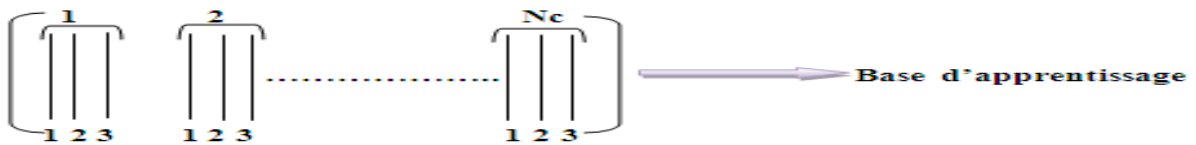


Figure 5.30 Schéma représentatif de la base d'apprentissage.

Le schéma de la phase de test est comme suit :

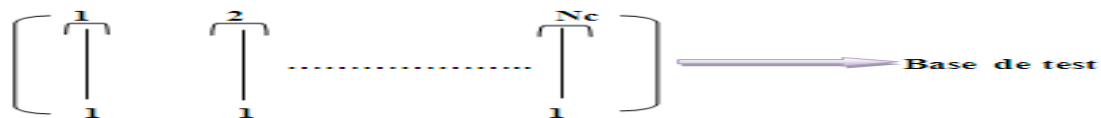


Figure 5.31 Schéma représentatif de la base de test.

Donc la phase d'apprentissage par SVMs sera :

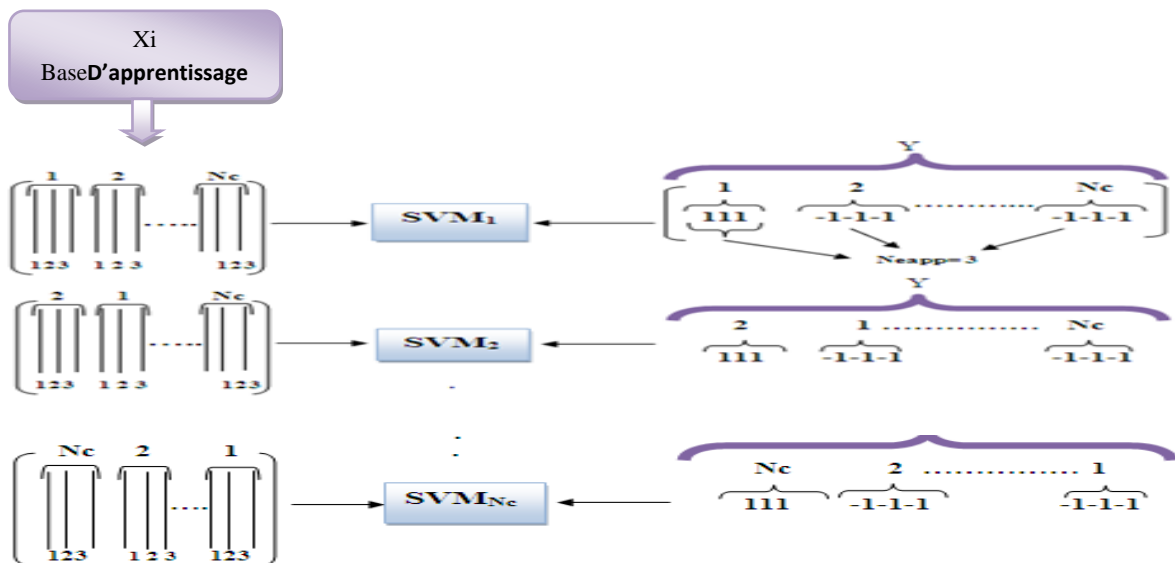


Figure 5.32 Schéma de la phase d'apprentissage par SVMs.

D'après la **figure 5.32** nous constatons que **123 SVMs** sont utilisés dans la phase apprentissage un SVM pour chaque classe (personne).

2. Phase d'évaluation : Dans cette phase nous présentons à l'entrée des **123 SVMs** le vecteur caractéristique réduit et nous sauvegardons l'indice de max des scores. Ainsi nous obtenons **2x123** scores/personnes d'où deux indices max/personne. Ces indices nous permettent de calculer notre taux de réussite **TR**.

