الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي و البحث العلمي

جامعة وهران للعلوم و التكنولوجيا محمد بوضياف

THÈSE

En vue de l'obtention du

Diplôme de Doctorat en Sciences

Présenté par : BENSENANE Hamdan

Intitulé

Sécurisation des systèmes biométriques

Faculté	***:Génie Electrique***
Département	***:Electronique***
Spécialité	***:Electronique***
Option	***:Techniques des Communications Modernes***

Soutenue le 30/04/2018 devant le Jury Composé de :

Membres de Jury	***Grade***	***Qualité***	***Domiciliation***
OUAMRI-Abdelaziz	***Professeur***	***Président***	***USTO***
KECHE - Mokhtar	***Professeur***	***Encadrant***	***USTO***
-	***-***	***Co-Encadrant***	
OUSLIM - Mohamed	***Professeur***	***Examinateurs***	***USTO***
HADJ SLIMANE - Zine-Eddine	***Professeur***		***Univ -Tlemcen***
DJEBBARI – Abdelghani	***MCA***		***Univ –Tlemcen***
MESSADI – Mahammed	***MCA***		***Univ –Tlemcen***

Année Universitaire : 2017-2018

Résumé

En raison de leurs avantages, les systèmes de Reconnaissance Faciale (RF) ont été parmi les systèmes biométriques les plus utilisés ces dernières années.

Tout d'abord, l'ART (transformation angulaire radiale) a été utilisée pour extraire les caractéristiques du visage qui alimentent un classificateur SVM ou un classificateur par voisin le plus proche pour la reconnaissance de visage. Les résultats obtenus montrent que les performances de cette approche se comparent favorablement avec celles des meilleures méthodes de l'état de l'art.

Ensuite, deux approches sont proposées pour contrer les pirates qui utilisent des masques 3D pour tromper les systèmes de RF. La première approche suit la démarche usuelle qui consiste à utiliser une phase de reconnaissance suivie d'une phase de vérification avec ART pour extraire les caractéristiques pertinentes qui sont introduites dans un classificateur basé sur le Maximum de Vraisemblance (MV) pour décider si l'image capturée représente une image d'un visage réel ou d'un visage avec masque. Dans la seconde approche, un système de reconnaissance faciale a été réalisé qui, sans nécessiter une phase de vérification, permet de rejeter les pirates qui tentent de le tromper en portant des masques 3D de personnes appartenant à sa base de données. Ce système combine les LMI et la LDA pour l'extraction des caractéristiques et utilise le MV pour la classification.

Mots clés

Reconnaissance Faciale (RF), Transformation Angulaire Radiale (ART), Spoofing, Maximum de Vraisemblance (MV), Moment de Legendre Invariant (LMI).

Summary

Because of their advantages, Facial Recognition systems (RF) have been among the most used biometric systems in recent years.

In the first part of this thesis it is proposed to use ART to extract facial features that feed an SVM classifier or nearest neighbor classifier for face recognition. The results obtained show that the performances of this approach compare favorably with those of the best methods of the state of the art.

In the second part of this thesis, two approaches are proposed to counter hackers who use 3D masks to deceive RF systems. The first approach follows the usual approach of using a recognition phase followed by a verification phase. In our case, this one uses ART to extract the relevant characteristics which are introduced in a classifier based on the Maximum Likelihood (MV).

In the second approach, it is proposed to use a facial recognition system that without requiring a verification phase allows rejecting hackers who try to deceive by wearing 3D masks of people belonging to its database. This system combines LMI and LDA for feature extraction and uses the MV for classification.

Keywords

Facial Recognition (RF), Radial Angle Transformation (ART), Spoofing, Maximum Likelihood (MV), Legendre Invariant Moment (LMI).

ملخص

بفضل مزاياها، أنظمة التعرف على الوجه من بين النظم البيومترية الأكثر استخداما في السنوات الأخيرة.

في الجزء الأول من هذه الرسالة نقترح استخدام ART لاستخراج ميزات الوجه التي تغذي مصنف أو مصنف أقرب جارPPV للتعرف على الوجه . النتائج التي تم الحصول عليها تبين أن أداء هذه الطريقة يقارن بشكل إيجابي مع أفضل الطرق السابقة.

في الجزء الثاني من هذه الرسالة، يقترح نهجان لمواجهة القراصنة الذين يستخدمون أقنعة 3D لخداع نظم التعرف على الوجه (RF). النهج الأول يتبع النهج المعتاد في استخدام مرحلة التعرف على الوجه تليها مرحلة فحص الوجه. في حالتنا هذه نستخدمART لاستخراج الملامح المهمة والتي يتم إدخالها إلى مصنف أقصى احتمالية (MV) لتحديد ما إذا كانت الصورة الملتقطة تمثل صورة لوجه حقيقي أو وجه مقنع. في النهج الثاني، اقترحنا نظام التعرف على الوجه الذي دون الحاجة إلى مرحلة التحقق يسمح لرفض القراصنة الذين يحاولون الخداع من خلال ارتداء أقنعة 3D للناس الذين ينتمون إلى قاعدة البيانات .هذا النظام يجمع LMI وLDA لاستخراج الملامح ويستخدم MV للتصنيف. أداء هذه النهج الثلاثة هي قابلة للمقارنة ويمكن وصفها بأنها مرضية للغاية، وميزة النهج الثاني (LMI+LDA+MV) هو أنه أسرع.

كلمات المفتاحية

نظم التعرف على الوجه (RF) ، تحويل الزاوية الشعاعي(ART) ، الغش ، أقصى احتمالية (MV) ،

عزم لوجندر ثابت(LMI) .

Remerciements

Tout d'abord, je remercie le bon dieu, tout puissant, de m'avoir donné la volonté d'entamer ce travail, la force et la patience pour le réaliser et le courage pour l'achever.

Je tiens à exprimer toute ma gratitude à mon directeur de thèse, le ***Professeur KECHE MOKHTAR*** pour l'aide compétente qu'il m'a apportée, pour sa patience et son encouragement. Sa disponibilité, son attention et sa compétence scientifique m'ont été précieux pour réaliser mon travail, le structurer et l'améliorer.

Je tiens à exprimer mes sincères remerciements à ***Professeur OUAMRI ABDELAZIZ*** de m'avoir accueilli au sein de son laboratoire et d'avoir accepté de présider le jury.

Je tiens à remercier ***Professeur OUSLIM MOHAMED, Professeur HADJ SLIMANE ZINE-EDDINE, Docteur DJEBBARI ABDELGHANI et Docteur MESSADI MAHAMMED*** qui m'ont fait l'honneur de juger ce modeste travail, leurs précieuses remarques et valeureux conseils.

Je tiens à exprimer ma gratitude à l'ensemble des personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

Je dédie ce modeste travail :

A **mes chers parents** *qui m'ont aidé à réaliser ce parcours. Ce travail est le fruit de longues années d'étude, de longs mois de persévérance et de longs jours d'apprentissage. Votre soutien et amour m'ont donné la force inconstamment pour réussir et prospérer dans la vie.*

A **ma belle-mère** *pour ses gestes gracieux, ses encouragements et ses conseils.*

« Vos prières et vos bénédictions m'ont été d'un grand secours pour mener à bien mes études »

A **mon frère**, **beaux-frères** *et* **mes sœurs** *pour leur soutien infini et leur aide incessante, à qui je souhaite un meilleur avenir.*

A **tous les membres de ma famille** *surtout mes neveux* **Yassine, Rayane, Sidi Mohammed, Amine et Selma.**

Ainsi qu'à **ma belle-famille** *et surtout notre* **chère grand-mère Hadja Zakia**.

Enfin, je témoigne toute mon affection et gratitude à **ma chère épouse** *qui m'a inlassablement encouragé. Son aide, soutien moral et réconfort psychique ont été d'une si grande importance pour que cette thèse soit achevée.*

A la mémoire de mes chers défunts

Mes grands-parents **Hadj Mohammed, Hadj Omar, Hadja Zoubida et Hadja Housni**

Mon beau père Professeur **Abdellatif Beghdadi** *qui a tant donné pour le savoir*

SOMMAIRE

LISTE DES FIGURES

LISTE DES TABLEAUX

LISTE DES ABREVIATIONS

ART : transformation angulaire radiale

ISV : variance inter-session

LBP : les modèles binaires locaux

LDA : analyse linéaire discriminante

MV : maximum de vraisemblance

PPV : le plus proche voisin

PZM : les pseudos moments de Zernike

LMI : les moments de Légendre invariant

SVM : support vector machine

RF : reconnaissance faciale

3DMAD : 3D Mask Attacks Database.

Introduction générale

I-Introduction générale

De nos jours on trouve pratiquement dans tous les domaines des systèmes de reconnaissance de personnes à fin de protéger les affaires personnelles ou sécuriser le fonctionnement de certains systèmes publiques. Des moyens informatiques sont mis en œuvre pour assurer cette tâche qui peut être : Le contrôle d'accès aux lieux de travail, le contrôle d'accès aux ordinateurs, l'e-commerce, les opérations bancaires basées sur l'identification, l'accès au moyens de transport, etc.

Le système de reconnaissance des personnes est basé sur la mesure de certains paramètres physiologique du corps humain pour les utiliser ensuite dans l'identification ou la vérification. Ces paramètres physiologiques sont généralement permanents, ce qui signifie qu'ils ne varient pas ou peu au cours du temps.

Plusieurs de ces mesures ont été techniquement prouvées et utilisées dans les systèmes de reconnaissance de personnes commercialisés.

Ces différents paramètres physiologiques du corps humain sont parfois appelés modalités biométriques. Plusieurs modalités biométriques (Figure. I.1) ont été utilisées pour représenter une personne. On peut citer le visage, l'empreinte digitale, l'iris, la forme de la main, les veines de la main, le réseau veineux de la rétine, l'empreinte de l'oreille et l'ADN.

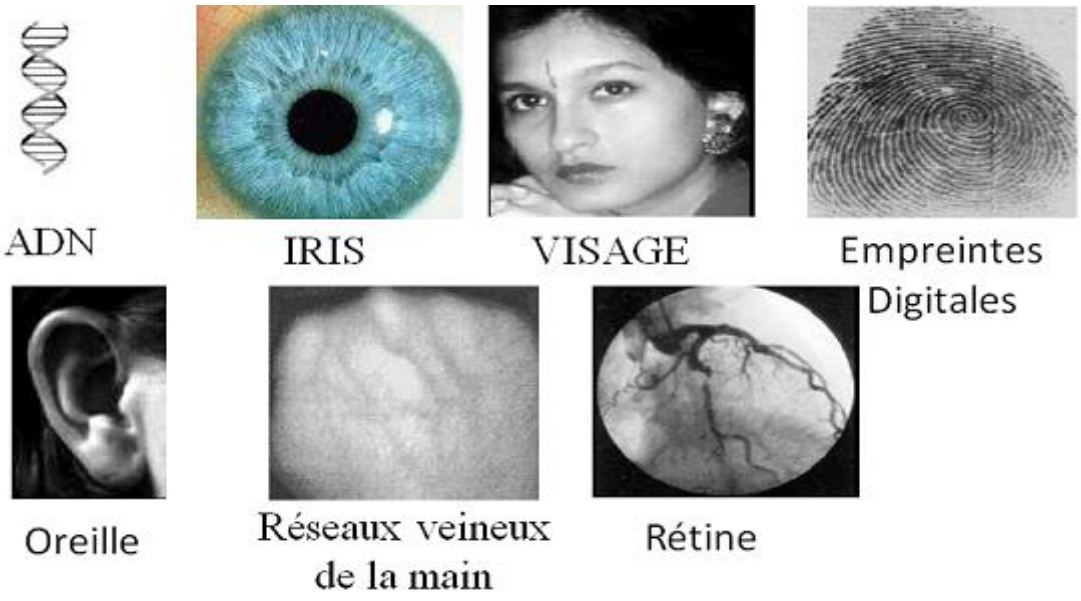


Figure. I.1 Modalités biométriques.

La science qui utilise les modalités biométriques dans les systèmes de reconnaissance s'appelle la biométrie.

La biométrie cherche à identifier une personne à partir de la mesure d'éléments biologiques, comportementaux ou physiologiques uniques et propres à chaque individu. La biométrie est utilisée dans plusieurs domaines, tels que l'interaction Homme Machine, la surveillance, l'indexation, la sécurisation des transactions, le Contrôle d'accès,...etc.

Le marché de la biométrie ne cesse d'évoluer depuis son apparition dans les années 70, comme le montre la figure I.2 (source : www.biometricgroup.com). Le développement du marché de la biométrie repose sur les innovations technologiques réalisées durant les deux dernières décennies.

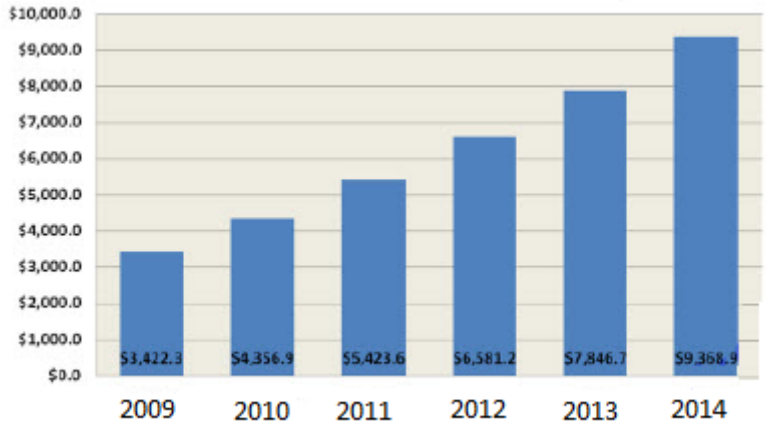


Figure I.2 : Evolution du marché de la biométrie entre 2009 et 2014.

La technique la plus utilisée dans les systèmes biométriques est les empreintes digitales, avec plus de la moitié du marché mondial. La reconnaissance de visage vient en deuxième position avec 11.4% de part de marché selon biometricgroup, comme le montre la figure I.3.

A la différence des modalités biométriques précédentes, la reconnaissance faciale ne nécessite pas une grande coopération des utilisateurs. Cette modalité biométrique est sans contact, naturelle, bien acceptée et ne nécessite en plus qu'un capteur très bon marchés (Webcam) qui est pratiquement disponible sur tous les appareils électroniques récents.

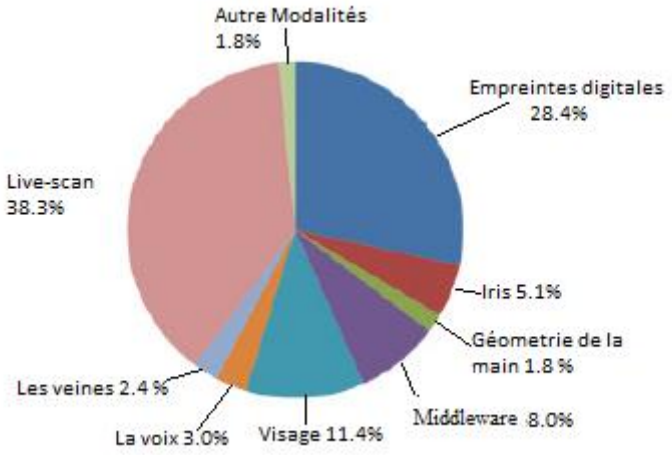


Figure I.3 : Pourcentage des revenus associés aux principales techniques de la biométrie en 2009.

On voit de plus en plus l'utilisation du visage comme modalité biométrique dans les systèmes de reconnaissance comme par exemple pour:

- Protéger des appareils électroniques personnels, comme les ordinateurs portables et les téléphones portables.
- Vérifier l'accès à un lieu privé réservé pour une population particulière, comme les laboratoires, les bureaux et les entreprises.
- Sécuriser les transactions bancaires, les paiements électroniques et l'activation des applications personnelles.

Un système automatique de reconnaissance de visages comme, montré à la figure. I.4, se décompose en trois sous-systèmes, qui sont la détection du visage, l'extraction des caractéristiques et la classification des utilisateurs.

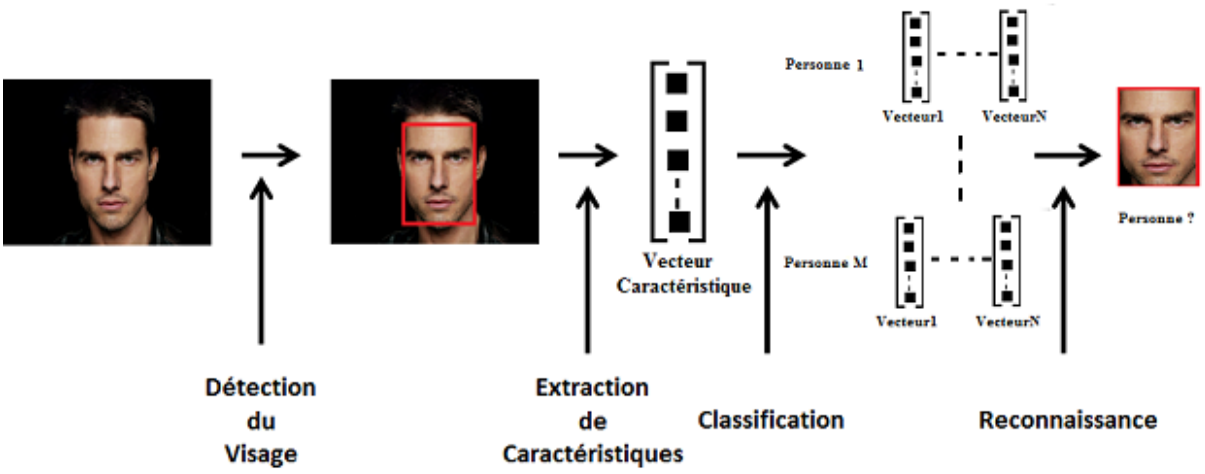


Figure I.4 : Système de reconnaissance du visage.

La difficulté de la reconnaissance du visage par webcam change suivant les conditions d'acquisition. Pour un environnement contrôlé, des paramètres tels que l'arrière-plan, la direction et l'intensité des sources lumineuses sont des paramètres maîtrisés lors de l'acquisition, par contre dans un environnement non contrôlé, une série de prétraitements sont nécessaires avant de procéder à la reconnaissance.

Les recherches dans le domaine de la reconnaissance du visage sont très nombreuses depuis le début des années 90.

On peut trouver plusieurs travaux et à tous les niveaux du système de reconnaissance.

Pour l'étape de détection du visage plusieurs méthodes de détection de visage ont été proposées; on trouve les approches qui se basent sur des connaissances acquises comme celle introduite par Yang et Huang dans [1], les approches basées sur des caractéristiques invariables, qui utilisent la propriété de la peau humaine pour capter les régions contenant des visages [2], les approches basées sur la mise en correspondance « Template-matching » comme celle de Yuille et al. dans [3] qui utilise un modèle déformable pour représenter les caractéristiques faciales, et les approches basées sur l'apparence globale qui sont pour la plus part des techniques d'apprentissage automatique, dont la plus connue est l'Eigenface [4].

Pour l'étape d'extraction des caractéristiques de visage, plusieurs algorithmes ont été proposés ces dernières années; on peut citer: ACP, LDA, ICA, ...etc. [30, 36, 38], pour les approches globale ; les algorithmes référence sont : EGM, EBGM,...etc. [42, 44], pour les approches locales, et les approches hybrides dans lesquelles les deux approches sont fusionnées comme par exemple : DCT, DWT, ...etc. [17, 47].

Pour l'étape de classification, on peut trouver deux types de classification la classification non supervisée comme le K-moyennes et la classification supervisée, basée sur une distance, comme la distance Euclidienne, la distance de Manhattan et la distance Cosinus, ou encore la classification par SVM (Supports Vectors Machine).

Malheureusement les progrès accomplis jusqu'à présent pour réaliser un système de reconnaissance fiable et performant sont menacé par l'apparition du problème du piratage.

Le piratage d'un système de reconnaissance de personnes est l'accès non autorisé à ce système par usurpation d'identité.

L'usurpation d'identité ou 'spoofing' est l'un des problèmes majeurs actuels auquel fait face la biométrie.