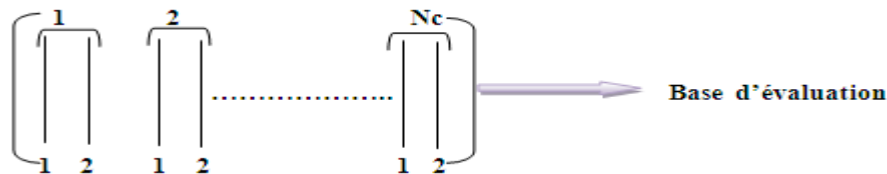


Figure 5.33 Schéma représentatif de la base d'évaluation.

Le schéma de la phase d'évaluation est :

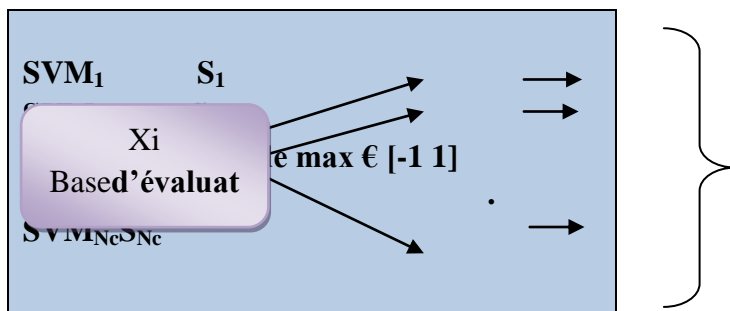


Figure 5.34 Schéma de la phase d'évaluation par SVMs

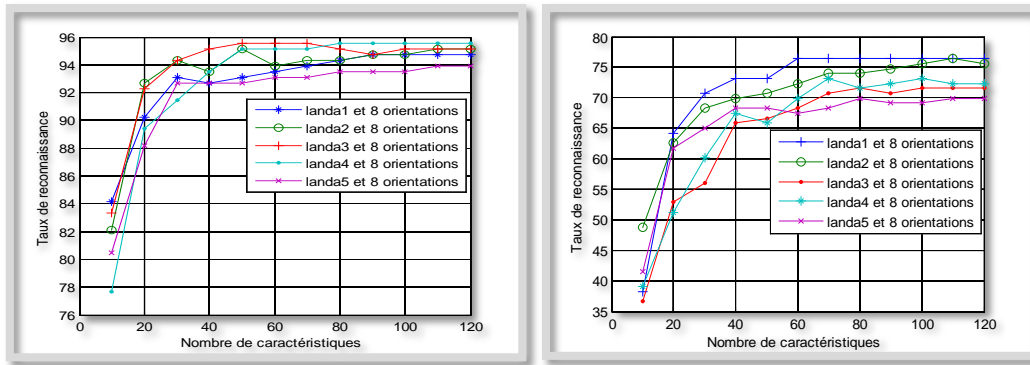
5.7.3 Résultats obtenus

5.7.3.1 Recherche des filtres optimaux

Nous nous intéressons au calcul du taux de reconnaissance dans l'ensemble d'évaluation et l'ensemble de test pour $\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$ et $\theta = \{0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$.

- **1^{er} expérience λ variable avec 8 orientations**

Dans les expériences menées nous fixons lambda et nous calculons le **TR** dans les 8 orientations pour divers nombre de paramètres caractéristiques puis, nous varions lambda et nous reprenons les mêmes expériences. Comparaison des taux de reconnaissance dans l'ensemble d'évaluation et de test pour **5** les résolutions et **8** orientations:



(a) l'ensemble d'évaluation (b) l'ensemble de test

Figure 5.35 Taux de reconnaissance pour 5 résolutions et 8 orientations.

Nous remarquons que le meilleur **TR** dans la base d'évaluation est obtenu pour $\lambda=8\sqrt{2}$ avec un **TR** = 95.52%. Dans la base de test le meilleur taux de réussite est obtenu pour $\lambda=4$ et $\lambda=4\sqrt{2}$ avec un **TR**=76.42%. C'est le résultat retenu dans cette première expérience.