On parle de spoofing lorsqu'une personne essaie de se faire passer une autre personne pour accéder à un système de reconnaissance. Étant donné que les données faciales peuvent être

facilement acquises sans contact, le spoofing est une menace réelle pour les systèmes de reconnaissance faciale.

Il existe trois types d'attaques par spoofing dans les systèmes de reconnaissance, des attaques par photos, les attaques par vidéo et les attaques par masques.

Pour tromper le système biométrique facial dans les attaques par photos, les pirates peuvent simplement présenter à la caméra une photo du visage de l'utilisateur; l'efficacité de ce type d'attaque a été améliorée par l'apparition des imprimantes à haute résolution. Une autre possibilité c'est d'utiliser des photos très détaillées sur des écrans à haute définition (smart phones).

Les attaques par vidéo consistent à présenter une vidéo devant l'objectif de la caméra afin de tromper le système de reconnaissance faciale; à présent il est possible de présenter la vidéo d'une personne au système de reconnaissance avec des supports électroniques comme par exemple un Ipad ou un Smartphone.

Pour tromper le système de reconnaissance faciale dans les attaques par masque, les pirates se présentent devant la caméra avec un masque 3D d'une personne appartenant à la base de données du système; ce type d'attaque par masque est devenu une tâche facile grace à l'apparition des imprimantes 3D et des caméras avec une haute résolution.

La lutte contre les deux premières attaques a fait l'objet de plusieurs travaux de recherche, qui ont donné des résultats satisfaisants.

Malheureusement, toutes les méthodes de détection des attaques par photo ou vidéo faillent face aux usurpations d'identité par masque 3D.

En effet un système de détection du clignement des yeux et des mouvements des lèvres, par exemple, peut être vaincu en utilisant simplement des masques avec les régions des yeux et de la bouche imprimées avec haute résolution.

Contribution

Notre but dans cette thèse est de développer un système de détection des attaques par masque 3D simple et efficace. Dans la première partie de notre travail nous avons réalisé un système de reconnaissance du visage (extraction des caractéristiques, classification) pour prouver qu'un système de reconnaissance du visage est en général vulnérable aux attaques par masque 3D. Pour cela, nous avons utilisé la base de données 3D Mask Attacks Database (3DMAD) [29].

Pour l'extraction des caractéristiques du visage nous avons proposé une nouvelle approche, déjà utilisée dans plusieurs domaines du traitement d'images, nommée la transformation angulaire radial (ART).

La méthode proposée a été testée sur plusieurs bases de données qui contiennent des visages réels comme (ORL, Yale, faces94, faces96) et comparée avec d'autres approches qui utilisent la décomposition polynomiale de l'image comme les Pseudo-Moment-Zernike (PMZ) et les Moments de Legendre Invariants (LMI). Pour la classification nous avons utilisé deux méthodes, qui sont les plus utilisées dans la reconnaissance du visage, en l'occurrence la classification par le plus proche voisin, basée sur la distance euclidienne, et la classification par SVM.

Les résultats obtenus avec des bases de données qui contiennent des visages peu inclinés sont très bons. Cependant, les résultats obtenus avec la base de données 3DMAD, montrent que cette méthode est très vulnérable aux attaques par masques 3D. Ceci est le cas en général de tout système de reconnaissance de visage non doté d'un mécanisme de détection de ce genre d'attaques [75, 76].

Un système de reconnaissance de visage capable de détecter les attaques par masques 3D comporte généralement deux étapes: une étape de reconnaissance pour décider si le visage devant la caméra appartient à la base de données, suivie d'une étape dite de vérification pour discriminer les vrais visages des faux visages (masques).

En suivant cette approche, nous proposons une nouvelle méthode de vérification qui utilise l'ART pour l'extraction des caractéristiques et le Maximum de Vraisemblance pour la classification. Les résultats obtenus avec cette méthode en utilisant la base de données 3DMAD sont très satisfaisants et se comparent favorablement avec ceux obtenus avec une autre méthode proposée dans la littérature et qui utilise la même base de données. Notre méthode a l'avantage d'être plus simple.

L'approche qui combine la reconnaissance et la vérification a l'inconvénient d'être complexe. Pour réduire cette complexité, nous proposons une méthode de reconnaissance de visage qui est robuste aux attaques par masques 3D et qui ne nécessite pas la phase de vérification.

Dans cette méthode on combine la décomposition polynomiale de l'image par LMI avec l'analyse linéaire discriminante (LDA) pour l'extraction des caractéristiques et on réalise la classification par le maximum de vraisemblance.

Les tests effectués avec cette méthode de reconnaissance en utilisant plusieurs bases de données (ORL, Yale, faces94, faces96) montrent qu'elle permet d'obtenir un taux de reconnaissance très élevé et un taux de faux rejets très faible, tandis que les tests effectués en utilisant la base de données 3DMAD montrent qu'elle est très efficace pour contrer le spoofing par masque 3D.

II-Organisation de la thèse

Nous avons choisi d'organiser notre mémoire en quatre chapitres principaux.

Dans le second chapitre on présentera le spoofing. Les techniques principales de spoofing dans les systèmes de reconnaissance du visage seront présentées dans un premier temps. Dans un second temps, on présentera les différentes techniques développées au cours de ces dernières années pour contrer ce spoofing.

Le troisième chapitre sera consacré à la description d'un système général de reconnaissance de visage, avec ses deux composantes: l'extraction des caractéristiques et la classification. Pour chacune de ces composantes on passera en revue les principales méthodes proposées dans la littérature spécialisée. Pour l'extraction des caractéristiques, on présentera en plus une nouvelle méthode qu'on comparera avec d'autres méthodes qui utilisent le même principe. Ces méthodes combinées avec différents classifieurs sont testées sur des bases de données avec visages réels (ORL, Yale, faces94, faces96) et une autre base de données qui fournit des visages réels et des masques 3D de visages. L'utilisation de cette dernière base de données nous permettra de mettre en relief la vulnérabilité d'un système de reconnaissance de visage aux attaques par masques 3D de visages.

L'objet du quatrième chapitre est la présentation des deux solutions que nous proposons pour combattre ce type d'attaques. La première solution s'inscrit dans l'approche générale qui consiste à utiliser une phase de reconnaissance suivie d'une phase de vérification. Pour cette dernière phase on présentera une nouvelle méthode pour discriminer entre des visages réels et des visages avec masque 3D. Cette méthode utilise l'ART pour l'extraction des caractéristiques et le MV comme classifieurs. Ses performances sont comparables avec celles d'une autre méthode qui utilise la même base de données (3DMAD). Toutefois, l'avantage de la méthode proposée est qu'elle utilise des images RGB, qui peuvent être acquises à partir d'une simple webcam.

La seconde solution que nous proposons pour contrer les attaques par masques 3D de visages a l'avantage de ne pas nécessiter une phase de vérification séparée pour parer à ces attaques. Utilisant LMI combinée avec la LDA pour l'extraction des caractéristiques et le MV comme classifieur, elle permet d'obtenir un très bon taux de reconnaissance et en même temps un très faible taux d'acceptation de faux visages (visages avec masques 3D).

Dans le dernier chapitre du mémoire nous présenterons les outils et les bases de données utilisés dans notre travail ainsi que les différents résultats obtenus que nous commenterons.

Chapitre I
Spoofing par masque 3D

I-Introduction

Auparavant le piratage de tout type de systèmes de reconnaissance de personnes se faisait à l'intérieur du système en modifiant les données biométriques pour permettre aux personnes non autorisées d'accéder au système ou pendant la communication en remplaçant les décisions négatives d'authentification. Afin de se protéger contre ce type d'attaques les recherches étaient focalisées sur la sécurisation des données sauvegardées dans la base de données et protection des communications par des algorithmes de cryptage.

Les nouvelles attaques dans les systèmes de reconnaissance de personnes par visage par exemple ou une autre modalité biométrique ne nécessitent pas l'accès au fonctionnement interne du système, mais peuvent être effectuées de l'extérieur; elles se font au niveau du capteur pendant la phase d'acquisition.

Dans le cas d'un système de reconnaissance faciale, ces attaques consistent principalement à présenter à la caméra de faux visages.

Ces visages peuvent être réalisés à l'aide d'une image d'une autre personne (imprimée sur papier ou affichée sur un écran), une vidéo ou un masque 3D.

Les recherches dans le domaine des techniques de dépistage des peaux réelles et synthétiques ont commencé vers 2004. À cette époque, l'application de l'authentification biométrique augmentait.

Le projet Tabula Rasa a été lancé en 2011 pour la recherche principalement dans le domaine de la détection des attaques au niveau du capteur.

On peut diviser les axes de recherche dans le domaine du spoofing (piratage) dans les systèmes de reconnaissance du visage en trois axes selon les types d'attaques.

II- Attaques par photos

La première approche utilisée pour pirater un système sécurisé par la reconnaissance du visage était d'utiliser une photo du visage. Comme le montre la Figure 1.1, une personne peut utiliser son visage comme un mot de passe pour activer son ordinateur, mais toute autre personne est simplement rejetée, par ce qu'elle n'a pas le même visage.

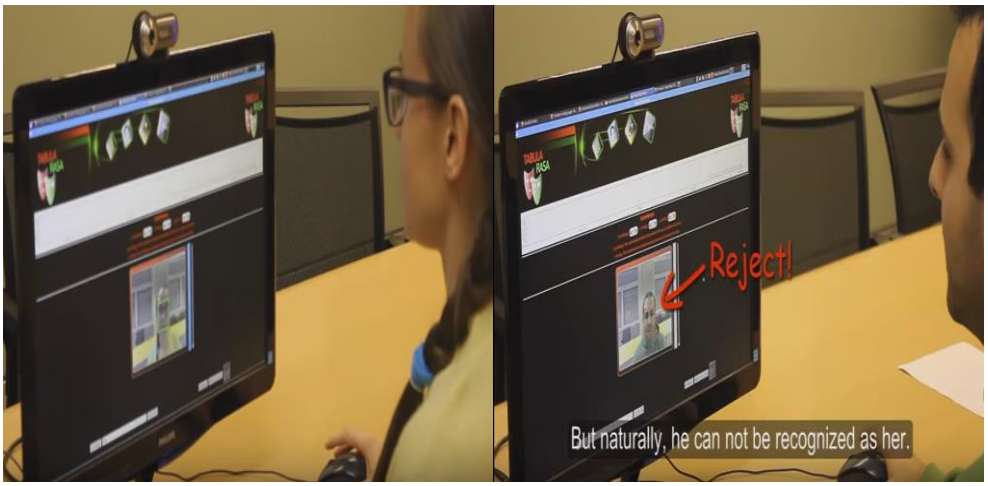


Figure 1.1. Système sécurisé par la reconnaissance du visage [1].

Afin de tromper le système d'authentification un pirate peut simplement transporter le visage valable pour activer le système sur un support en papier, comme le montre l'image ci-dessous.

Figure 1.2. Piratage d'un système sécurisé par la reconnaissance du visage avec photo [1].

Pour résoudre le problème de piratage par image, la plupart des recherches se sont focalisées sur la détection de la vivacité du visage, en se basant sur les clignements des yeux par exemple comme le montre la figure 1.3 ou par l'analyse de texture.

Pour distinguer entre un visage réel et une usurpation d'identité par photos les méthodes de détection de spoofing peuvent être classées en deux classes :

- Détection de mouvement ou détection de vivacité.
- Analyse de texture ou détection de la qualité d'image.

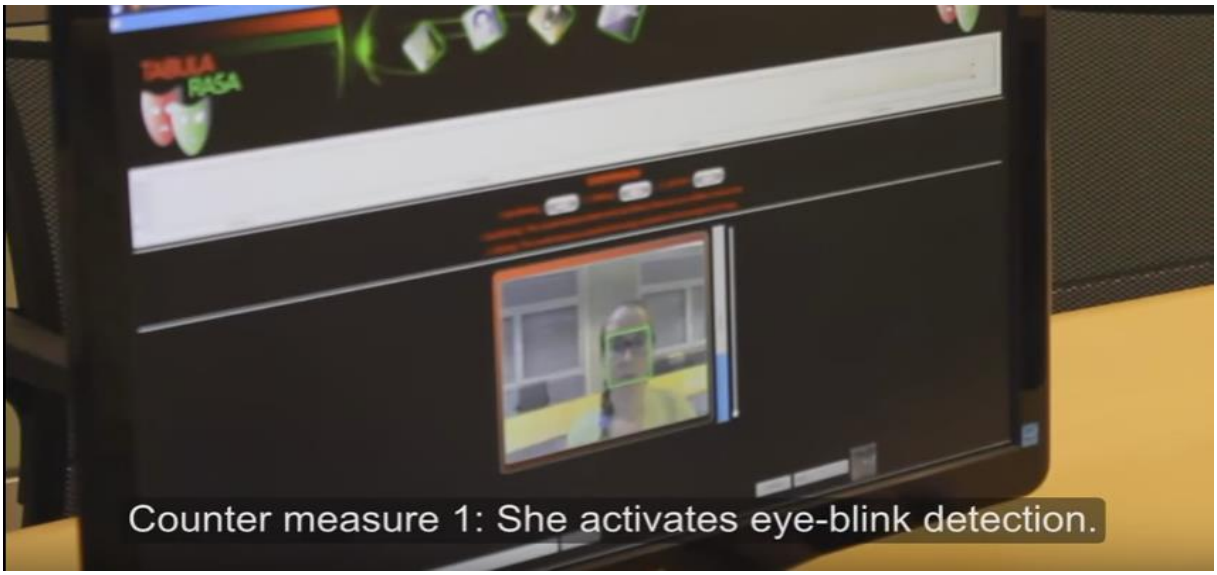


Figure 1.3. Neutralisation du piratage avec photo, par la détection des clignements des yeux, dans un système sécurisé par la reconnaissance du visage.

Dans leur travail, Jee et al. [2] ont détecté les yeux dans les images d'entrée séquentielles et ont calculé la variation de chaque région oculaire pour déterminer si le visage d'entrée est un vrai visage ou non.

Après normalisation de la région du visage, les régions des yeux sont extraites avec une taille 10x20 en fonction du centre des yeux. Ensuite, les régions oculaires sont binarisées en utilisant un seuil qui est calculé à partir de la valeur moyenne des pixels de chaque région oculaire.

La Figue 1.4 montre des exemples des régions d'œil binarisées extraites de 5 images séquentielles pour des visages réels et des visages faux. Comme le montre cette figure, les régions oculaires des faux visages changent très peu, mais les régions oculaires des visages réels ont une variation beaucoup plus grande, en raison du clignotement ou du mouvement de la pupille.

Figure 1.4. Exemple de régions œil binarisées de faux visages (a) et de visage réels (b).

Toujours pour la détection de la vivacité Pan et al. Dans [3] ont proposé une méthode qui analyse le comportement du clignotement des yeux, qui est représenté comme une séquence d'images temporelles après avoir été capturée numériquement par la caméra. Ils ont proposé un champ aléatoire conditionnel (CRF) pour modéliser une suite d'observations, en supposant qu'il existe une séquence sous-jacente d'états tirés d'un ensemble d'états finis. Une activité de clignotement peut être représentée par une séquence d'image S composée de T images, où S = $\{I_i, i = 1, ..., T\}$. Typiquement, les yeux dans les images s'ouvrent et se ferment, en outre il y a un état ambigu lorsque l'œil clignote en passant d'un état ouvert à un état fermé ou d'un état fermé à un état ouvert. Ils ont défini un ensemble de trois états pour les yeux.

Q = {o: open, c: close, b: ambiguous}.

Ainsi, une activité de clignotement typique comme montre la figure 5 peut être décrite comme un motif de changement d'état de o → b → c → b → o, les cercles dans la figure1.5 représente les états et les rectangles représente les transitions entre états.

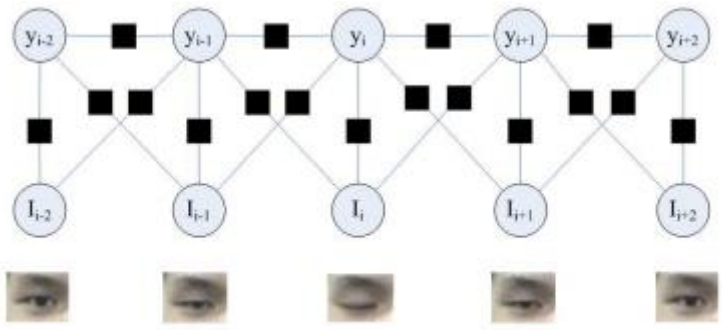


Figure 1.5 Modèle graphique d'un CRF à chaîne linéaire

A partir de ce modèle ils peuvent vérifier s'il y'a un clignement des yeux dans les séquences d'images capturées ou non. Les performances de la méthode proposée sont très bonnes, avec 1% d'échecs de détection d'attaques par photos.

Les autres approches utilisent l'analyse de texture et la détection de la qualité de l'image. Les méthodes dans cette catégorie tentent de mesurer les différences dans les détails d'images issues des visages réels et celles issues des écrans d'ordinateur ou du papier. Cela repose sur de multiples hypothèses, telles que: les images récapitulatives entraînent une diminution de la qualité, la réflectance de la lumière n'est pas la même entre les différentes surfaces, ou que l'impression sur papier crée des artefacts détectables.

Maatta et al. dans leur travail [4] adoptent les modèles binaires locaux (LBP) introduis par Ojala et al. [25] pour extraire les vecteurs caractéristiques. Le LBP est un opérateur de texture puissant, pour décrire non seulement les micros textures mais aussi leur information spatiale. Les vecteurs dans l'espace caractéristique alimentent alors un classificateur SVM qui détermine si les modèles de micro-texture caractérisent une personne réelle ou une fausse image.

L'opérateur LBP original forme des étiquettes pour les pixels d'image en comparant chaque pixel aux pixels dans voisinage 3 × 3; les résultats de la comparaison, qui sont soit 0 soit 1, sont regroupés pour former un code binaire pour chaque pixel. La figure 1.6 montre un exemple de calcul LBP. L'histogramme de ces $2^8 = 256$ étiquettes différentes peut alors être utilisé comme descripteur de texture.

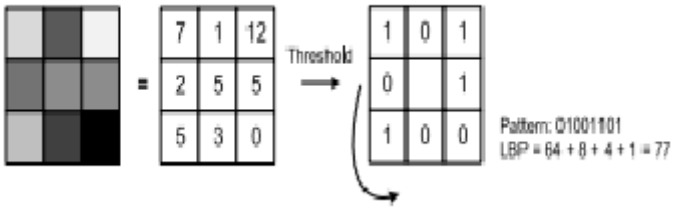


Figure 1.6. Calcul LBP pour un pixel

La Figue 1.7 montre des exemples de deux images (un visage réel et une photo) dans l'espace d'origine et les images LBP correspondantes en utilisant le LBP de base.

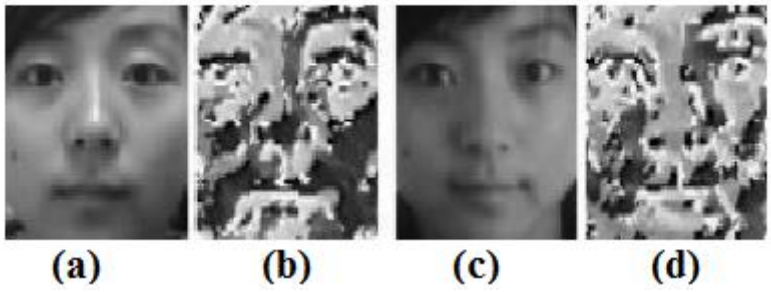


Figure 1.7. Deux images : un vrai visage(a), une photo(c) et leurs images LBP correspondantes (b, d).

Maatta et al. ont remarqué que la photo imprimée semble assez similaire à l'image du visage originale alors que les images LBP présentent certaines différences.

Le taux d'erreur de discrimination entre des images de visages réels et de faux visages obtenu est égal à 2.9 %.

Dans le même contexte Gautam et Jayash dans [5] ont proposé une technique qui exploite la différence dans les caractéristiques d'illumination entre le visage réel et le faux visage. Ils ont montré que la lumière qui tombe sur un visage réel se reflète au hasard dans différentes directions en raison de sa surface 3D (lèvre, nez, etc.), par contre la lumière se reflète uniformément sur une surface planaire 2D, telle que celle d'une photo, comme le montre la Figure 1.8.

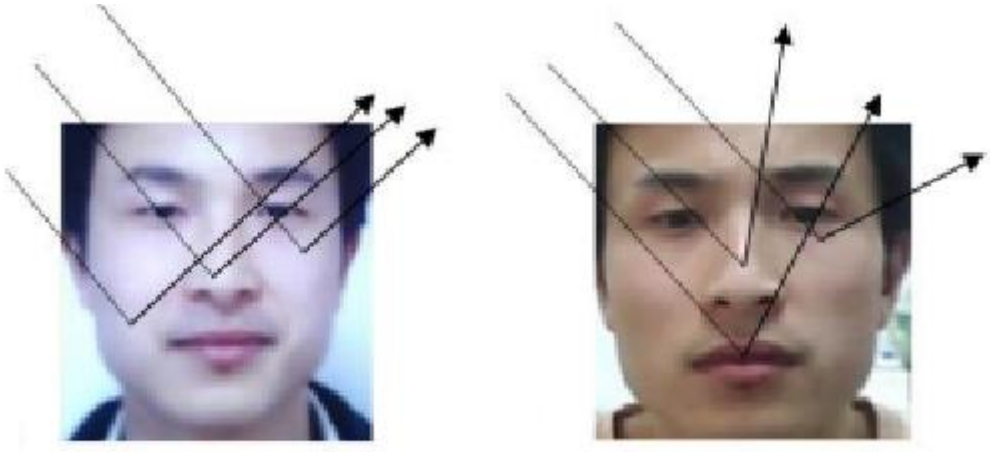


Figure 1.8. Caractéristiques d'illumination d'un faux visage et un vrai visage.

Le modèle local distribué (LDP) est suffisamment capable d'obtenir les petites différences entre les cartes de vitesse de diffusion des vrais et des faux visages [5].
Une image diffuse de l'image d'entrée, I, est calculée tout d'abord par une diffusion non linéaire comme suit:

$$U^{k+1} = U^k + (\mathrm{div}(l\, \nabla U^k\, l\, \nabla U^k) \text{ avec } U^0 = I \quad (1)$$

Pour résoudre l'équation de diffusion donnée en (1), le schéma de fractionnement de l'opérateur additif (AOS), défini dans l'équation (2), est utilisé:

$$U^{k+1} = \frac{1}{2}\left((I - 2\tau . A_x . (U^k)\right)^{-1} . \frac{1}{2}\left((I - 2\tau . A_y . (U^k)\right)^{-1} . U^k \quad (2)$$

Les auteurs ont ensuite utilisé la LBP sur les images diffusées.

Les résultats sont affichés sur la Figure 1.9; la première colonne montre des images issues de l'ensemble de données NUAA [15], la deuxième colonne des images issues de la diffusion, et la troisième colonne des images LBP.

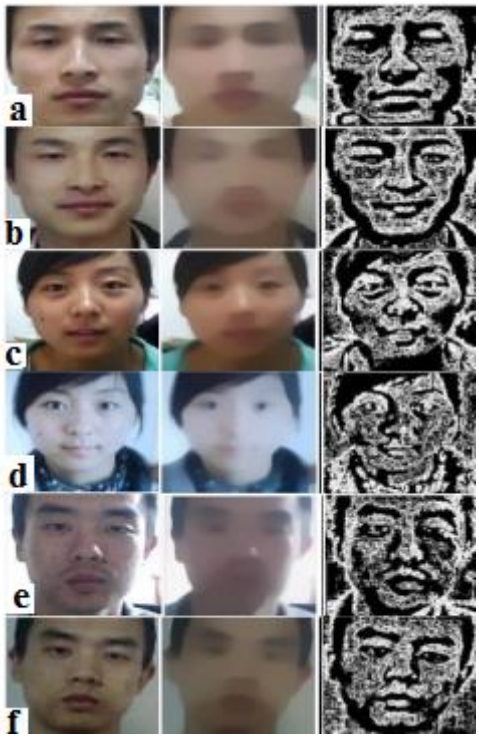


Figure 1.9 : (a), (c), (e) Visages réels et (b), (d), (f) Faux visages.

Les taux de reconnaissance obtenus par cette méthode sont 88.67% de taux de reconnaissance pour les visages réels et 92.53% de taux de reconnaissance pour les faux visages.
D'autres méthodes vérifient la vivacité dans l'ensemble du visage, par exemple en utilisant l'Optical Flow Fields, comme proposé par Bao et al. dans [6].

III-Attaques par vidéo

Afin de tromper le système d'authentification dans ce type d'attaques, les pirates utilisent des supports vidéo comme (smartphone, ipad) sur lequel il y'a un enregistrement vidéo d'une séquence d'un visage valable pour activer le système comme montre la figure 1.10.

Figure 1.10. Piratage d'un système sécurisé par la reconnaissance du visage avec vidéo [1].

Les méthodes pour détecter le spoofing par photo qui utilisent la détection de la vivacité du visage ou le clignement des yeux ont échoué face à ce type de piratage comme le montre la Figure 1.11.

Figure 1.11. Défaillance du système de vérification par l'approche de clignement des yeux contre les attaques par vidéo [1].

Pour résoudre le problème de piratage par vidéo, la plupart des recherches se sont focalisés sur l'analyse du mouvement du visage pour distinguer les mouvements 3D d'un visage réel des mouvements 2D d'un faux visage.

La méthode proposée par Wang et al. [7] repose sur l'hypothèse que lorsqu'un vrai visage tourne, les points de repère sur le visage se déplacent d'une manière différente par rapport à une face imprimée ou un support vidéo (voir la figure 1.12).

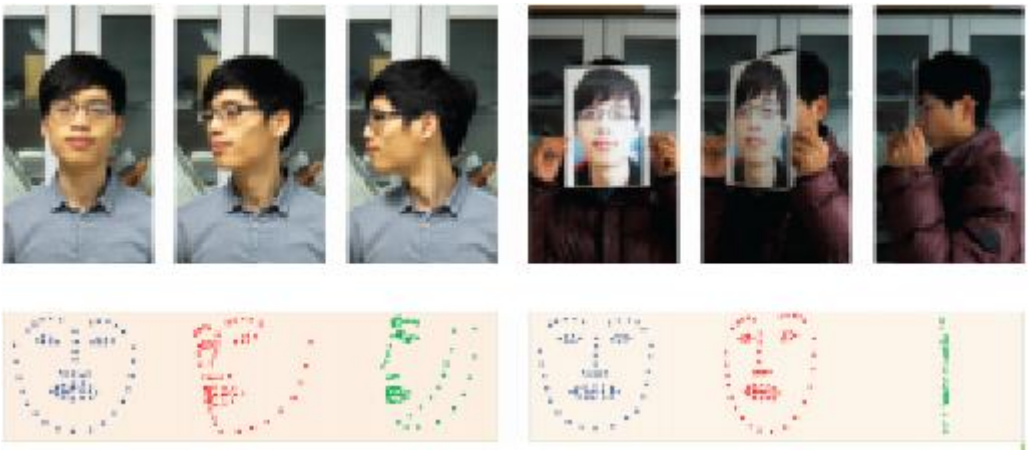


Figure 1.12. Comparaison de la façon dont les repères se comportent entre un visage authentique et la photo d'un visage.

Afin d'effectuer une détection en 3D, les repères doivent être extraits des séquences données. Deux algorithmes de détection de repères différents ont été utilisés. STASM [8] est un algorithme de détection qui est capable de détecter 77 points de repère différents sur le visage en utilisant les étapes suivantes:

- L'emplacement du visage est déterminé à l'aide d'un détecteur de visage global.
- L'emplacement global des repères est déterminé en utilisant la sortie de l'étape 1.
- Les étapes suivantes sont itérés environ quatre fois:
- Obtenir la partie de l'image entourant un point de repère.
- Comparer la position aux positions d'un modèle de forme global, connu pour ce point de repère.
- Améliorer le positionnement du repère.
- Confirmer la forme nouvellement proposée de tous les points de repère.

DEST [9] est un algorithme de détection qui est capable de localiser 64 repères et fonctionne quelque peu comme STASM. Toutefois, à la place d'améliorer la position des repères sur la base d'un modèle de forme global, il utilise des arbres de régression.

On trouve aussi la méthode de De Marsico et al. [10] qui utilise les rapports croisés de différentes configurations de repères pour distinguer entre les mouvements 3D et 2D (voir la Figure 1.13); les distances entre ces repères sont utilisées plus tard pour distinguer entre des visages authentiques et des visages sur des supports vidéo. Leur recherche montre que cette méthode peut être utilisée avec succès pour distinguer entre les visages réels et les visages sur support vidéo. Cela se fait en mesurant la variation de la valeur des rapports croisés dans le temps: si cette variation est supérieure à un certain seuil, un visage est marqué comme étant authentique.

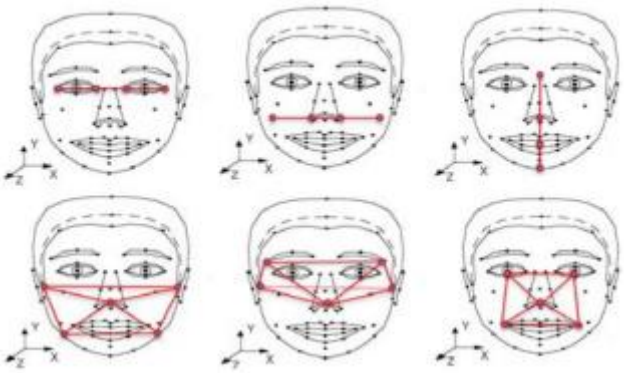


Figure 1.13. Configuration des repères utilisés par De Marsico

Pour distinguer entre un visage réel et un visage sur un support vidéo, la méthode proposée par Fokkema dans [11] repose sur un challenge que l'utilisateur doit gagner. La détection du spoofing dépend de l'utilisateur qui doit pivoter la tête. Il faut que l'utilisateur déplace un point sur l'écran dans un motif donné en tournant sa tête, comme le montre la Figure 1.14. Si l'utilisateur réussit à le faire, le défi est gagné, mais si l'utilisateur s'écarte du modèle ou manque de temps, le défi est perdu. L'utilisation d'un modèle précis généré aléatoirement rend le défi plus robuste contre les attaques par vidéo préenregistrées, parce que dans ces vidéos, l'utilisateur ne tourne pas sa tête dans les directions spécifiques et l'ordre que le défi demande.

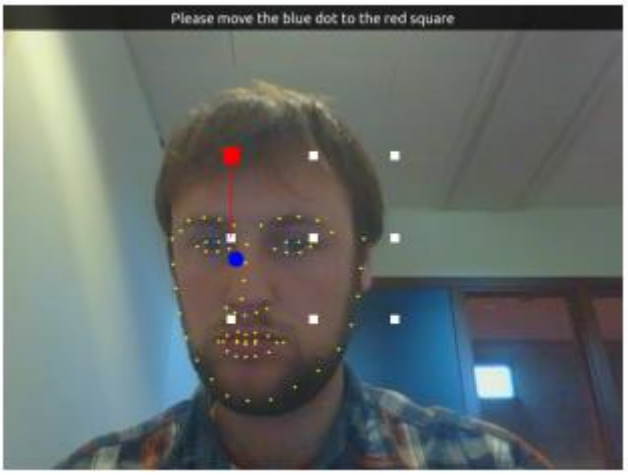


Figure 1.14 : Exemple de la façon dont le défi est présenté à l'utilisateur.

Dans leur travail W. Di et al. [12] proposent d'effectuer une détection du spoofing par support vidéo à l'aide de l'analyse de distorsion d'image (IDA). Quatre types de caractéristiques IDA (réflexion spéculaire, flou, moments de couleur et diversité des couleurs) ont été conçus pour

capturer la distorsion de l'image. Les quatre caractéristiques sont concaténées ensemble, pour former un vecteur de fonctionnalité IDA d'une dimension égale à 121, ensuite un classificateur d'ensemble composé de deux classificateurs SVM conçus pour différentes attaques est utilisé pour la discrimination entre les visages authentiques et les visages sur un support mobile.

Ils ont également construit une base de données, appelée MSU MFSD [13], en utilisant deux appareils mobiles (Android Nesus 5 et MacBook Air 13). Il s'agit de la première base de données de la parodie mobile. Un sous-ensemble de cette base de données, composé de 35 sujets est mis à la disposition du public [12].

W. Di et al. [12] ont obtenu un taux d'erreur total moyen (HTER en anglais) égal à 7.42%. Le HTER (Half Total Error Rate) est la moyenne entre le Taux de Faux Rejets (FRR, pour False Rejection Rate en anglais) et le Taux de Fausse Acceptation (FAR, pour False Acceptance Rate, en anglais) [14].

IV- Attaques par masque 3D

Afin de tromper le système de reconnaissance faciale ce troisième type d'attaques consiste à utiliser devant la caméra un masque 3D du visage d'une personne appartenant à la base de données du système comme le montre la Figure 1.15.

Figure 1.15. Piratage avec Masque 3D d'un système sécurisé par la reconnaissance du visage.

Le sujet des attaques des systèmes de reconnaissance par masques 3D est un nouveau problème pour la reconnaissance du visage. Toutes les méthodes de détection des attaques (photo, video) ont failli face aux usurpations d'identité avec masques 3D.

Kollreider et al. [18] montrent que les systèmes de reconnaissance faciale protégés par des systèmes de vérification s'appuyant sur les clignements des yeux et les mouvements des lèvres peuvent être vaincus en utilisant des masques 3D qui enveloppent le visage tout en laissant les yeux et la bouche apparents. De même un système de détection du clignement des yeux et les mouvements des lèvres peut être vaincu en utilisant simplement des masques de haute résolution d'impression des yeux et des régions de la bouche, comme le montre la Figure 1.16.

Aussi les contre-mesures basées sur les mouvements dépendent des différences entre les mouvements des surfaces 2D et 3D et ne s'appliquent pas lorsque des masques sont utilisés à la place de photos ou de vidéos.

De ce fait, l'impact des attaques par masques à des fins de falsification et les technique de contre mesure pour détecter ces attaques sont devenus un des sujets d'actualité pour les chercheurs dans le domaine de la biométrie.

La raison principale du retard dans les études de spoofing par masques 3D est due à l'indisponibilité de bases de données publiques de masques de visages. À notre connaissance, dans la littérature, il y en a que deux, l'une privée et l'autre publique.

Il existe plusieurs façons de fabriquer des masques. Un masque d'une personne peut être préparé même en utilisant des papiers. Par exemple pour fabriquer des masques colorés de n'importe quelle personne, la société 'Thats My Face' demande simplement deux photos de cette personne, une frontale et une autre de profil.

Les masques sont fabriqués en cartographiant les images 2D (Une image frontale et une image de profil) de la personne cible sur le modèle de visage 3D qui montre les caractéristiques ethniques de la personne cible. Cependant, comme le modèle 3D est basé sur une forme ethnique, il ne montre pas les caractéristiques exactes de la forme 3D de la personne cible.

Figure 1.16. Masques de haute résolution d'impression.

Comme l'un des premiers travaux dans le spoofing par masque 3D dans les systèmes de reconnaissance du visage, on trouve le travail de Kim et al. [19]. Ce travail introduit un nouvel espace de fonctionnalité 2D où les visages réels et les faux visages masqués peuvent être efficacement discriminés. Il exploite la disparité de réflectance basée sur l'albédo entre des visages avec une peau réelle et des faux visages fabriqués avec des matériaux spécifiques. Le vecteur caractéristique utilisé pour la classification est composé des mesures de rayonnement de la région du front sous des illuminations de 850 à 950 nm. Les peaux faciales et les masques montrent des distributions linéairement séparables dans l'espace caractéristique proposé, comme le montre la Figure 1.17. En appliquant simplement le discriminant linéaire de Fisher, ils ont obtenu une précision de 97,78% dans la détection de la face fausse.

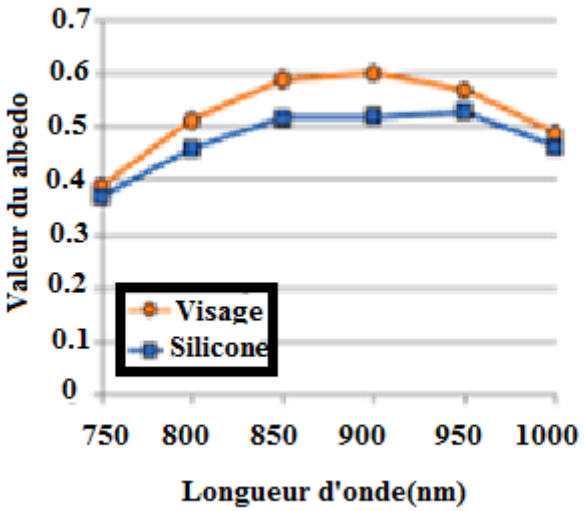


Figure 1.17. Réflectance albédo pour une peau réelle et une peau en silicone.

Le problème de leur travail est que les expériences ne sont pas effectuées sur de vrais masques, mais uniquement sur leurs matériaux, et les performances de spoofing ne sont pas incluses. Le fait que pour la détection de formes, des radiations doivent être utilisées pour se situer à une distance d'environ 30 cm de l'émetteur des rayons, ajouté à la possibilité d'occlusion dans la région frontale ainsi que les limitations de portée rendent cette méthode peu pratique.

De même, Zhang et al. ont proposé dans [20] une analyse multi-spectrale, en affirmant qu'il n'est pas possible de détecter les attaques en utilisant uniquement des images visuelles. Après avoir mesuré les courbes d'albédo de la peau du visage et des masques à différentes distances et avec différent matériaux, ils ont sélectionné deux longueurs d'onde discriminatives, 850 et 1450 nm en l'occurrence (voir Figure 1.18). Ils ont ensuite utilisé un classificateur SVM pour discriminer entre des visages avec une peau réelle et des visages avec une peau synthétique.

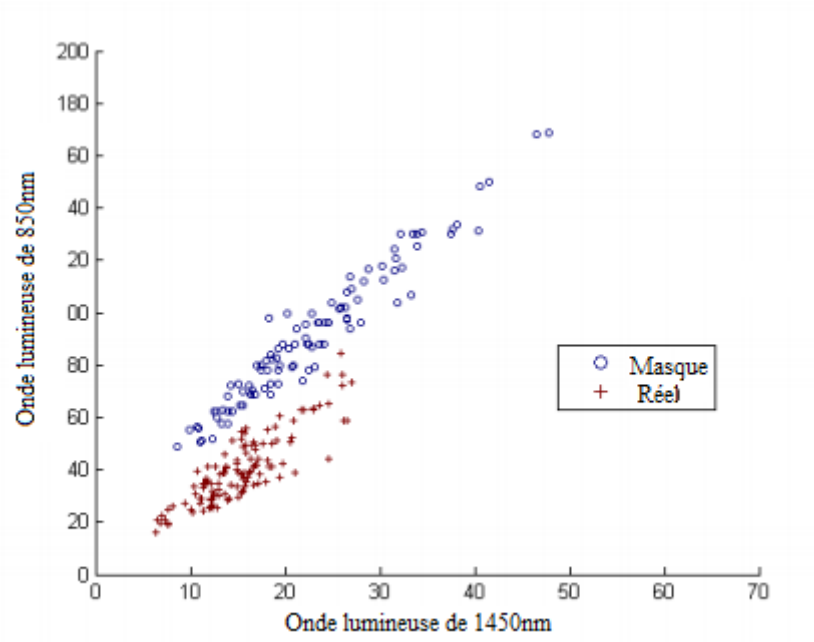


Figure 1.18. Répartition de la réflectance des masques et des visages réels.

Les expériences ont été menées sur une base de données composée de 20 masques de différents matériaux (voir Figure 1.19): 4 plastiques, 6 gel de silice, 4 pâte à papier, 4 plâtres et 2 éponges.

Figure 1.19. Échantillon de faux visages réalisés par des
gel de silice, plastique, pâte à papier, éponges.