Tableau 5.12 Tableau récapitulatif.

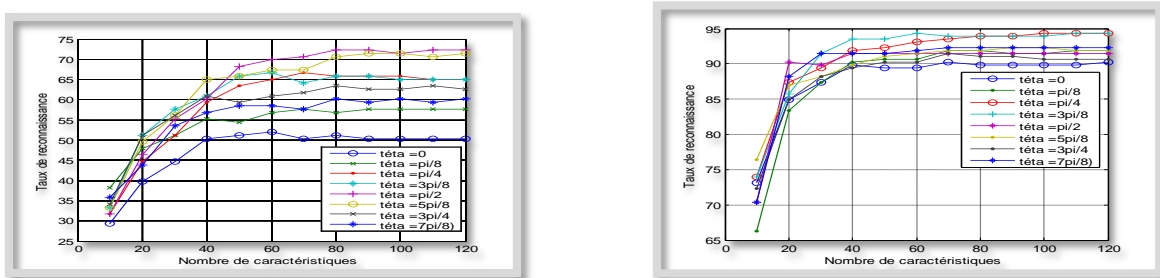
Résolution(λ)	4	$4\sqrt{2}$	8	$8\sqrt{2}$	16
TR % évaluation	94.71	95.12	95.12	95.52	93.90
TR% test	76.42	76.42	76.42	73.17	69.91
Np	100	110	100	100	110

Dans ce tableau nous constatons que le meilleur obtenu est **TR_{test} = 76.42 %** avec résolution $\lambda = 4$ et 8 orientations θ et c'est un bon résultat concernant la reconnaissance de visages présentant des expressions qui reste à améliorer.

• 2^{ème} expérience $\lambda=4$ avec 8 orientations

Le meilleur résultat de TR obtenu sur la base de test est **TR=76.42 %** pour la première résolution $\lambda=4$. Pour cela, nous essayons dans cette deuxième expérience de fixer $\lambda=4$ et changer les orientations $\{\theta = 0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$ en appliquant ceci sur les deux **BDD test** et **évaluation**.

Comparaison du taux de reconnaissance dans l'ensemble d'évaluation et de test pour $\lambda = 4$ dans les 8 orientations:



(a) l'ensemble de test (b) l'ensemble d'évaluation

Figure 5.36 Comparaison du taux de reconnaissance pour $\lambda = 4$ dans les 8 orientations.

Nous pouvons dire que le meilleur résultat dans la base d'évaluation est obtenu pour les orientations $\theta = 3\pi/8$ et $\theta = \pi/4$. Dans la base de test les orientations $\theta = \pi/2$ et $5\pi/8$ s'avèrent les plus favorables.

Tableau 5.13 Résultats du système d'identification pour les quatre meilleurs systèmes retenus dans l'ensemble d'évaluation et test.

Orientations des filtres (θ)	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$	$5\pi/8$	$3\pi/4$	$7\pi/8$
TR % évaluation	89.43	91.46	94.30	94.30	91.46	92.27	91.05	92.27
TR% test	52.03	57.72	65.85	66.66	72.35	71.54	63.41	60.16
Nombre de caractéristiques	60	90	100	60	80	90	80	80

5.7.2 Fusion par concaténation des filtres de Gabor associés aux orientations optimales

D'après les résultats obtenus précédemment nous retenons quatre orientations $\theta = \{\pi/4, 3\pi/8, \pi/2, 5\pi/8\}$, $\lambda = 4$ et pour diverses N_p . Ces descripteurs ainsi obtenus sont concaténés et de nouveaux TR calculés en fonction de N_p .

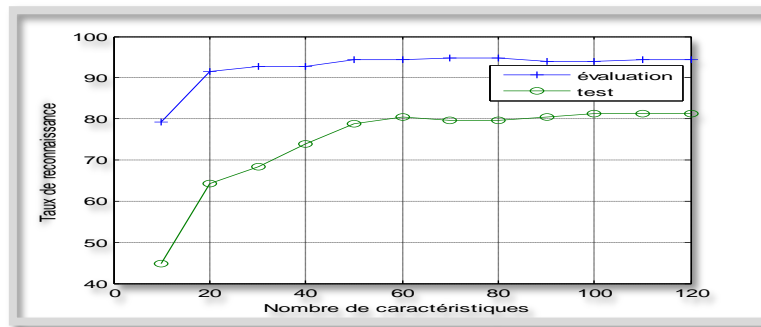


Figure 5.37 Valeurs de TR pour fusion par concaténation des descripteurs optimaux.