Le système de vérification est couplé à un système de reconnaissance faciale par infrarouge, comme montré à la Figure 1.20. Le système comprend deux groupes de LED comme sources lumineuses actives. Les deux groupes de LED sont entrelacés et distribués de manière uniforme sur un rectangle. Deux photodiodes sont utilisées pour recevoir la lumière réfléchie aux deux longueurs d'onde.

Figure 1.20. Système d'émission et de réflexion réalisé par Zhang et al. [20].

Pour former le vecteur caractéristique Zhang et al. ont utilisé le modèle Lambertien de réflectance [21]; l'intensité lumineuse de réflectance I à un emplacement (x, y) vaut:

$$I(x, y) = A_0(x, y)\, r(x, y) \cos\theta(x, y) \quad (3)$$

Dans cette équation, A_0 (x, y) est l'intensité lumineuse d'entrée à l'emplacement (x, y) du visage, r(x, y) est l'objet albédo et θ(x, y) est l'angle entre le vecteur orthogonal à la surface et l'orientation du récepteur.

Comme ils ont trouvé des difficultés pour acquérir des masques faciaux, seulement 20 masques ont été réalisés, chacun échantillonné 5 fois à différentes distances pour former une base de données de 100 attaques par masque 3D.

Les résultats obtenus montrent que leur méthode peut atteindre une précision de détections correctes égale à 89,18%. Un avantage de cette méthode est qu'elle permet de s'affranchir de la contrainte distance.

Concernant le domaine de spooginf par masque 3D, N. Kose et J-L. Dugelay [22] sont les premiers à avoir fait une véritable étude. Ils sont les premiers à avoir calculé les performances de spoofing par masques dans les systèmes de reconnaissance du visage, en utilisant des données 3D issues du balayage 3D par un scanner, au lieu des données 2D (image de texture), pour détecter le piratage par masques. Leur travail a été mené en utilisant la base de données de masques qui a été créée par MORPHO [24] dans le contexte du projet de recherche de l'Union Européenne (UE), TABULA RASA [23]. Cette base de données comprend de nombreux échantillons de masques 3D de haute qualité de 16 sujets réels.

Les scans des sujets ont été acquis par un scanner 3D et les masques ont été fabriqués à l'aide d'une imprimante 3D. En plus des images de texture, elle comprend des balayages 3D pour les visages réels et les masques. Un exemple à partir de la base de données est donné à la figure 1.21; on y trouve sur la rangée supérieure le visage réel, l'image de texture, l'analyse 3D après le prétraitement et la carte de profondeur 3D du visage, tandis que sur la rangée inférieure on trouve les mêmes images pour le masque correspondant. Une carte de profondeur correspond à la distance réelle entre un objet et la caméra (coordonnées du centre optique, en 3D).

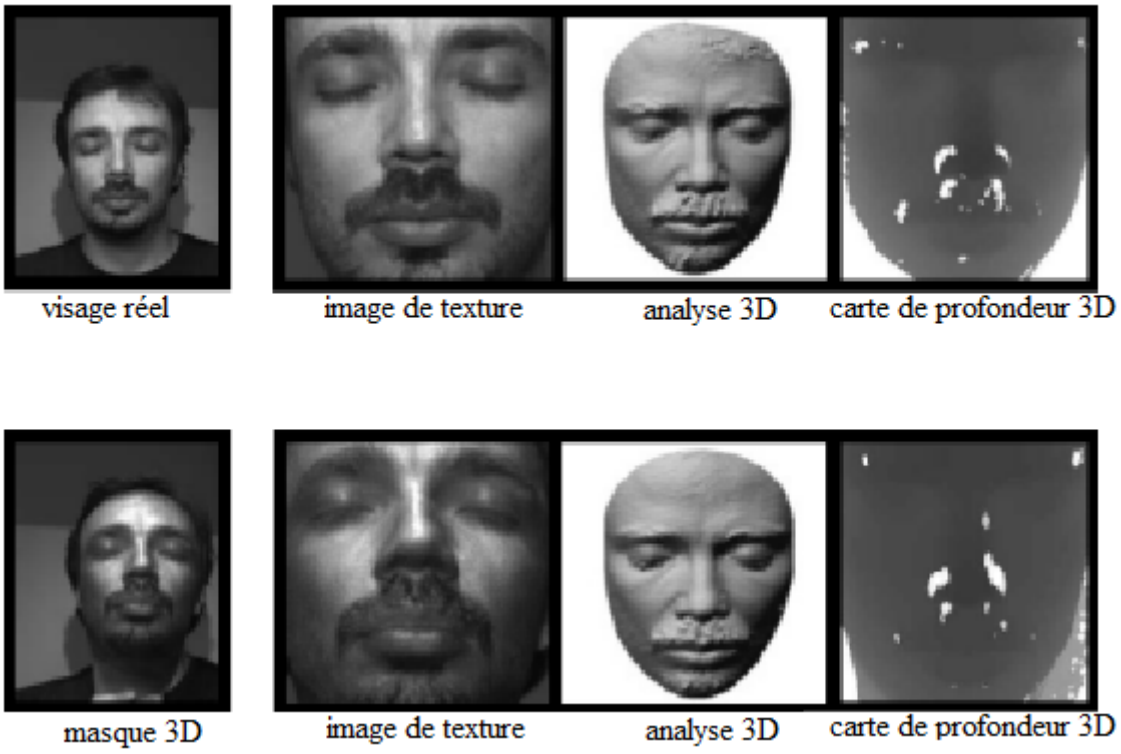


Figure 1.21. Exemple à partir de la base de données de masques créée par [24].

Grâce à cette base de données, N. Kose et J-L. Dugelay ont pu évaluer l'impact du spoofing par masques sur la reconnaissance de visage en 2D et 3D à fin de développer une contre-mesure à l'aide de données 2D et 3D. Les résultats de cette étude montrent que les masques utilisés ont une texture très similaire à celle des visages réels et en particulier les caractéristiques de la forme du visage 3D. Ils ont prouvé que les masques 3D sont très efficaces pour tromper les systèmes de reconnaissance faciale. Ils ont appliqué une analyse de micro-texture à la fois sur des images en niveau de gris et des images de profondeur par la méthode LBP [25]. Les précisions de classification totales obtenues avec les deux types d'image sont respectivement 88,12% et 86%.

Dans une autre étude N. Kose et J-L. Dugelay [28] ont fusionné les informations extraites à la fois des images de texture et de profondeur et ont obtenu une précision de classification plus élevée, égale à 93,5%.

Cette étude comporte deux lacunes principales: premièrement, les auteurs n'ont pas testé leur méthode par rapport au spoofing par masques imprimés. Deuxièmement, et surtout, la base de données utilisée n'est pas publique, ce qui constitue un obstacle à la recherche comparative et reproductible.

N. Erdogmus and S. Marcel sont allés plus loin dans [26]; ils ont réalisé des masques 3D à haute résolution à l'aide d'imprimantes 3D. Les masques sont des répliques de sujets réels, qui ont été analysés par un scanner 3D. Les auteurs ont réalisé une base de données publique, appelée 3D Mask Attack Database (3DMAD) [29].

Figure 1.22. 17 masques 3D répliques des 17 personnes réels de la base de données 3DMAD.

La base de données 3DMAD est principalement composée de 10 vidéos de 17 sujets avec visages réels et 5 vidéos de sujets avec des masques répliques des 17 sujets (voir figure ci-dessus), enregistrés par le capteur Microsoft Kinect. Pour chaque enregistrement dans la base de données 3DMAD, on trouve deux types d'image, une image RGB et une image de profondeur voire figure ci-dessous.

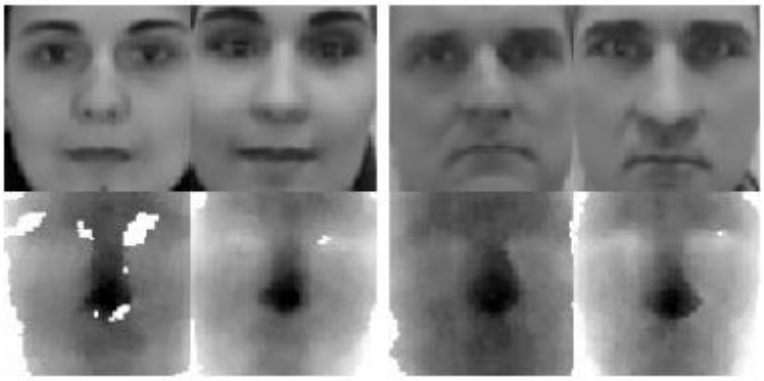


Figure 1.23. Image en niveau de gris et image de profondeur de deux personnes avec visage réel et avec masque 3D.

Dans la première étape de leur travail ils ont mesuré l'impact de spoofing par masque 3D dans le système de reconnaissance du visage 2D, qui utilise la méthode de modélisation de la variance inter-session (ISV) [27]. L'ISV est une extension plus fiable de l'approche GMM (Gaussian Mixture Models) qui estime des modèles clients plus fiables, en modélisant explicitement et en supprimant les variations intra-clients. Les expériences ont montré que 65.70% des tentatives d'attaque par masque dans l'ensemble de test sont incorrectement classées en tant que clients. Ce taux élevé de fausse acceptation par spoofing (SFAR) prouve que les attaques par masques dans la 3DMAD sont très efficaces pour tromper les systèmes de reconnaissance de visage 2D. Le taux de fausse acceptation (FAR) passe de 1,06% à 13,99% si les masques sont inclus dans l'ensemble des imposteurs.

Ayant constaté la vulnérabilité des systèmes de reconnaissance de visages aux attaques par masques 3D, les auteurs ont proposé un système de vérification composé de deux étapes et basé sur l'analyse de texture par LBP, pour contrer ces attaques (voir Figure 1.24). La première étape est une vérification 2D sur des images en niveau de gris par LBP et la deuxième étape est une vérification 3D, aussi par LBP, sur des images en profondeur.

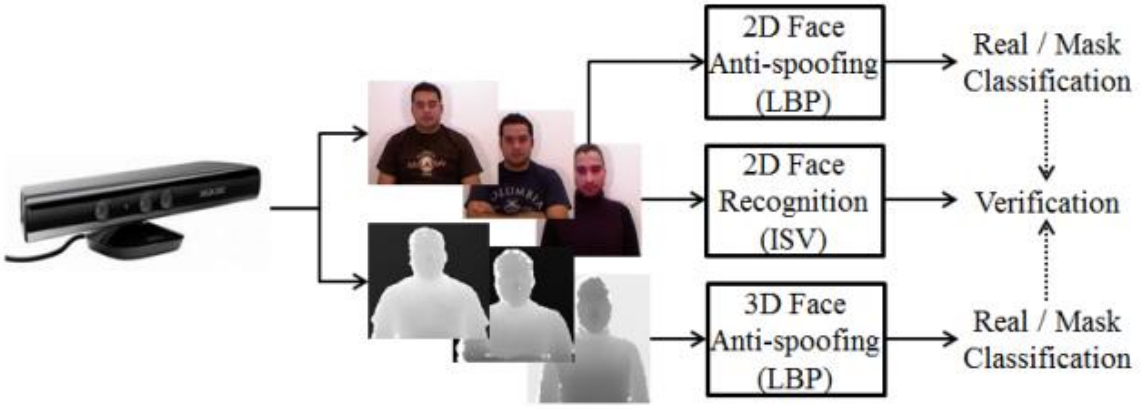


Figure 1.24. Système anti-spoofing proposé dans [26].

Dans leur étude, ils ont cherché à analyser les propriétés discriminatives des différents opérateurs LBP pour différencier ente un visage réel et un masque 3D. Pour chaque image de chaque vidéo ils ont formé un histogramme LBP. En utilisant une variante LBP, appelée motifs uniformes (LBP^{u2} 3×3, LBP pour simplifier), dans laquelle les motifs avec plus de deux transitions bit à bit sont éliminées, ils ont pu réduire la longueur du vecteur

caractéristique à 59. Ils ont également testé trois autres extensions de LBP proposées dans [26]: transitoire (tLBP), codé par direction (dLBP) et modifié (mLBP). La tLBP compare deux pixels voisins consécutifs de façon circulaire dans le sens des aiguilles d'une montre. La dLBP compare seulement quatre pixels adjacents, mais inclut également les informations de direction dans un extra bit. Enfin, la mLBP compare les pixels dans un voisinage 3 × 3 à leur moyenne au lieu du pixel central. En plus, pour chacune de ces variantes LBP, l'effet de la division de l'image faciale en blocs a été évalué.

En utilisant les histogrammes LBP comme vecteurs caractéristiques, trois classifieurs ont été testés. Le premier applique une mise en correspondance (matching) d'histogramme χ^2 pour comparer les échantillons de test avec un histogramme de référence qui est simplement calculé en prenant la moyenne de tous les échantillons d'accès (visage) réel dans l'ensemble d'apprentissage. Le second classifieur est un classifieur linéaire à base de LDA (Linear Discriminant Analysis), appliquée après la PCA (Principal Components Analysis) pour la réduction de la dimension. Le troisième classifieur testé est un classifieur non linéaire, le SVM avec noyau radial en l'occurrence. Les résultats obtenus sont présentés à la figure ci-dessous.

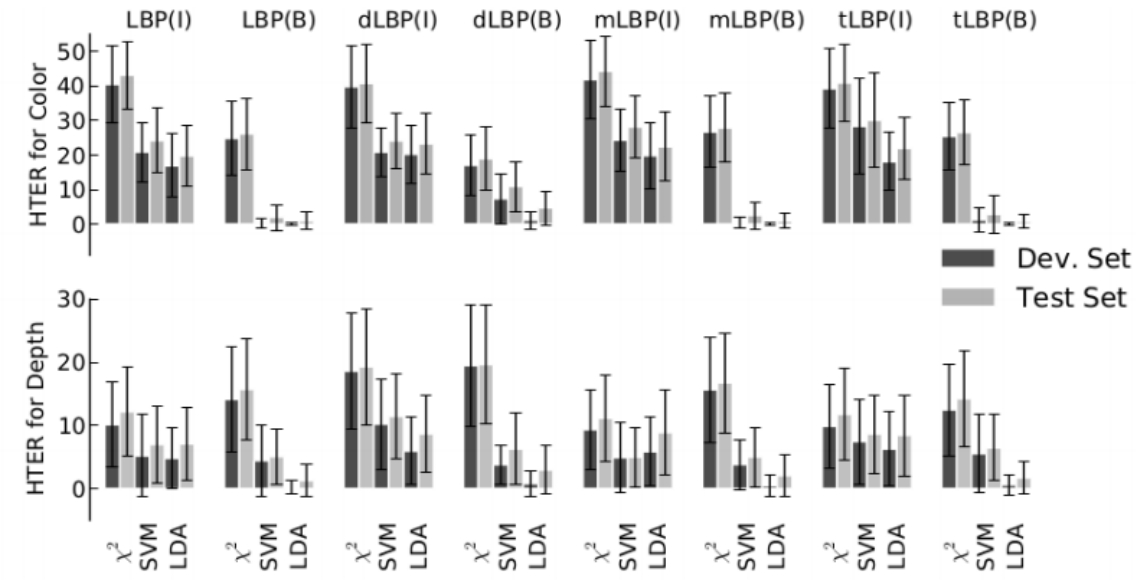


Figure 1.25. HTER des quatre types de LBP, appliqués par image (I) et par bloc (B), pour les images couleur et les images de profondeur, sont représentés avec des barres d'erreur et les incertitudes sont indiquées par des écart-types.

A l'issue de leur travail ils ont conclu que l'application de la LBP par blocs avec LDA donne les meilleurs résultats avec les deux types d'images (couleur et profondeur), pour lesquelles les valeurs HTER obtenues sont, respectivement, 0,95% et 1,27%. Le HTER (Half Total Error Rate) est la moyenne de deux taux : le FFR (False Fake Rate), qui le pourcentage d'accès réels classés comme attaques, et le FLR (False Living Rate), qui le pourcentage d'attaques classées comme accès réels.

Dans cette thèse nous proposons dans le chapitre quatre deux méthodes pour contrecarrer les attaques qui tentent de tromper un système de reconnaissance faciale à l'aide de masques 3D. La première méthode, à l'instar de la méthode proposée dans [26], utilise une phase de reconnaissance et une phase de vérification. Pour cette dernière phase, nous proposons d'utiliser l'ART pour l'extraction des caractéristiques et le ML (Maximum Likelihood) pour la classification. Nous avons testé cette méthode en utilisant la base de données 3DMAD et nous avons obtenu des performances qui sont similaires à celles de la méthode proposée dans [26]. L'avantage de la méthode proposée est qu'elle peut être utilisée par tout système de reconnaissance faciale avec des images RGB acquises par une simple webcam, contrairement à d'autres approches, telles que celles proposées dans [26] et [28] qui utilisent des capteurs spéciaux pour l'acquisition des images RBB et des images de profondeur.

Pour réduire la complexité, qui peut être un facteur déterminant dans un système de reconnaissance faciale embarqué, nous proposons une seconde approche qui évite la phase de vérification, et qui est donc capable d'effectuer simultanément la tache de reconnaissance et celle de rejet des attaques par masque. Cette approche utilise les Moments Invariants de Legendre, combinés avec la LDA pour l'extraction des caractéristiques et le ML pour la classification. Les taux de reconnaissance obtenus en utilisant différents bases de données et le taux SFAR obtenu en utilisant la base de données 3DMAD sont très satisfaisants.

Chapitre II :
Reconnaissance de visage

1-Introduction

Plusieurs techniques de reconnaissance de visages ont été proposées ces dernières années.

La phase de reconnaissance se déroule en deux étapes comme le montre la Figure 2.1.

- L'étape d'extraction des caractéristiques.
- L'étape de classificati5on.

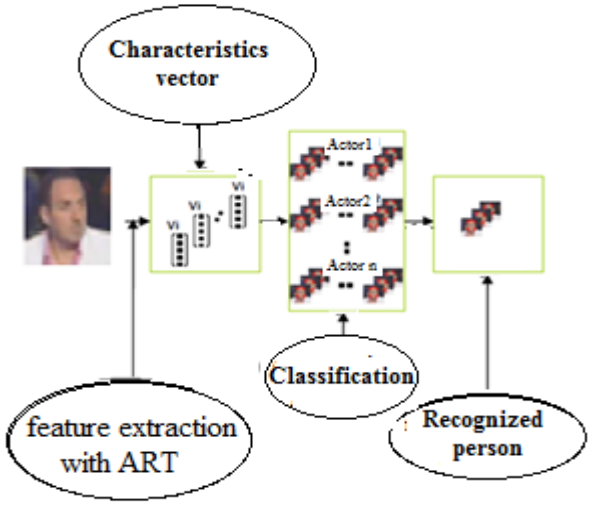


Figure. 2.1. Schéma d'un système de reconnaissance de visage.

L'extraction des caractéristiques est une phase très importante dans un système de reconnaissance de visages. Elle consiste à obtenir des informations, concernant un visage, pour ensuite former la base de données des signatures ou les utiliser dans l'étape de classification.

Ces informations doivent être discriminantes et non redondantes.

Les caractéristiques d'un visage doivent être différentes d'une personne à une autre et invariantes pour les images du visage de la même personne.

L'élaboration d'une signature représentative d'une personne est très difficile à former parce que les images du visage d'une même personne peuvent être différentes comme le montre la Figure 2.2. On trouve deux types de variations dans l'apparence du visage pour une même personne, les variations intrinsèques et les variations extrinsèques. Les variations intrinsèques sont principalement dues à des changements dans les expressions faciales lors de l'acquisition de l'image du visage comme: le sourire, l'ouverture de la bouche, le clignement des yeux, …etc.

Les variations extrinsèques sont principalement des changements externes qui affectent l'apparence du visage. Ces changements externes sont par exemple le port de lunettes ou d'un cache nez ou une modification de la source d'éclairage lors de la prise d'image.

Figure. 2.2. Visage affecté par des variations intrinsèque/extrinsèque.

II-Etat de l'art sur les méthodes d'extraction de caractéristiques

Les méthodes d'extraction de caractéristiques du visage pour représenter une personne peuvent être divisées en trois catégories:

- les méthodes globales,
- les méthodes locales ou géométriques,
- les méthodes hybrides.

II-1- Méthodes globales

Ces méthodes utilisent toute l'image pour l'extraction de caractéristiques. La matrice qui contient tous les vecteurs images de visages de dimension (n x m) est appelé espace images.

Pour les méthodes globales cette matrice est utilisée comme modèle de signature pour la classification des personnes. Ces méthodes permettent de sauvegarder implicitement toutes les informations de texture du visage et son aspect global. Toutefois, le grand problème de ces méthodes réside dans la taille très grande du vecteur des pixels de l'image, ce qui rend très difficile et très longue l'étape de classification. Par exemple, une image d'une taille 200 ×200 est transformée en un vecteur de 4.10^4 éléments et comme le nombre d'images d'apprentissage propre à chaque personne doit être au moins égal à dix, donc la dimension de la matrice des caractéristiques qu'il faut par personne est égale à 4.10^4 x 10, un nombre assez grand.

En pratique, on n'a pas besoin de tous les éléments de l'image du visage pour construire un modèle représentatif d'une personne. Des techniques de réduction de dimension sont généralement utilisées, comme les Eigenfaces (PCA), les Fisherfaces (LDA) et l'Analyse en Composantes Indépendantes (IDA).

Eigenface est une technique de réduction des dimensions de la taille des vecteurs développée par Turk et Pentland dans [30]; pratiquement, elle est la première technique fonctionnelle de reconnaissance de visages. Elle repose sur l'analyse en composantes principales (PCA) d'un ensemble d'images du visage sauvegardées dans une base de données. Son principe consiste à

trouver les vecteurs propres de la matrice de covariance des vecteurs d'image. Pour construire la matrice de covariance, chaque image de visage I(m,n) est traitée comme un vecteur Γ_i de dimension m×n (Figure 2.3). Chaque élément du vecteur correspond à l'intensité lumineuse d'un pixel. On rassemble ensuite les M images de la base de données dans une matrice unique Γ, où chaque colonne Γ_i représente une image (Figure 2.4).

Fig. 2.3. Procédure de transformation d'une image en un vecteur V_i (m×n,1)

$$\Gamma = \begin{pmatrix} a1,1 & b1,1 \ldots z1,1 \\ \vdots & \vdots \quad \vdots \\ an,1 & bn,1 \ldots zn,1 \\ \vdots & \vdots \quad \vdots \\ a1,m & b1,m \ldots z1,m \\ \vdots & \vdots \quad \vdots \\ an,m & bn,m \ldots zn,m \end{pmatrix}$$

Fig. 2.4. Matrice Γ (m×n,M) des M images de la base de données.

L'image moyenne Ψ, calculée en utilisant l'équation (3.1), est ensuite utilisée pour obtenir les images ajustées Φ_i à l'aide de l'équation (3.2).

$$\Psi = \frac{1}{M}\Sigma_{i=1}^{M}\Gamma_i \tag{3.1}$$

$$\Phi_i = \Gamma_i - \Psi \quad \text{pour : } i = 1...M \tag{3.2}$$

Finalement, la matrice de covariance, C, du jeu de données est calculée par l'équation 3.3:

$$C=\sum_{i=1}^{M} \Phi_i \Phi^T{}_i = A*A^T \quad \text{avec} \quad A=[\Phi_1 \Phi_2 \Phi_3 \dots \Phi_M] \tag{3.3}$$

On peut former la matrice de de covariance C à partir de deux matrices, la matrice de dispersion intra-personne C_I et la matrice de dispersion extra-personne C_E.

$$C_I = A_I^T . A_I \tag{3.4}$$

$$C_E = A_E^T . A_E \tag{3.5}$$

$$C = C_E + C_I \tag{3.6}$$

C, est égale à C_E dans le cas où on a seulement un seul exemple d'apprentissage par personne, ce qui donne $C_I = 0$.

L'Eigenface calculé en se limitant seulement à la matrice de dispersion extra-personne C_E n'est pas très performant, parce qu'il ne peut faire une discrimination entre les erreurs d'identification et les erreurs causées par la transformation et le bruit.

Ceci est illustré aux Figures 2.5 (a), 2.5 (b) et 2.5 (c) où sont représentés, respectivement, les Eigenfaces obtenus par décomposition avec la matrice de dispersion intra-personne C_I, ceux obtenus par décomposition avec la matrice de dispersion extra-personne C_E, et les Eigenfaces standards [31]. Les Eigenfaces extra personnels ressemblent beaucoup plus aux Eigenfaces standards.

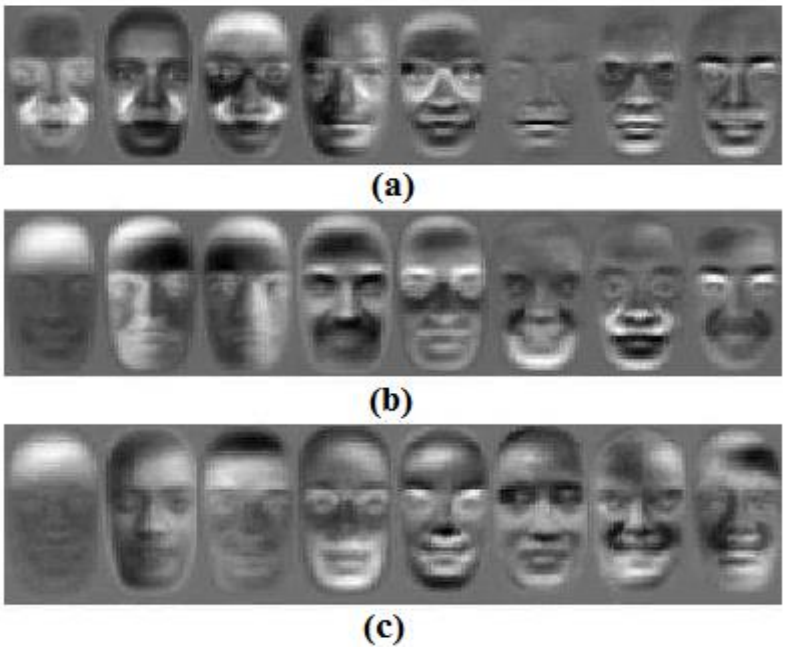
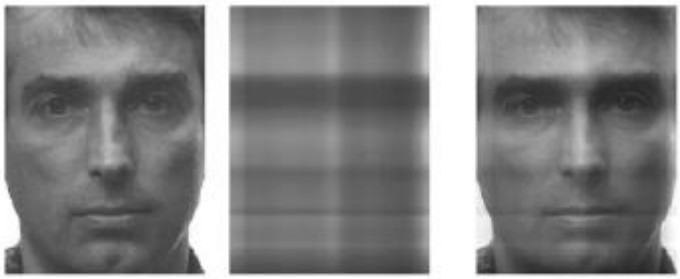


Figure 2.5 : (a) Eigenfaces inter personnels, (b) Eigenfaces extra personnels, (c) Eigenfaces standards.

Comme autres extensions de l'algorithme PCA, on trouve la $(PC)^2A$ présentée par Wu et Zhou dans [32], où ils ont utilisé une méthode basée sur deux projections pour enrichir

l'information sur l'espace image, une projection horizontale et une projection verticale de l'image comme montré à la Figure 2.6 (b). L'image obtenue à partir de ces deux projections est ensuite combinée avec l'image originale pour générer une image plus riche en information comme montre la figure 2.6 (c). La $(PC)^2A$ a amélioré le taux de reconnaissance de 5 % par rapport à la méthode eigenface standard avec une réduction de plus de 10 % de la taille du vecteur eigenface.



(a) **(b)** **(c)**

Figure. 2.6. (a) image originale, (b) image de projection, (c) image combinée.

Dans la même alignée, Yang et al. [33] ont introduit une nouvelle méthode qu'ils ont appelée Analyse en Composantes Principales Bidimensionnelle (ACP2D). Cette méthode applique directement l'ACP sur les images 2D au lieu de les transformer en vecteurs 1D. Cette modélisation permet de réduire le coût calculatoire et l'instabilité numérique lors de la construction du modèle.

Par ailleurs, plusieurs chercheurs ont développé diverses extensions de la méthode « eigenface », comme l'eigenface probabiliste [34] et l'eigenface laplacienne connue par Laplacian faces [35]. Ces approches possèdent des performances supérieures à celles de la méthode eigenface basique. Le problème est que ces techniques ne fonctionnent pas si un seul exemple d'apprentissage par personne est disponible.

Dans les mêmes approches de réduction des vecteurs image on trouve les systèmes d'identification du visage par l'analyse discriminante linéaire (LDA) proposée dans [36].

L'apprentissage par la LDA est effectué en cherchant une projection optimale W qui maximise la dispersion intercalasse, relative à la matrice S_w, tout en minimisant la dispersion inter-classe, relative à la matrice S_b.

En d'autres termes, on cherche W qui maximise le critère d'optimisation de Fisher J(W) :

$$W = \arg\max (J(W)) = \frac{|W^T S_b W|}{|W^T S_w W|} \qquad (3.7)$$

Pour cela il est nécessaire de trouver les vecteurs propres du produit matriciel (Sw^{-1} Sb).

La maximisation de l'équation n'est possible que si la matrice intercalasse S*w* est non singulière (inversible). Cette condition pose un problème pour les applications de reconnaissance du visage qui utilisent la LDA. Pour remédier à ce problème, Belhumeur et al. [37] ont proposé d'utiliser l'espace PCA réduit comme espace intermédiaire. La LDA appliquée à la reconnaissance faciale s'appelle Fisherface. Nous allons la détailler dans le chapitre IV.

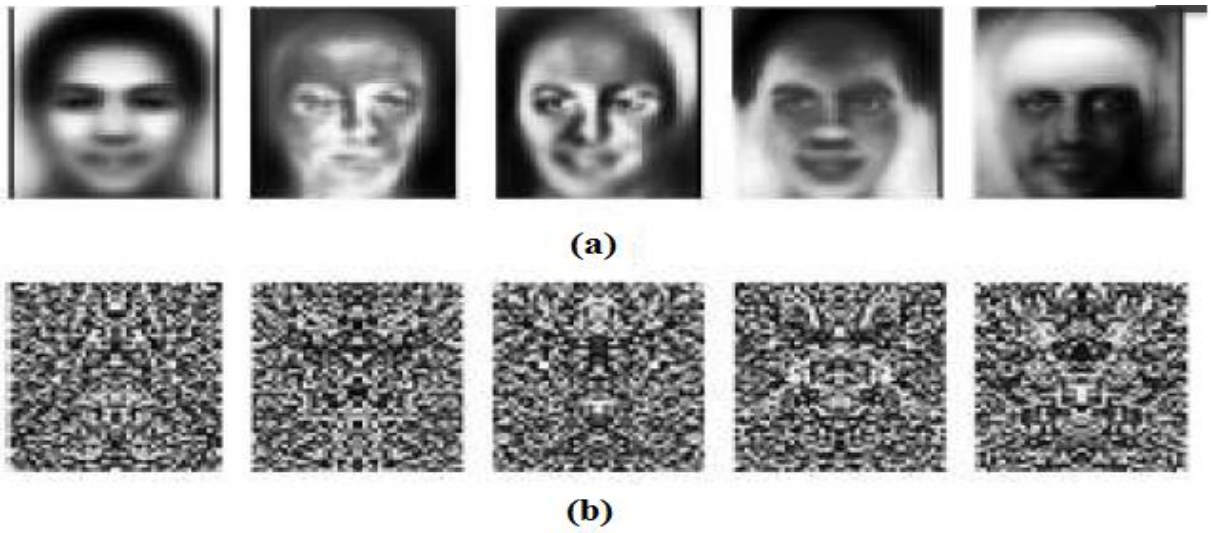


Figure. 2.7. Exemples des 5 premiers visages propres ACP (a) et LDA (b).

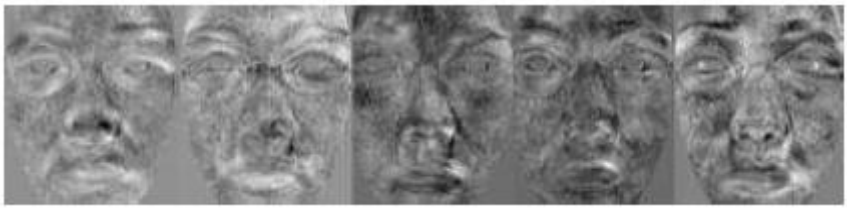


Figure. 2.8. Exemples des 5 premiers visages propres Fisherfaces.

Dans le même sens Bartlett et al. ont proposé dans [38] d'utiliser l'ICA (Independant Component Analysis) pour extraire des caractéristiques pour l'identification de visage. Au lieu de décorréler les données comme pour la PCA et la LDA, l'ICA cherche à les rendre statistiquement indépendantes. La méthode ICA proposée dans [39] utilise la PCA comme étape préliminaire. Deux techniques ont été proposées pour l'identification de visage : la première technique cherche à obtenir des vecteurs Eigenfaces qui soient statistiquement indépendant deux à deux, tandis que la deuxième technique cherche à rendre les variables de projection statistiquement indépendantes. La figure 2.9 montre les vecteurs de projection de l'ICA par les deux techniques. Néanmoins, les résultats expérimentaux montrent que l'ICA n'apporte pas d'amélioration conséquente par rapport à la PCA [40], pire elle engendre une dégradation des performances d'après les résultats obtenus par Bartlett et al. [38].

(a) (b)

Figure. 2.9 : 6 premiers axes de l'espace de visage ICA (a) Technique I (b) Technique II.

Le problème qui se pose pour les méthodes de réduction des dimensions comme (PCA, ICA, LDA) est que sur les bases de données dynamiques, on a tout le temps besoin d'ajouter ou de supprimer de nouvelles personnes, ce qui exige à chaque fois une mise à jour de la base de projection. Pour cela, il est intéressant de trouver des possibilités adaptatives pour ces méthodes quand des images de nouvelles classes sont ajoutées.

II-2- Méthodes basées sur les caractéristiques locales

Les approches basées sur l'extraction de points caractéristiques locaux peuvent être séparées en deux catégories: les approches géométriques et les approches basées sur les graphes. Les approches géométriques sont basées sur l'extraction de la position relative des éléments qui constituent le visage (tel que le nez, la bouche et les yeux). Dans la catégorie de ces approches, Brunelli et Poggio [41] ont réalisé un système de reconnaissance faciale qui détecte automatiquement les coordonnées des repères du visage (Nez, Coins de la bouche, Centres des yeux, etc.) pour former 35 caractéristiques géométriques du visage. La ressemblance est calculée en utilisant le classifieur de Bayes. Un taux d'identification de 90 % sur une base de données de 47 personnes a été trouvé. L'avantage des approches géométriques est que le coût de stockage est très bas comparé à celui des autres techniques, mais leur inconvénient est que l'extraction des caractéristiques géométriques est difficile à réaliser.

Dans les approches basées sur les graphes, Lades et al. [42] ont proposé une méthode topologique élastique, nommée « Elastic Graph Matching (EGM)»; cette méthode fait joindre à chacun des points locaux du visage, préalablement détectés, des coefficients de Gabor [43] pour représenter une personne. Par la suite, une version améliorée de la technique EGM, nommée d'Elalstic Buch Graph Matching (EBGM) a été proposée dans [44]. La différence étant que l'EBGM utilise un seul graphe pour représenter les différentes variations dans l'apparence du visage (voir Figure 2.10). Dans le cas de l'EBGM, chaque nœud contient un ensemble de 40 coefficients complexes des ondelettes de Gabor. Ces coefficients s'appellent un jet. Aussi, la géométrie d'un visage est codée par les arêtes du graph qui sont les distances entre les différents jets. Ainsi, les nœuds contiennent les variations des niveaux de gris, par exemple un nœud « eye bunch » contient des jets calculés par les différents filtres de Gabor appliqués sur la région des yeux. Un taux d'identification de 95.25 % a été obtenu par cette

méthode, en utilisant la base de données ORL avec seulement une image d'apprentissage par personne.

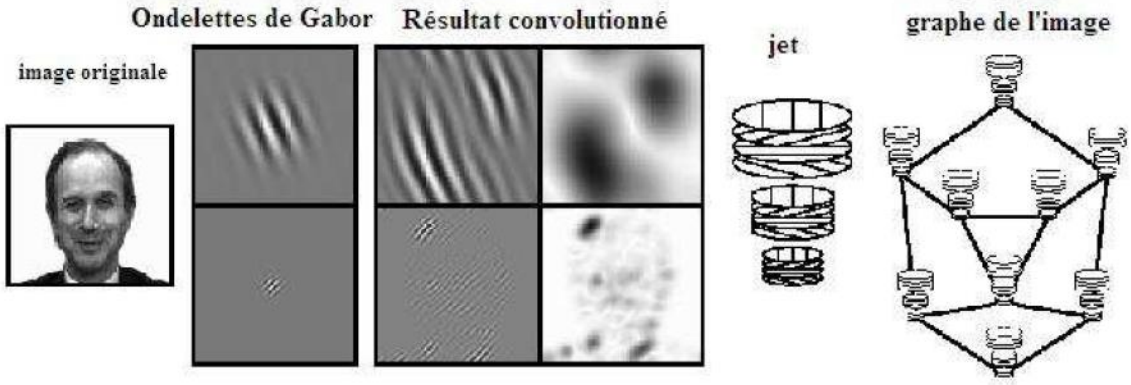


Figure. 2.10. Représentation de l'Elalstic Buch Graph Matching pour un visage.

Les méthodes basées sur les caractéristiques locales sont efficaces. Cependant leurs performances dépendent essentiellement de la précision dans la localisation des points d'intérêt du visage. Cette tâche reste très difficile en pratique.

II-3- Méthodes basées sur les caractéristiques hybrides

Ces techniques utilisent les méthodes globales de manière modulaire sur les différentes régions du visage. Ainsi, elles exploitent mieux le fait que les différentes régions faciales ne soient pas 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.

Dans cette catégorie, on trouve la DCT proposée dans [45]. La DCT est à la base la compression avec perte des signaux temporels, comme les sons audio (extension MP3), les images fixes pour (extension JPEG) et les vidéos (normes MPEG). La DCT est souvent utilisée pour une modélisation paramétrique de petits blocs carrés (8 * 8 ou 16 * 16) de l'image. Il a fallu attendre jusqu'à 2001 pour voir la DCT s'appliquer dans les systèmes de reconnaissance du visage et ceci après le développement des algorithmes rapides du calcul de la DCT. Cette première utilisation a été présentée dans [17]. Après ce premier succès, la DCT a été associée à d'autres types de transformations, comme dans [47] où les auteurs ont combiné les transformations en ondelettes DWT et la DCT. Plusieurs algorithmes ont ensuite été appliqués avec la DCT, on peut citer la DCT+LDA, DCT+ACP 2D et DCT+PCA. Les auteurs dans [48] ont conclu que la combinaison la plus stable est celle du duo DCT+PCA.

Toujours dans les approches hybrides, Pentland et al. ont introduit dans [20] l'approche « Modular Eigenspase ». Dans cette approche, une ACP est appliquée sur chacune de régions

du visage, comme la région des yeux, la région du nez. La bouche étant trop sensible à des changements des expressions faciales, sa prise en compte se traduit par une baisse du taux de reconnaissance. Cette approche peut être qualifiée d'hybride, puisqu'elle utilise à la fois les caractéristiques globales et les caractéristiques locales.

III-Extraction des caractéristiques du visage par les moments polynomiaux

Comme déjà mentionné ci-dessus, un système de reconnaissance comporte deux phases: la phase d'extraction des caractéristiques et la phase de classification. L'extraction des caractéristiques consiste à obtenir des informations qui doivent être discriminantes et non redondantes, pour l'étape de classification. De nombreuses caractéristiques peuvent être utilisées pour la reconnaissance du visage. Dans le domaine du traitement de l'information, les moments polynomiaux et circulaires sont largement utilisés pour leur propriété d'orthogonalité, qui permet la génération de descripteurs non redondants, et pour leurs propriétés d'invariance en translation, en échelle et en rotation. Ces moments ont été appliqués par exemple pour la reconnaissance des images de personnes, l'indexation des images dans les bases de données, ainsi que pour l'analyse et la description des formes des objets 2D ou 3D. Ces caractéristiques comprennent les moments polynomiaux, tels que les invariants des moments de Legendre (LMI), les moments circulaires polynomiaux, les pseudo moments de Zernike (PZM) et les moments obtenus par l'ART. Dans notre travail de magister on a déjà développé un système de reconnaissance du visage par PMZ [83] et les LMI ont déjà été proposés pour la reconnaissance du visage [84]. Dans Cette thèse on propose d'utiliser les moments ART pour cette tâche.