Les résultats de cette nouvelle expérience sont représentés dans le tableau suivant :

Tableau 5.14 Valeurs de TR pour fusion par concaténation des descripteurs optimaux.

	Nombre de caractéristiques											
	10	20	30	40	50	60	70	80	90	100	110	120
TR% évaluation	79.26	91.46	92.68	92.68	94.30	94.30	94.71	94.71	93.90	93.90	94.30	94.30
TR% test	44.71	64.22	68.29	73.98	78.86	80.48	79.67	79.67	80.48	81.30	81.30	81.30

Le meilleur résultat est obtenu avec un $TR_{test} = 81.30\%$ avec un $TR_{éval} = 94.30\%$. Nous pouvons affirmer que c'est un bon résultat relativement à ce que l'on espère avoir dans les applications 3D expressions en reconnaissance de visage monomodale. Une autre question se pose à ce stade du travail : Pourquoi ne pas simplifier notre système de reconnaissance en fusionnant l'image neutre avec ses six expressions et tenter d'identifier (ou authentifier) l'individu dans ce cas. C'est l'objet de notre application dans la section suivante.

5.8 Fusion du visage neutre et des visages avec les six expressions

La conception du système est illustrée par l'organigramme de la figure 5.38. Nous consacrons cette partie à l'implémentation du modèle étudié, la mise en œuvre du système et les résultats obtenus. Les outils utilisés, les algorithmes développés et l'approche réalisée sont mis en évidence. Dans cette étape nous fusionnons les scores du visage neutre et des visages de différentes expressions. Ces scores sont représentés par les distances calculées précédemment, la méthode de fusion utilisée est la **fusion par Ondelette**. Nous utilisons l'instruction **wfusmat** disponible dans **Matlab**, cette instruction fusionne deux matrices ou deux vecteurs. Dans un processus d'authentification, la stratégie de décision nous permet de choisir entre les deux alternatives suivantes: l'identité de l'utilisateur correspond bien à l'identité de la personne proclamée (ou recherchée) ou elle ne correspond pas. En comparant le modèle de l'utilisateur à celui qui existe dans la base de données. Cette phase est basée sur le seuil prédéfini.

$V2D_{test}$: image de visage 2D $_{test}$ (couleur). (Image de l'utilisateur)

$V3D_{test}$: image de visage 3D expressions test. (Image de l'utilisateur)

$XV2D \setminus XV3D$: modèle de la personne proclamée stocké dans la base de données

DIS_V2D : distance entre $XV2D$ et $V2D_{test}$. (projetées dans le même sous espace)

DIS_V3D : distance entre $X3D$ et $V3D_{test}$. (projetées dans le même sous espace)

DIS : distance fusionnée de DIS_V2D et DIS_V3D .