Hu [52] a été le premier à présenter l'utilisation des moments invariants d'image dans les applications de reconnaissance de motifs 2D. Parmi les moments les plus populaires utilisés comme caractéristiques pour la reconnaissance du visage, on trouve:

III-1 Moments de Legendre

Annadurai et al. [24] sont les premiers à avoir utilisé les moments invariants de Legendre (LMI) pour la décomposition polynomiale d'une image en niveaux de gris. Les vecteurs caractéristiques de la fonction d'intensité f (x, y) extraits avec le moment 2D de Legendre sont définis par Teague dans [25]:

$$L_{m,n} = \lambda_{m,n} \int_{-1}^{1} \int_{-1}^{1} . P_m(x) . P_n(y) f(x,y) dxdy \,, \qquad (3.8)$$

Pour une image I(i, j) carrée NxN, l'équation précédente s'écrit:

$$L_{m,n} = \lambda_{m,n} \sum_{i=1}^{N} \sum_{j=1}^{N} I(i,j) \,.\, P_m(x_i).\, P_n(y_j) \qquad (3.9)$$

Le coefficient de normalisation $\lambda_{m,n}$ dans cette équation est donné par:

$$\lambda_{m,n} = \frac{(2m+1)(2n+1)}{(N-1)^2} \qquad (3.10)$$

Le moment polynomial $P_m(x)$, représente le polynôme de Legendre d'ordre m, donné par:

$$P_m(x) = \sum_{k=0}^{m} C_{mk}\left[(1-x)^k + (-m)^k(1+x)^k\right], \qquad (3.11)$$

avec:

$$C_{mk} = \frac{(-1)^k(m+k)!}{2^{k+1}(m+k)!(k!)^2} \qquad (3.12)$$

Pour que les polynômes de Legendre soient orthogonaux ils doivent être calculés dans l'intervalle [-1, 1]. Pour cela une image carrée de (N x N) pixels avec la fonction d'intensité I (i, j) doit être rééchelonnée dans l'intervalle -1≤ x, y≤ 1 comme suit:

$$x_i = \frac{2i-N-1}{N-1}, y_j = \frac{2j-N-1}{N-1} \qquad (3.13)$$

où *i* et *j* sont les coordonnées de la position d'un pixel de l'image originale, et x_i et y_j sont les nouvelles coordonnées de ce pixel dans le nouveau repère.

La figure 2.11 représente la base 2D des moments de Legendre, avec n = 0: 2 et m = 0: 4.

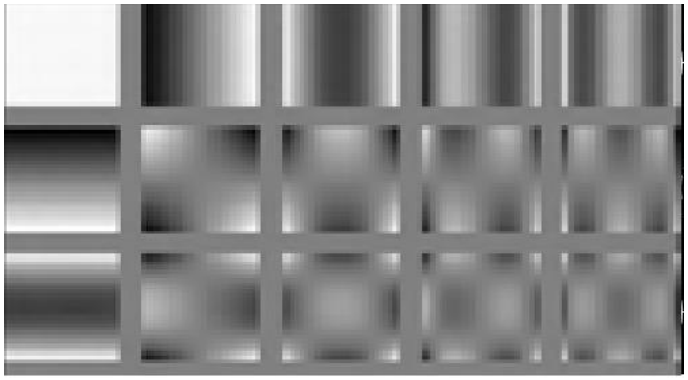


Figure. 2.11. Base 2D LMI, avec n = 0:2 et m = 0:4.

L'implémentation des LMI est détaillée dans l'annexe C sous forme d'algorithme.

III-Moments de zernike et les pseudos moments de zernike

Les moments de Zernike ont été développés par F. Zernike [55]. L'expression de ces moments est donnée par l'équation suivante :

$$Z_{m,n}=\frac{n+1}{\pi}\iint_{xy}^{x^2+y^2\le 1} f(x,y)\,[V_{n\,m}\,(x,y)]^{*}dxdy, \qquad (3.14)$$

où $[.]^{*}$ est utilisé pour indiquer la valeur complexe conjuguée, n représente l'ordre de décomposition radial (n=0,1,2,…,) et m représente le nombre de répétitions de la décomposition par rapport à l'ordre de décomposition n. L'ordre et la répétition sont liés par les deux conditions suivantes :

$$(n - |m| \text{ modulo } 2)=0 \quad \text{et} \quad |m| \le n \qquad (3.15)$$

$V_{nm}(x, y)$ représentent les polynômes de Zernike constituants la base orthogonale de projection. Ils s'écrivent en général en représentation polaire sous la forme suivante :

$$V_{nm}(r,\theta) = R_{nm}(r).exp(jm\theta), \qquad (3.16)$$

où $R_{nm}(r)$ sont des polynômes radiaux donnés par:

$$R_{nm}(r)=\sum_{k=|m|}^{n}\frac{(-1)^{(n-k)/2}\,(n+k)!}{\frac{(n-k)}{2}!\,\frac{(k+m)}{2}!\,\frac{(k-m)}{2}!}r^{k} \qquad (3.17)$$

L'application des moments de Zernike à une fonction discrète $I(x,y)$ nécessite la réécriture de (3.18) comme suit :

$$Z_{nm}=\frac{n+1}{\pi}\sum_{x}\sum_{y} I(x,y)[v_{nm}(x,y)]^{*} \quad \text{où} \quad x^2+y^2\le 1 \qquad (3.18)$$

La formulation classique des moments de Zernike détaillée ci-dessus est très facile à mettre en œuvre algorithmiquement. Cependant, elle reste très coûteuse en termes de temps de calcul et se prête mal à un traitement rapide [56]. Pour pallier les inconvénients des moments de Zernike les chercheurs ont développé un autre type de moments de Zernike, beaucoup plus performants et stables, appelés les Pseudos-Moments-de-Zernike (PZM). La différence par rapport aux moments de Zernike se situe dans les polynômes radiaux, $R_{nm}(r)$, qui deviennent:

$$R_{nm}(r)=\sum_{k=0}^{n-|m|}(-1)^k \frac{(2.n+1-k)!}{(k)!(n+|m|-k)!\,(n-|m|-k)!} r^{n-k} \quad (3.19)$$

Et la relation entre n et m qui devient $|m| \leq n$ seulement. La formule générale des Pseudo-Moment-de-Zernike suivante est donnée dans [57] :

$$PZ_{nm}=\frac{n+1}{\pi}\sum_{k=0}^{n-|m|}(-1)^k \frac{(2.n+1-s)!}{(k)!(n+|m|-k)!\,(n-|m|-k)!} r^{n-k}\left(\sum\sum_{x^2+y^2} I(x,y)exp(-jm\theta)\, r^k\right) \quad (3.20)$$

En utilisant les deux équations suivantes,

$$P\beta_{n,m,k}=\sum_{k=0}^{n-|m|}(-1)^k \frac{(2.n+1-k)!}{(k)!(n+|m|+1-k)!\,(n-|m|-k)!} \quad (3.21)$$

$$X_{m,k}=\sum\sum_{x^2+y^2} I(x,y).exp(-jm\theta\,).r^k \quad (3.22)$$

L'équation (3.20) peut être réécrite sous la forme :

$$PZ_{nm} = \frac{n+1}{\pi}\sum_{k=0}^{n-|m|} P\beta_{n,m,k}.X_{m,k} \quad (3.23)$$

Les expressions des polynômes radiaux pour les ordres de n=0 à n=3 sont données ci-après:

$R_{0,0} = 1$

$R_{1,0} = -2 + 3r$

$R_{1,1} = r$

$R_{2,0} = 3 + 10r^2 - 12r$

$R_{2,1} = 5r^2 - 4r$

$R_{2,2} = r^2$

$R_{3,0} = -4 + 35r^3 - 60r^2 + 30r$

$R_{3,1} = 21r^3 - 30r^2 + 10r$

$R_{3,2} = 7r^3 - 60r^2 + 10r$

$R_{3,3} = r^3$

La figure 2.12 représente la base 2D PMZ, avec n = 0: 4 et m = 0: n.

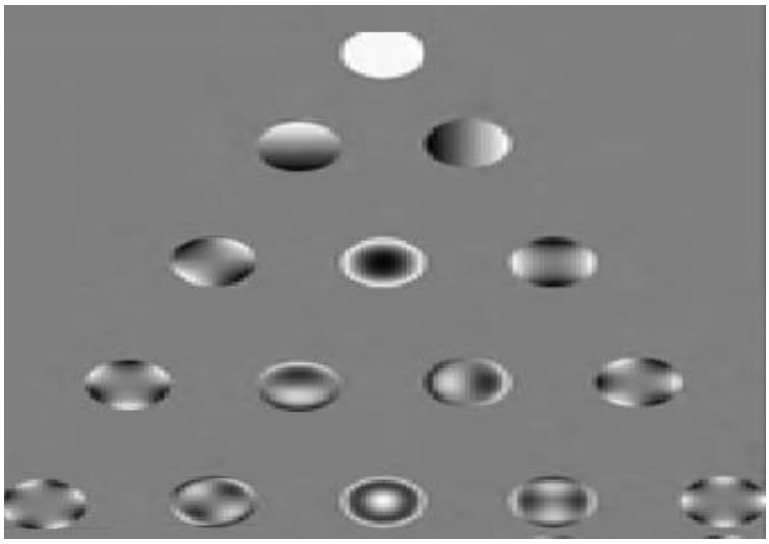


Figure. 2.12. Base 2D PMZ, avec n = 0 : 4 et m = 0 : n.

L'implémentation des PMZ est détaillée dans l'annexe C sous forme d'algorithme.

III-3 Base de projection ART

Une nouvelle approche est apparue pour la décomposition des images en niveaux de gris en utilisant le principe des polynômes radiaux. Cette approche utilise les moments obtenus par la Transformation Angulaire Radial (ART) pour représenter une image. La Transformation Angulaire Radial (ART) est utilisée dans de nombreuses applications intelligentes, telles que les systèmes de vidéosurveillance [58], la reconnaissance de logo [59] et la détection de visage, avec un taux de détection correcte de 88,7% [60].

La première utilisation des ART était en 2001 par M. Bober dans [61], qui a utilisé les ART comme descripteur de forme, basé région, pour la standardisation de la compression MPEG-7. Les vecteurs caractéristiques des images de visage, extraits avec l'aide de l'ART, sont des projections orthogonales de l'image du visage sur une fonction de base radiale. Cette projection est définie par:

$$W_{m,n} = \int_0^{2\pi} \int_0^1 f(r,\theta).V_{nm}(r,\theta)drd\theta \qquad (3.24)$$

L'application de l'équation (3.24) à une fonction discrète I (x, y) nécessite de la réécrire comme suit:

$$W_{m,n} = \sum_{r\leq 1} \sum_{\theta\leq 2\pi} I(r,\theta) \,.\, [V_{n\,m}(r,\theta)]^*, \qquad (3.25)$$

où la notation * représente le conjugué d'un nombre complexe.

La fonction de base radiale ART se compose de deux fonctions qui sont écrites, en général, comme suit:

$$V_{n,m}(r,\theta) = A_m(\theta).R_n(r) \qquad (3.26)$$

où $A_m(\Theta)$ est une fonction exponentielle qui assure l'invariance par rapport à la rotation:

$$A_m(\theta) = \frac{1}{2\pi}.\exp(jm\theta), \tag{3.27}$$

et $R_n(r)$ est la fonction radiale, définie par:

$$R_n(r) = \begin{cases} 1 & \text{pour} \quad n = 0 \\ 2.\cos(n.r.\pi) & \text{pour} \quad n \neq 0 \end{cases} \tag{3.28}$$

La figure 2.13 représente la base de projection ART 2D, avec n = 0: 2 et m = 0: 4.

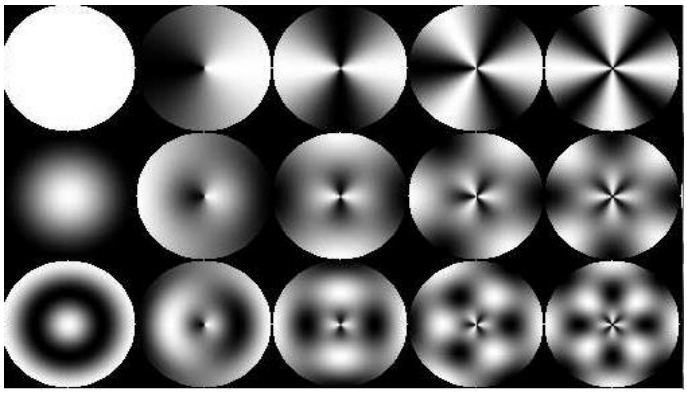


Figure. 2.13. Base de projection 2D ART,
avec n = 0 : 2 and m = 0 : 4.

Afin de conserver l'orthogonalité de la base, la fonction *I(x,y)* doit être recalculée à l'intérieur du cercle unité. Pour cela les coordonnées cartésiennes des positions des pixels sont transformées en représentation polaire (r,θ) tel que le centre de l'image soit le centre du cercle unité. L'équation (3.25) exige que l'image *I(x,y)* soit représentée à l'intérieur du cercle unité, par transformation approprié (mise à l'échelle). Cette transformation est donnée par l'équation (3.29)

$$\begin{cases} x_i = c + \frac{i(d-c)}{N-1} \\ y_j = d - \frac{j(d-c)}{M-1} \end{cases} \tag{3.29}$$

où i et j sont les coordonnées du pixel de l'image originale, x_i et y_j sont les nouvelles coordonnées de ce pixel dans le nouveau repère (le cercle unité) après transformation, M et N sont, respectivement, les largeurs horizontale et verticale de cette image et c et d sont les

paramètres qui permettent de faire le choix entre recalculer la fonction *I(x,y)* complètement ($c = -\sqrt{1/2}$ et $d = \sqrt{1/2}$) ou partiellement (c = -1 et d = 1) dans le cercle unité comme le montre la Figure (2.14) . Recalculer partiellement implique l'élimination de certain points de l'image (le coin), tandis que recalculer complètement implique l'introduction de points étrangers à l'image dont l'intensité est fixée à 0.

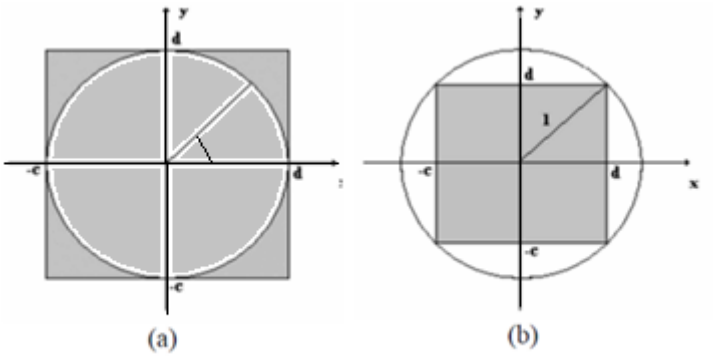


Figure 2.14. Deux méthodes pour recalculer la fonction I (x, y) transformée (rectangle gris) dans le cercle de l'unité: a) c = -1 and d =1, b) c = -1/ $\sqrt{2}$ and d =1/$\sqrt{2}$.

Le passage en coordonnés polaires se fait selon l'équation (3.30).

$$\begin{cases} r_{ij} = \sqrt{x_i^2 + y_j^2} \\ \\ \theta_{ij} = tan^{-1}\left(\frac{y_j}{x_i}\right) \end{cases} \tag{3.30}$$

Les moments circulaires, $W_{m,n}$, obtenus par la projection orthogonale de l'image du visage sur la fonction de base radiale, peuvent être exprimés sous forme de vecteur caractéristique d'une image du visage de différentes manières:

(1)$\{\Re(P_{nm}), \Im(P_{nm})\}$: un vecteur complexe à deux dimensions est converti en un vecteur réel à une dimension.

(2) $\|P_{n,m}\|^2$: un vecteur d'amplitude qui est le module du vecteur complexe

(3) arg(P_{nm}): la phase ou l'argument du vecteur complexe.

(4) $\{\|P_{n,m}\|^2, \arg(P_{nm})\}$: un vecteur d'amplitude et d'argument du vecteur complexe.

La première représentation a été utilisée dans notre travail, car elle combine à la fois les parties réelles et les parties imaginaires dans un vecteur caractéristique. Cette représentation préserve la phase et évite les calculs complexes par rapport à d'autres conversions.

L'implémentation de l'ART utilisant cette représentation est détaillée dans l'annexe C sous forme d'algorithme.

III-4 Comparaison entre les différentes bases de projection

-La base de projection LMI se compose de deux polynômes réels, $P_m(x_i)$ et $P_n(y_j)$, qui sont indépendants. Les moments qui en résultent sont ainsi réellement plus informatifs.

- Les bases de projection PMZ et ART sont le produit d'une fonction complexe et d'un polynôme radial ; les moments qui en résultent sont complexes.
- La base de projection LMI est orthogonale sur l'intervalle [-1, 1].
- Les bases de projection PMZ et ART sont orthogonales sur le cercle d'unité.
- L'ordre de répétition, m, est indépendant de l'ordre de décomposition, n, dans les bases de projection LMI et ART.

-Tous ces moments sont à l'origine invariants uniquement par rapport à la rotation. Une étape de normalisation initiale est nécessaire pour les rendre invariants par rapport à l'échelle et à la translation.

IV-Etat de l'art des méthodes de classification

Après la phase d'extraction des caractéristiques du visage on utilise la classification pour décider si un visage est membre de la base de données et à quelle classe il appartient. Les techniques proposées dans la biométrie pour résoudre les problèmes de la classification automatique de données sont un domaine de recherche qui a été largement étudié durant les dernières années. L'étape de classification est influencée principalement par la méthode utilisée lors de l'étape d'extraction de caractéristiques.

Nous allons passer en revue les principales parmi ces méthodes, à l'exception du classifieur à base du Maximum de Vraisemblance qui sera détaillé au chapitre suivant.

IV-1 Méthode basée sur la mesure de correspondance par le calcul des distances

La méthode de calcul des distances est utilisée pour trouver le degré de similitude entre un vecteur caractéristique d'une image test et un ou plusieurs vecteurs caractéristiques d'images de références [62]. La mesure de similitude par la distance euclidienne ou la distance de Mahanalobis a été testée pratiquement avec toutes les méthodes d'extraction de caractéristiques proposées dans les systèmes de reconnaissance du visage [63]. La distance euclidienne est la méthode de classification la plus utilisée dans les systèmes de reconnaissance qui utilisent l'extraction de caractéristique par Eigenface ou FisherFaces. Cette distance est un cas particulier de la distance de Minkowski. Pour définir cette distance, considérons deux vecteurs $X = (x_1, x_2, ..., x_M)$ et $Y = (y_1, y_2, ..., y_M)$ dans un espace Euclidien R^M (M étant la dimension de l'espace Euclidien), La distance de Minkowski

d'ordre p, nommée Lp, est donnée par l'équation suivante :

$$Lp=(\sum_{i=1}^{M}|x_i - y_i|^p)^{1/p} \quad (3.31)$$

Pour p = 1, on obtient la distance de Manhattan :

$$L1=\sum_{i=1}^{M}|x_i - y_i| \quad (3.32)$$

Pour p = 2, on obtient la distance Euclidienne :

$$L2=\sqrt{\sum_{i=1}^{M}|x_i - y_i|^2} \quad (3.33)$$

La figure 2.15 donne le schéma général du processus de classification d'une image test dans un système de reconnaissance du visage par distance euclidienne. Le principe consiste à calculer la distance euclidienne moyenne entre le vecteur caractéristique de l'image test et les vecteurs caractéristiques des images référence de chaque classe de la base de données. La classe de l'image test correspond à la classe qui a obtenu la plus petite distance euclidienne, à condition que cette distance soit inférieure à un certain seuil de décision.