Phase test

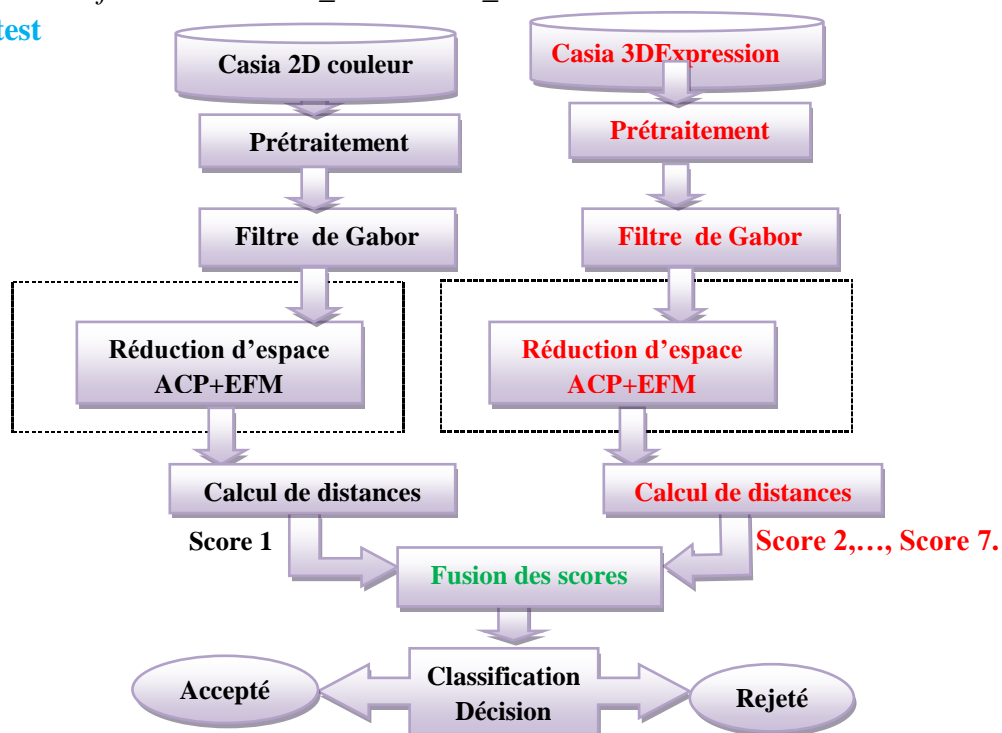


Figure 5.38 Organigramme du SRV 3D en présence d'expressions (Fusion des expressions)

Le programme se présente sous plusieurs fonctions, l'essentiel est présenté sous la forme suivante :

- *Lire()* pour lire l'ensemble des images d'apprentissage Visages **2D** et **3D** expressions , d'évaluation et de test.
- *Pretraitement()* pour préparer les images Visages **2D** et **3D** expressions , à la phase de traitement.
- *PCA()* *EFM()* cette fonctions renvoie une base dans laquelle on projette toutes les images Visages **2D** et **3D** expressions , d'apprentissage, d'évaluation et de test.
- *distance_intra()* renvoie les valeurs des distances entres les images des clients évaluation et les images d'apprentissage.
- *distance_extra()* renvoie les valeurs des distances entres les images des imposteurs évaluation et les images d'apprentissage.
- *Fusion_dis()* fusionne les distance intra et extra des visages **2D** avec les distances intra et extra visage **3D** expressions.
- *seuillage ()* cette fonction sert à estimer les valeurs de **TEE** et le seuil.

Nous utilisons le filtre de **Gaboret** la fusion des scores (Mean, Max, Min).

- Nous étudions d'abord le cas où le nombre de paramètres de $N_p(\text{EFM})$ variant de **10** à **200** et **PCA = 200**. Nous remarquons que la différence entre les erreurs pour chaque fusion n'est pas très grande. Ceci est une caractéristique très importante, car on peut dire que le système est stable.
- Nous étudions maintenant le cas où le nombre de paramètres de $N_p(\text{EFM})$ qui variant de **10** à **100** et **PCA = 100**.

Les résultats obtenus sont présentées sur le tableau suivant où nous présentons la comparaison entre $N_p(\text{PCA})=200$ et $N_p(\text{PCA})=100$ pour TR :

Tableau 5.15 Comparaison du TR pour fusion des expressions à base d'ACP+EFM.

PCA+EFM	Mean	Min	Max
100	92	92.25	78.92
200	77	82	69

Comme le montre ce tableau le meilleur **TR = 92.25 %** est obtenu pour $N_p(\text{PCA})=100$ pour la fusion par combinaison **Min** et classification par mesure de la norme Euclidienne L_1 .