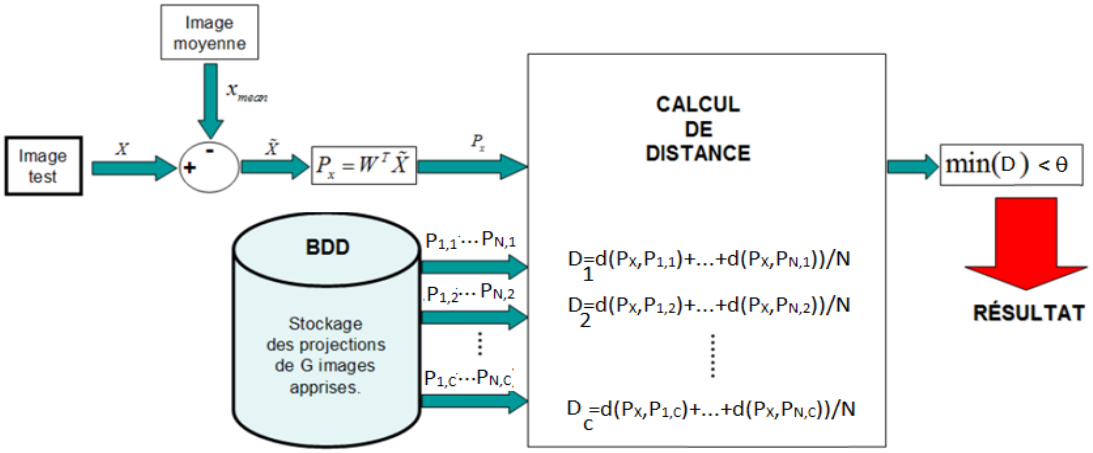


Figure. 2.15. Classification des images dans un système de reconnaissance du visage par distance Euclidienne.

IV-2 Méthode basée sur la mesure de dissimilation métrique

Cette technique de classification est essentiellement utilisée pour mesurer la dissemblance des d'échantillons qui utilisent des histogrammes comme modèles. Il existe de nombreuses mesures pour évaluer la similitude entre deux histogrammes, tels que l'intersection de l'histogramme, le Ratio de vraisemblance et la statistique chi-carré [73]. Par exemple, pour la classification par la statistique chi-carré un échantillon S est assigné à la classe du modèle M qui minimise la distance suivante :

$$D(S,M) = \sum_{i=1}^{N} \frac{(S_i - M_i)^2}{(S_i + M_i)} \quad (3.34)$$

Où N est la taille des histogrammes, S_i et M_i sont les valeurs des échantillons et des histogrammes modèles à la $n^{ème}$ position.

IV-3 Réseaux de neurones

Grace à leur pouvoir discriminant, les réseaux de neurones RNA ont été utilisés par de nombreux systèmes de reconnaissance de visage (Figure 3.14). Mazloom et Ayat dans [64] ont réalisé un réseau de neurones multicouches pour classifier des vecteurs caractéristiques issus d'Eigenface. L'architecture générale des réseaux de neurones consiste en la représentation des neurones en couches successives, la première représentant la couche d'intrants (input layer), la dernière étant la couche de sortie (output layer), les couches intermédiaires étant les couches cachées (hidden layer) du réseau.

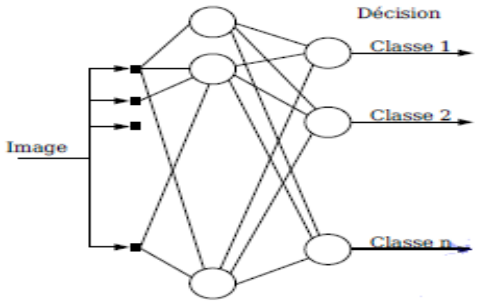


Figure 2.16. RNA pour la reconnaissance de visages

En général, un réseau de neurones artificiel peut être définit par les éléments suivants :

-Nature des entrées : Elles peuvent être soit binaires, ou bien réelles.

-La fonction d'entrée totale : Elle définit le prétraitement effectué sur les entrées.

-La fonction d'activation : Les cellules des réseaux de neurones (les cercles dans la figure 2.16) sont généralement des fonctions d'activation. Les fonctions d'activation sont en général des fonctions non linéaires monotones croissantes, qui peuvent être déterministes, continues, discontinues ou aléatoires. Les principales fonctions d'activation des cellules des neurones sont les fonctions linéaires, les fonctions seuils, les fonctions marches, les fonctions sigmoïdes et les fonctions radiales.

Comme autres méthodes proposées pour classifier des vecteurs caractéristiques issus de la méthode ICA [36] ou la méthode DCT [37], on trouve les RNA discriminants dont les fonctions d'activation sont des fonctions de type bayésiennes radiales (réseau RBF).

IV-4 Supports Vector Machines (SVM)

Guo et Coll [67] ont proposé d'utiliser les classifieurs SVM pour la reconnaissance faciale. Les SVM sont une technique efficace pour classifier implicitement des éléments sous-jacents répartis dans l'espace des attributs en utilisant une fonction de décision de type hyper-plan. Ce sont des classifieurs binaires qui permettent de traiter des problèmes de classification

linéaire en formulant le problème de classification comme un problème d'optimisation quadratique.

Osuna et al. [68] sont les premiers à avoir appliqué les SVM pour la classification des images de visages. Les résultats obtenus ont illustré le succès de cette approche de classification dans la biométrie. De leur côté, Huang et al. [69] ont utilisé la classification par SVM pour la détection de pose dans les images. Par la suite, Heisele et al. [70] ont utilisé cette technique de classification pour la détection de la position du visage dans une image (classification "visage" ou "non-visage"). L'efficacité de la discrimination par les SVM a été prouvée dans la reconnaissance du visage par plusieurs études. Par exemple, Huang et al. ont obtenu un taux de reconnaissance de 100% sur les visages de la base de données FERET. Avec la même technique, mais avec d'autres protocoles de test (1.355 images issues des bases de données AR, AT&T, YaleFaces), les taux changent selon le nombre de classes. Pour 10 classes, les SVM atteignent 98.89% de taux de reconnaissance, 1.85% de taux de faux négatifs et 0% de taux de faux positifs; pour 60 classes, ces taux deviennent 95.57%, 2.71% et 0.74%, respectivement. Des SVM à noyau non-linéaire ont également été utilisés pour la reconnaissance de visages.

V-Conclusion

Dans ce chapitre, nous avons mis en lumière l'ensemble du processus automatique de reconnaissance de visages, allant de l'extraction de caractéristiques à la classification.

La première partie, a été consacrée à un état de l'art sur l'extraction de caractéristique de visage. Nous avons ensuite étudié en détails plusieurs méthodes d'extraction de caractéristiques de visage qui utilisent la décomposition polynomiale de l'image, à savoir, PZM, LMI et ART. Les deux premières méthodes ont déjà été utilisées dans les systèmes de reconnaissance du visage par contre la dernière n'a jamais été utilisée auparavant.

Dans la deuxième partie nous avons présenté les méthodes de classification, en particulier la méthode du 'plus proche voisin', basée sur la distance Euclidienne et la méthode du séparateur à vaste marge, qui est plus performante en termes de taux de reconnaissance, mais plus gourmande en termes de temps de calcul. Nous avons testé ces méthodes de reconnaissance du visage sur plusieurs bases de données Comme (ORL, Essex Faces94, Essex Faces96, Yaleface). Nous avons en particulier testé leur vulnérabilité aux attaques par masques 3D.

Dans le chapitre suivant, nous présenterons une évaluation expérimentale des performances de ces méthodes.

Chapitre III :
Les contre-mesures contre les attaques par masque 3D dans RF

I-Introduction

Les attaques (spoofing en anglais) dans les systèmes de Reconnaissance Faciale 2D (RF) les plus courantes sont les attaques photographiques et vidéo, en raison de leur commodité et de leur faible coût. Sur la base des observations que les RF sont vulnérables à ces attaques, les chercheurs ont commencé à travailler sur des contre-mesures pour réduire leur impact sur les performances de reconnaissance. Lorsque des masques 3D sont utilisés pour effectuer ces attaques, certaines des mesures proposées contre les attaques 2D ne sont plus applicables. Il semblerait que la détection des attaques par masques 3D est plus difficile que la détection des attaques par vidéo où photo. Les attaques des systèmes de RF par masque 3D sont un sujet nouveau.

Dans ce chapitre, des techniques de contre-mesure sont proposées pour protéger les systèmes de reconnaissance du visage 2D contre les attaques par masques 3D. La proposition de ces techniques se justifie par le fait qu'en général les systèmes de RF sont très vulnérables à ces attaques. Pour s'en assurer on a mesuré le SFAR (Spoofing False Acceptance Rate) des deux méthodes de reconnaissances du visage étudiées dans le chapitre précédent (ART, LMI), en utilisant la base de données 3DMAD, qui est la seule base de données publique, disponible actuellement, permettant ce type d'étude. Avec un SFAR avoisinant 65% pour les deux méthodes, les résultats obtenus sont éloquents, ils mettent en évidence le danger que représente les attaques par masques 3D pour les systèmes de RF.

L'efficacité des attaques par masques 3D pour tromper les systèmes de RF est due au fait que ces masques ont une texture et surtout des caractéristiques de forme qui sont très similaires à celles d'un visage.

Toutes les contre-mesures proposées jusqu'à présent pour protéger les systèmes de reconnaissances du visage contre les attaques par masques 3D reposent sur une étape de vérification après l'étape de reconnaissance du visage, comme le montre la figure suivante.

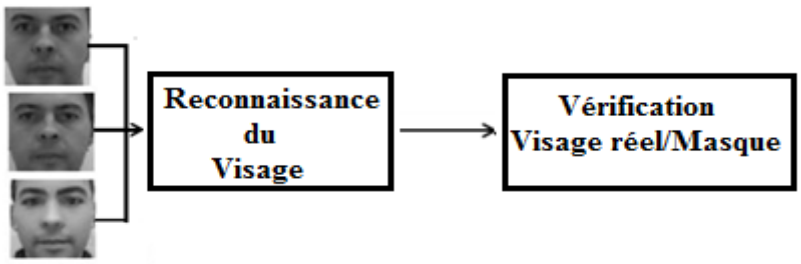


Figure. 3.1. Principe de la détection des pirates par masque 3D dans un système de reconnaissance faciale.

Les techniques de vérification utilisées sont basées sur l'analyse des différentes caractéristiques de texture et de réflexion de la lumière pour les visages réels et les masques. Dans ce chapitre, deux approches sont proposées pour contrer les attaques par masques 3D dans un système de RF 2D. La première approche suit le schéma classique d'une étape de reconnaissance suivie d'une étape de vérification, qui utilise l'ART pour l'extraction de caractéristiques et le Maximum de Vraisemblance (MV) pour la classification. Les résultats obtenus par cette méthode de vérification rivalisent avec ceux des méthodes utilisant l'analyse de texture par LBP [74]; ils montrent qu'elle est valable pour protéger n'importe quel système de reconnaissance du visage utilisant une simple webcam comme moyen d'acquisition. En effet, les images issues d'une webcam sont des images RGB et les algorithmes développés n'utilisent que ce type d'images contrairement aux autres méthodes [20,24].

La seconde approche évite le recours à deux phases, une phase de reconnaissance et une phase de vérification, elle propose un système de RF, qui possède un taux de reconnaissance très élevé, tout en étant très immune au spoofing par masques 3D. Ce système de reconnaissance faciale utilise les LMI combinés avec la LDA pour l'extraction de caractéristiques et le MV pour la classification. Les deux approches proposées peuvent être utilisées sans risque pour la santé de l'utilisateur, contrairement à la méthode proposée par Kose et Dugelay [28] qui utilise deux types d'images (image de profondeur et image RGB) capturées avec un scanner d'acquisition adéquat qui émet des rayons nuisibles pour la santé. Un autre avantage des approches proposées est l'utilisation d'images RGB uniquement, qui peuvent être acquises par une simple Webcam.

Jusqu'ici les contre-mesures proposées contre les attaques par masques 3D ont été principalement basées sur l'analyse de la texture par LBP. Le LBP standard et ses variantes sont détaillés dans l'annexe B. Dans ce qui suit, nous allons passer en revue les différentes contre-mesures basées sur le LBP proposées dans la littérature pour parer aux attaques par masques 3D, avant de décrire les deux approches que nous proposons pour cela.

II- Mesures contre les attaques par masques 3D par LBP

Dans [4], les auteurs affirment que les détails de micro-texture sont nécessaires pour discriminer entre un visage réel et un masque 3D et qu'ils peuvent être détectés efficacement en utilisant les différents opérateurs LBP.

La texture peut être caractérisée par une structure spatiale (par exemple les LBP), qui varie avec la translation et le contraste mais qui ne varie pas avec la rotation [26]. Les LBP [25]

sont l'une des approches d'analyse de texture les plus populaires, qui caractérise la structure spatiale de la texture d'image locale.

L'analyse de la micro-texture, qui a d'abord été proposée dans [26] pour détecter les attaques 2D (photo, vidéo), a été utilisée ensuite pour détecter les attaques par masques 3D. Cette technique d'analyse de micro-texture, basée sur les LBP, met l'accent sur les différences de micro-texture dans l'espace des fonctionnalités. Dans [26], les auteurs ont cherché à apprendre les différences entre le visage réel et le masque, et à concevoir un espace qui exploite ces différences.

Quatre contre-mesures ont été proposées dans [76] pour faire face aux attaques par masques 3D. Trois d'entre elles utilisent les données 2D issues des images de texture RGB et la quatrième utilise les données 3D (images de profondeur estimées à partir des balayages 3D par scanner). Dans le cas de la première contre-mesure (CM1), une analyse de micro-texture est appliquée sur les images de texture en niveau de gris, puis un histogramme de caractéristique d'une longueur de 833 est obtenu. Dans le cas de la CM1 [76], au lieu d'utiliser la méthode LBP classique pour l'extraction de caractéristique une méthode plus performante, appelée LBP uniforme ($LBP^{U}_{P,R}$)), a été utilisée. En effet, pour ne pas utiliser tous les motifs des pixels de l'image à partir du LBP classique, des informations suffisantes peuvent être obtenues en utilisant uniquement des modèles uniformes.

Comme montré à la Figure 3.2, le vecteur caractéristique utilisé dans la CM1 [76] pour la classification est un histogramme global de 833 éléments, composé de 3 histogrammes locaux. Le premier et le deuxième histogramme sont calculés à partir de l'image du visage entier en utilisant les opérateurs $LBP^{U2}_{8,2}$ et $LBP^{U2}_{16,2}$ ce qui donne des histogrammes d'une longueur égal à 59 et 243, respectivement. Le troisième histogramme est calculé en appliquant l'opérateur $LBP^{U2}_{8,2}$ à l'image du visage, et en divisant l'image LBP obtenue en régions de taille 3 x 3 qui se chevauchent. Les histogrammes locaux des différentes régions, d'une longueur de 59 éléments chacun, sont groupés pour former un histogramme unique composé de 531 éléments. La longueur de l'histogramme final est égale à 833 (c'est-à-dire 531 + 59 + 243).

Concernant la seconde contre mesure (CM2) [77], l'analyse de micro-texture est appliquée sur des images de texture de réflexion plutôt que sur les images de texture elles-mêmes. La raison est que l'analyse révèle que les composantes des images de réflexion des visages réels et des

masques 3D sont différentes. Les composantes d'éclairage et de réflexion des images de texture sont obtenues en utilisant l'algorithme variationnel de retinex[80]. Des vecteurs caractéristiques d'une longueur égale à 833 ont été obtenus en appliquant la technique de la première contre mesure aux images de réflexion.

Pour la troisième contre mesure (CM3) [75], les auteurs ont utilisé les valeurs d'intensité de la composante réflexion de l'image comme entrée pour le classificateur SVM linéaire directement. Les valeurs d'intensité de la composante réflexion de l'image, comprises entre 0 et 1 (R (x, y) ∈ [0,1]), ont été mises à l'échelle en les multipliant par 255. Le vecteur caractéristique résultant d'une image de taille 64 × 64 pixels est un vecteur caractéristique de taille 4096 (64 × 64 = 4096), fournissant des informations concernant le niveau d'intensité de réflexion de chaque pixel.

Enfin, dans la quatrième contre-mesure (CM4)[76], l'analyse de micro-texture est également appliquée à des images de profondeur qui sont obtenues à partir des balayages 3D bruts et un autre histogramme de longueur 833 est obtenu et appliqué à un classificateur SVM linéaire pour discriminer entre les masques et les visages réels.

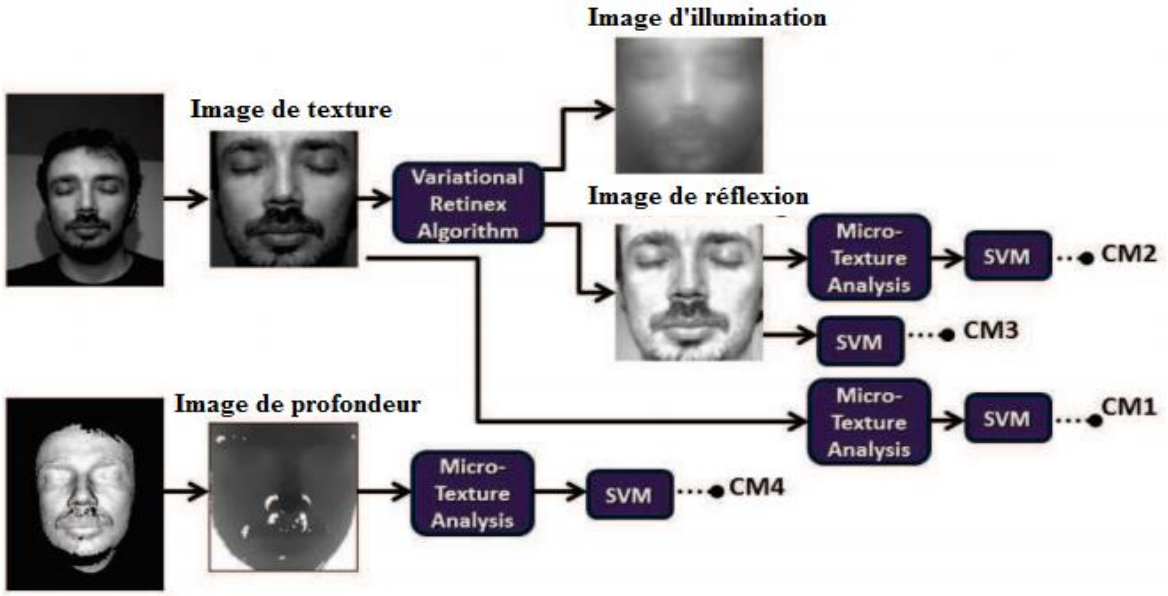


Figure. 3.2. L'organigramme des contre-mesures proposées par [77]

Selon les résultats présentés dans [77] et reproduits dans le tableau 4.1 ci-dessous, les meilleures performances, en termes d'HTER et de précision, sont obtenues avec les contre-mesures basées sur l'analyse de la réflexion (CM2 et CM3). Le HTER (Half Total Error Rate) est la moyenne de deux taux d'erreur : le taux des faux visages FLR (False Living Rate), qui représente le pourcentage des masques classés en tant que visages réels et le taux du faux

masque FFR (False Fake Rate), qui représente le pourcentage de visages réels classés en tant que masques. La précision quant à elle représente le taux de détection des masques.

L'analyse de ces résultats montre que les caractéristiques extraites à partir des images de réflexion par les différents LBP fournissent des informations plus pertinentes que les caractéristiques extraites à partir des images de texture ou de profondeur. Aussi, on peut constater que les données 2D (images de texture) sont plus appropriées que les données 3D (cartes en profondeur) pour la détection de masques.

Contre-mesures	HTER (%)	précision (%)
CM1	9.04	91.46
CM2	5.02	95.98
CM3	9.04	93.47
CM4	18.59	82.91

Tableau 4.1 HTER et précision des quatre contre-mesures proposées dans [77].

Le problème de cette étude est que les outils utilisés pour l'acquisition des images ne sont pas disponibles sur les appareils de reconnaissance du visage et les résultats ont été obtenus avec une base de données privée, ce qui ne permet pas de les vérifier.