5.9 Conclusion

Ce chapitre est structuré en deux grandes parties :

1 Approche monomodale 2D

Dans cette partie nous traitons la reconnaissance de visages 2D à base de techniques globales, divers classifieurs et à l'aide de techniques de réduction d'espace linéaires et non linéaires :

1.1 Méthodes de réduction d'espace linéaires :

Nous présentons les résultats obtenus en effectuant en premier lieu une étude comparative entre des méthodes de réduction d'espace à base d'ondelettes de **Gabor** comme méthodes d'extraction de caractéristiques pour la reconnaissance de **visages 2D** sur la base de données **XM2VTS**. Ces résultats montrent :

- L'efficacité de l'approche **ACP+EFM** par rapport **ACP, LDA, DLDA, ADL2DOC(ADLDOL)**. La méthode **EFM** est **mieux performante** que l'**ADL** suivie de l'**ADLD** suivie de l'**ACP** en employant un nombre restreint de caractéristiques (vecteurs propres) dans notre base de données **XM2VTS**. Nous constatons que le résultat obtenu avec **ACP+EFM** est le meilleur avec un **TEE=2.66±0.13%** et un **TR=94.33±1.49%** en utilisant seulement en utilisant seulement un nombre de caractéristiques égale **80**. Nos expériences sont menées avec la distance **corrélation normalisée** ou Angle.
- Les phases filtrées de la convolution des images de visages par les filtres de **Gabor** donnent des bons résultats pour quelques filtres. Nous remarquons que les meilleurs résultats sont obtenus avec la première résolution et les orientations : $\theta = \pi/2, 5\pi/8, 5\pi/8$.

Les résultats obtenus par les phases de **Gabor** sont donc satisfaisants. Pour cela nous utilisons ces trois phases des filtres de **Gabor** pour la conception de notre système multi-algorithmes dans le chapitre suivant. Ce résultat nous a motivé à poursuivre la suite de nos travaux avec des méthodes linéaires améliorées telles la **FLD (LDA)** et l'**EFM**. En réalisant la fusion multi algorithmique des quatre meilleurs systèmes résultats des investigations précédentes est réalisée et nous pouvons affirmer à ce stade que dans ce cas de **classification par SVM**, nous obtenons un parfait taux d'égale erreur avec **TEE=0±0%** dans l'ensemble d'évaluation et un bon taux de réussite **TR=97.72±1.46%** dans l'ensemble de test avec un temps de calcul T_c satisfaisant **t=0.32s**. Notre méthode est plus performante que la méthode qui utilise la **classification** basée sur la **distance de corrélation normalisée** avec un : **TEE=2.66 ±0.13%** et **TR=94.33 ±1.49%**.

1.2 Méthodes de réduction non linéaires

Ensuite nous avons étudié la méthode non linéaire **KPCA** pour une meilleure identification. L'association de la **KPCA** avec une classification par mesure de similarité basée sur la mesure de la norme **Euclidienne L2** nous donne dans la base d'évaluation un **TR=93.51%**.

Alors que l'association de la **KPCA** avec le classifieur **SVM** donne les meilleurs résultats avec un $TR_{\text{test}} = 97.5\%$ pour un noyau **Gausséen** avec la résolution $\sigma = 0.0004$ et $N_p = 600$ et c'est un bon résultat.

2 Approche monomodale 3D

Pour nos travaux en **3D** nous avons utilisé la **BDD CASIA 3D** version **2004**. Nous avons établi un protocole pour notre **BDD CASIA 3D** et assuré une bonne phase d'entraînement. Nos expériences ainsi menées nous conduisent à des résultats encourageants surtout que le domaine de l'identification des visages en présence d'expressions est très délicat d'autant plus que la base **CASIA 3D** est une collecte de visages Asiatiques où la ressemblance est très accentuée.

Notre étude comporte trois approches : l'approche monomodale **3D** et l'approche monomodale **3D expressions avec concaténation des régions d'intérêts** et l'approche monomodale **3D avec fusion des visages avec expressions faciales**.

2.1 Dans l'approche 3D c'est le prétraitement qui prime, pour cette raison la génération de cartes de profondeur est réalisée, et pour la facilité d'implémentation dans l'application l'**ACP+LDA** est utilisée. Nous avons testé notre approche sur trois variantes : **changement d'illumination**, **changement de poses** et **changement d'illumination + expression**. Nous obtenons pour un nombre de paramètres réduit respectivement **TR=88%**, **TR=84,5%** et **TR=91,5%**. On en déduit que dans ce cas que l'**ACP** en amont de la **LDA** améliore les performances du système d'identification. Les meilleures performances de l'**ACP** sont obtenues pour un nombre de paramètres réduits [**25,50**] en utilisant des tailles proches de l'**ACP** et de la **LDA**. Les performances du système sont dépendantes de la taille de l'**ACP** pour l'identification. La distance **corrélacion normalisée** pour des petites dimensions de l'**ACP** donne les meilleurs résultats pour la reconnaissance. Ce résultats reste insuffisant pour cela nous pensons que la fusion des caractéristiques ou scores des images de visages **2D** et **3D** pourrait améliorer la performance de notre SRV. Ce qui sera réalisé dans le chapitre 6.

2.2 Dans notre deuxième expérience pour la reconnaissance des visages **3D**, nous nous sommes intéressés aux expressions faciales qui posent grand problème en reconnaissance biométriques. Dans ce premier cas nous essayons d'extraire l'information contenue dans les régions d'intérêts en détectant ces derniers et en les soumettant au filtrage de **Gabor**. Puis les vecteurs descripteurs ainsi obtenus sont concaténés, puis réduits par **ACP+EFM** et soumis au classifieur **SVM**. D'après les résultats obtenus et présentés dans ce chapitre

nous remarquons que le meilleur TR dans la base d'évaluation est obtenu pour le filtre de résolution $\lambda = 8\sqrt{2}$ avec un TR = 95.52%. Dans la base de test le meilleur taux de réussite est obtenu pour $\lambda=4$ et $\lambda=4\sqrt{2}$ avec un TR = 76.42%. Les orientations présentant les meilleurs TR sont $\theta = 3\pi/8$ et $\theta = \pi/4$ dans l'ensemble d'évaluation $\theta = \pi/2$ et $5\pi/8$ dans l'ensemble de test (pour $\lambda = 4$). Une fois l'étude faite sur les différentes orientations et résolutions du filtre un choix est établi. Les meilleurs filtres sont concaténés et ce nouveau descripteur est soumis à l'étape de réduction d'espace. Et finalement, le meilleur résultat dans ce dernier cas est obtenu avec un **TR_{test} = 81.30 %** avec un **TR_{éval} = 94.30 %**. Nous pouvons affirmer que c'est un bon résultat relativement à ce que l'on espère avoir dans les applications **3D expressions**. Nous pouvons déduire aussi que le classifieur **SVM** donne toujours les meilleurs résultats pour toutes les techniques utilisées et ceci pour les visages **2D** et **3D**.

2.3 Notre étude s'est étendue à la fusion des caractéristiques du visage neutre+les visages présentant les six expressions faciales. Les résultats obtenus sont prometteurs

TR = 92.25% est obtenu pour une réduction d'espace **ACP+EFM** pour **N_p(PCA)=100** et **N_p(EFM) = 100** pour la fusion par combinaison **Min** et la classification par mesure de la norme **Euclidienne L₁**.

Nous pouvons dire que la concaténation des descripteurs basée sur les régions d'intérêts associés au SVM donne de meilleurs résultats que la fusion des expressions faciales. Ceci est dû en partie au SVM qui reste un classifieur de premier choix et aussi à la phase de prétraitement qui prend en compte l'information des régions d'intérêts mise en jeu dans les expressions faciales. Et à tous cela s'ajoute le temps d'exécution qui est plus rapide pour l'algorithme concaténation de descripteurs et au SVM.

Enfin pour améliorer davantage les performances de notre système d'authentification des visages. Nous proposons d'utiliser les filtres de Gabor au lieu d'utiliser l'information des pixels comme caractéristiques des visages et utiliser la fusion des scores issus de plusieurs experts (Multi algorithmes).

Malgré des résultats satisfaisants pour les systèmes d'authentification de visages obtenus dans ce chapitre les taux d'erreur associés à des systèmes biométriques monomodaux restent relativement élevés, ce qui les rend inacceptables pour un déploiement d'applications critiques de sécurité. Pour pallier ces inconvénients, une solution est l'utilisation de *plusieurs modalités biométriques* au sein d'un même système, on parle alors de système biométrique multimodal qui est étudié dans le reste de notre mémoire.