Dans des études plus récentes, E. Nesli et M. Sébastien ont cherché dans [26] à analyser les propriétés discriminatives des caractéristiques de texture, extraites par différents opérateurs LBP, pour la classification de «visage réel» / «masque 3D». Ils ont pour cela utilisé la base de données publique 3DMAD.

Dans leurs expériences avec LBP, trois autres extensions de [26] sont évaluées: tLBP, dLBP et mLBP en plus des LBP basique et uniforme (voir figure 3.3).

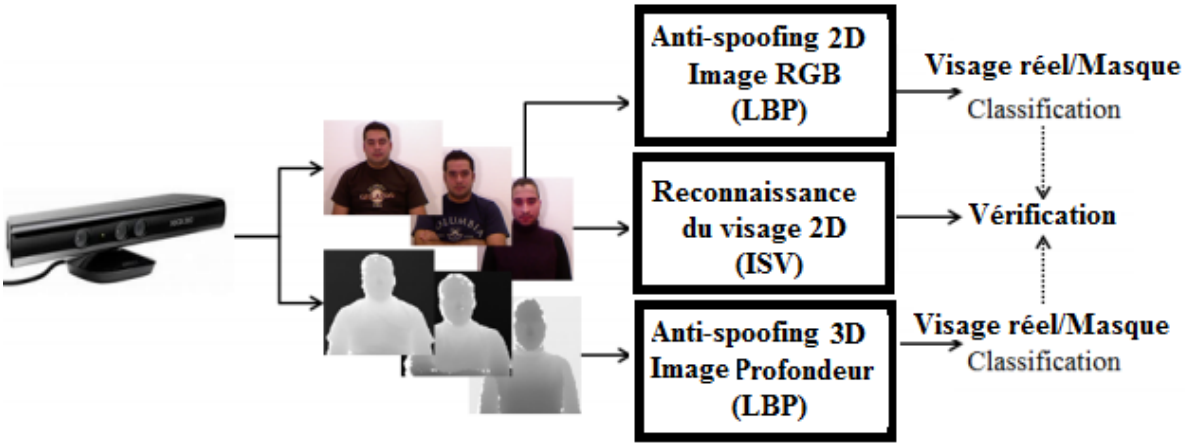


Figure. 3.3. Système de reconnaissance et les contre-mesures de protection contre les masques 3D proposées dans [26].

E. Nesli et M. Sébastien[26] ont proposé plusieurs contre mesure :

Pour leur CM1, ils ont utilisé la méthode LBP classique pour l'extraction des caractéristiques et ont formé un vecteur caractéristique de type histogramme d'une taille égale à 2^8 éléments.

Pour CM2, ils ont utilisé la méthode LBP uniforme pour l'extraction des caractéristiques et ont formé un vecteur caractéristique de type histogramme d'une taille égale à 59 éléments.

Pour CM3, CM4 et CM5, ils ont utilisé les méthodes tLBP, mLBP et dLBP pour l'extraction des caractéristique et ont formé un vecteur caractéristique de type histogramme d'une taille égale à 2^8 pour chaque méthode.

Les contre-mesures 1 à 5 sont appliquées sur des images de texture (image en niveaux de gris) et les contre-mesures de 6 à 10 sur des images de profondeurs (les contre-mesures i et i+5 utilisent la même technique pour l'extraction des caractéristiques).

De plus, l'influence de la division des images du visage en blocs a été évaluée. Pour chaque type de LBP, l'image est divisée blocs de taille en 3 × 3 et les histogrammes LBP sont calculés pour chaque bloc séparément et concaténés pour former le vecteur de caractéristiques final. Pour réduire d'avantage la taille des histogrammes la LDA a été utilisée.

Coté classification, deux classifieurs ont été testés. Dans le premier la mise en correspondance d'histogrammes par la statistique chi-carré $\chi 2$ est appliquée pour comparer les histogrammes tests avec deux histogrammes de référence, qui sont simplement calculés en prenant, dans

l'ensemble d'apprentissage, la moyenne de tous les histogrammes des visages réels pour le premier et la moyenne de tous les histogrammes des masques 3D pour le second. Le second classifieur, est le classifieur non linéaire SVM avec la RBF comme fonction noyau.
En combinant les types d'images, les méthodes d'extraction des caractéristiques et les méthodes de classification 96 contre-mesures ont été testées. Les expériences révèlent qu'il n'est pas facile de choisir une méthode parmi les quatre types de LBP. Les résultats sur cette base de données suggèrent une tendance générale qu'utiliser LBP+LDA par blocs pour donne les meilleurs résultats que ce soit pour des images en couleur ou en profondeur. Pour les images en couleur, en général l'extraction de caractéristiques par LBP + LDA et la classification par SVM fonctionne mieux que le reste. Les meilleures performances, en termes d'HTER, obtenues avec les contre-mesures basées sur mLBP + LDA, sont respectivement 0,95% et 1,27%, pour des images en couleur et en profondeur.

III-Techniques proposées contre les attaques par masques 3D

Deux approches sont proposées dans le présent travail pour parer aux attaques des systèmes de RF par masques 3D. La première approche utilise les données 2D (images de texture RGB) avec une étape de vérification après l'étape de reconnaissance, tandis que la deuxième est un système de reconnaissance qui rejette les masques sans le recours à une phase de vérification.

Comme le montre la Figure 3.5, la première approche utilise pour la vérification la décomposition polynomiale d'une image de texture en niveau de gris par l'ART pour l'extraction des caractéristiques et le maximum de vraisemblance pour la classification.

Dans la seconde approche, comme indiqué à la Figure 3.5, un système de reconnaissance du visage est conçu, qui en une seule passe assure un taux de reconnaissance très satisfaisant et en même rejette la majorité des attaques par masques 3D. Ce système utilise la décomposition polynomiale d'une image de texture en niveau de gris par LMI, combinée avec la LDA, pour l'extraction des caractéristiques, et le MV pour la classification. Cette approche sera désignée par l'abréviation LMI-LDA-ML. La décomposition par ART et LMI ayant été décrite dans le chapitre précédent, nous allons dans ce qui suit décrire en détails les autres techniques que nous avons utilisées dans les deux approches, à savoir la LDA et le classifieur MV.

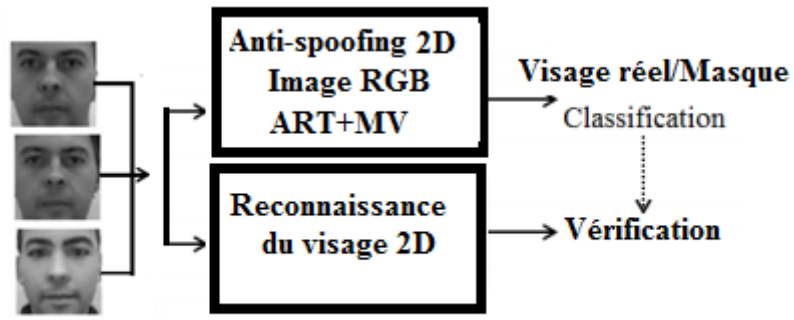


Figure. 3.4. Système de reconnaissance et de protection contre les masques 3D par ART+MV.

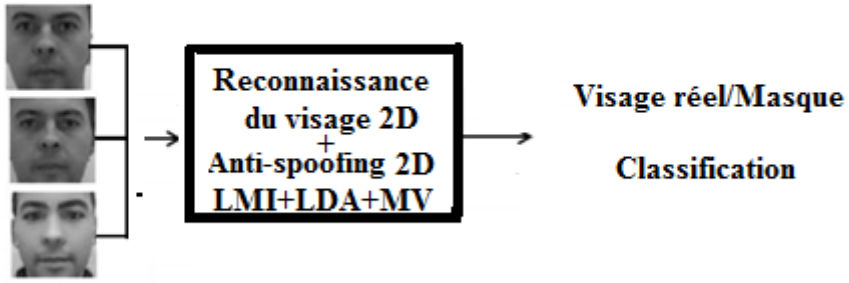


Figure. 3.5. Système 'LMI+LDA+MV' de reconnaissance et anti spoofing par masques 3D.

III-1-Analyse Linéaire Discriminante (LDA)

L'objectif de la LDA est de réduire la taille des vecteurs images ou les vecteurs issus de la décomposition polynomiale, mais avec la préservation de l'information discriminatoire.

L'algorithme LDA effectue une véritable séparation de classes comme l'illustre la Figure 3.6. Comme on peut le constater à partir de cette figure, la LDA permet de trouver la plus grande séparation entre les classes, ici le vecteur optimal W1 permet d'avoir une séparation sans chevauchement.

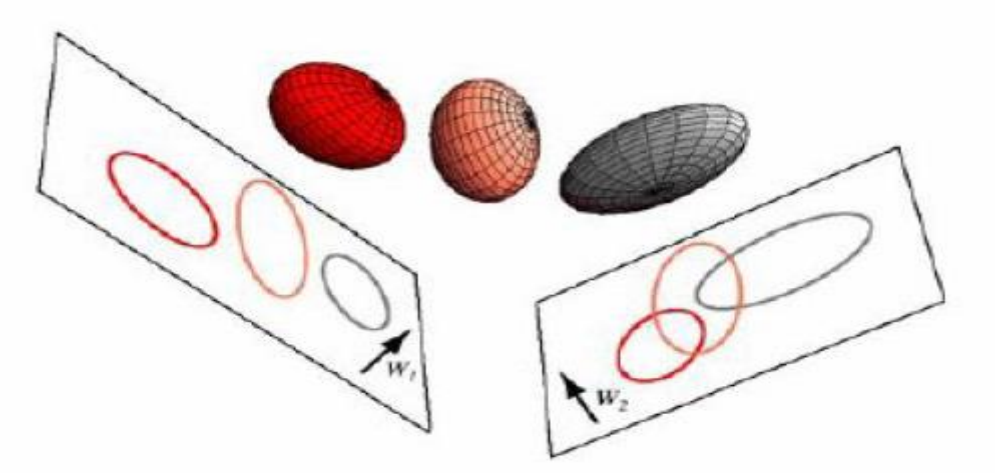


Figure. 3.6. Illustration du principe de séparation optimale des classes par la LDA.

La base de projection « W » LDA est composée de N-1 vecteurs optimaux "W_1,W_2,…,W_{N-1}", N étant le nombre de classes (personnes), à chaque classe correspond un sous ensemble composé de plusieurs images. L'organisation de la base de données 3DMAD obéit à cette règle. La LDA analyse les vecteurs propres de la matrice de dispersion des données, visant à maximiser les variations entre classes tout en minimisant les variations inter classes. Cela se réduit à trouver une base de projection W optimale, qui maximise la dispersion inter-classe, liée à la matrice Sw, et minimise la dispersion intra-classe (distance), liée à la matrice Sb.

Pour trouver la matrice de projection, on rassemble les vecteurs issus de la décomposition polynomiales LMI dans une grande matrice Γ, où chaque colonne Γ_j est un vecteur qui correspond à une image, puis on calcule le vecteur moyen Ψ de toutes les images de la base de données. Ensuite, pour chaque classe Ci, on calcule le vecteur moyen Ψ_{Ci}, comme suit:

$$\Psi_{Ci} = \frac{1}{q_i}\sum_{j=1}^{q_i}\Gamma_j , \quad (4.1)$$

où q_i est le nombre de vecteurs dans la classe Ci.

Chaque vecteur Γ_i de chaque classe Ci est ensuite recentré par rapport à la moyenne Ψ_{Ci}. On obtient alors un nouveau vecteur Φ_j centré:

$$\Phi_j = \Gamma_j - \Psi_{Ci} \quad (4.2)$$

Ensuite on calcule les différentes matrices de dispersion. On notera par c le nombre total de classes (i.e. le nombre d'individus), q_i le nombre de vecteur dans chaque classe Ci et M le nombre total de vecteurs d'apprentissage.

La Matrice de Dispersion Intra-Classe (S_w) est calculée en utilisant l'équation (4.3):

$$S_w = \sum_{i=1}^{c}\sum_{k=1}^{q_i}(\Gamma_k - \Psi_{c_i})(\Gamma_k - \Psi_{c_i})^T \quad (4.3)$$

La Matrice de Dispersion Inter-Classe (S_b) est calculée en utilisant l'équation (4.4):

$$S_b = \sum_{i=1}^{C} q_i (\Psi_{c_i} - \Psi)(\Psi_{c_i} - \Psi)^T \quad (4.4)$$

La Matrice de Dispersion Totale (S_T) est trouvée en utilisant l'équation (4.5):

$$S_T = \sum_{i=1}^{M} (\Gamma_i - \Psi)(\Gamma_i - \Psi)^T \quad (4.5)$$

Une fois ces matrices calculées, nous devons trouver une projection optimale W qui maximise la dispersion intercalasse, relative à la matrice S_w, tout en minimisant la dispersion inter-classe, relative à la matrice S_b. En d'autres termes, nous devons trouver W qui maximise le critère d'optimisation de Fisher J(W) [36]:

$$W = \arg\max (J(W)) = \frac{|W^T S_b W|}{|W^T S_w W|} \quad (4.6)$$

W peut alors être trouvé en résolvant le problème aux valeurs propres généralisé de l'équation (4.7):

$$S_b W = \lambda_w S_w W \quad (4.7)$$

Ce problème se ramène à un problème de recherche des vecteurs propres de la matrice $S_w^{-1} S_b$.

Un vecteur issu d'une décomposition polynomiale Φ_i est transformé en une composante Fisherfaces par une simple opération de projection vectorielle :

$$\Omega_i = W^T \Phi_i, \quad (4.8)$$

Pour toutes les M images de la base de données on trouve la matrice des caractéristiques Ω^T :

$$\Omega = W^T \Phi, \ \Omega^T = [\Omega_1 \Omega_2 \Omega_3 \ldots \Omega_M]^T, \text{ avec } \Phi = [\Phi_1 \Phi_2 \Phi_3 \ldots \Phi_M] \quad (4.9)$$

La matrice Ω est alors utilisée comme matrice caractéristique.
La maximisation de J(W) n'est possible que si la matrice S*w* est singulière (inversible). Cette condition ne pose pas de problème pour les matrices obtenues à partir d'une décomposition polynomiale.

La figure 3.7 donne le schéma général de la phase d'apprentissage d'un système de reconnaissance faciale utilisant les méthodes LMI+LDA.

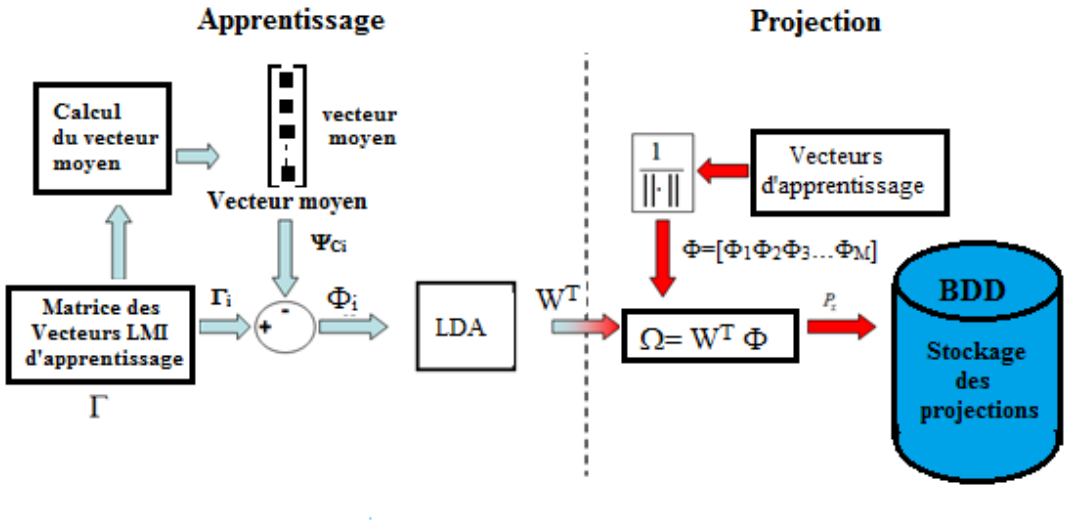


Figure. 3.7. Phase d'apprentissage d'un système de reconnaissance faciale utilisant les méthodes LMI+LDA.

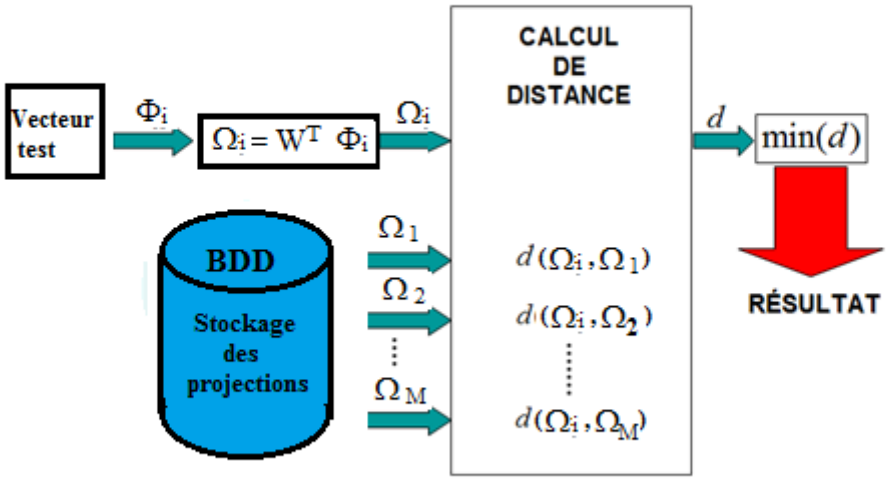


Figure. 3.8. Phase de reconnaissance d'un système de reconnaissance faciale utilisant LMI+LDA avec PPV.

III-2-Classification par maximum de vraisemblance

Le classificateur MV calcule la probabilité qu'un vecteur de fonctionnalité donnée appartienne à chaque classe et attribue à ce vecteur la classe ayant la probabilité la plus élevée.

La probabilité Gaussienne d'appartenir à une classe est plus élevée lorsqu'on s'approche du centre de la classe pris comme la moyenne de la classe.

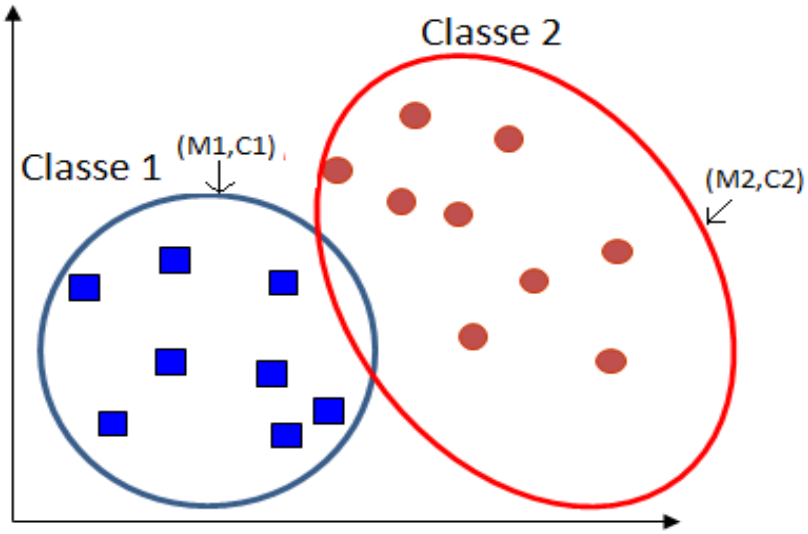


Figure .3.9. Des données 2D de deux classes modélisées par une fonction Gaussienne en fonction de la moyenne et la covariance de chaque classe.

La classification par probabilité Gaussienne est efficace lorsque la variance est étroite et que le chevauchement entre les différentes classes est faible. Toutes ces conditions sont remplies par les vecteurs caractéristiques, qui résultent de la décomposition par ART simplement ou par LMI combinée à la LDA.

Nous avons remarqué après expérimentation que la densité de probabilité d'appartenir à une classe donnée est élevée pour la personne ayant le vrai visage de cette classe et faible pour le masque 3D de la même classe. Après simulation, nous avons également remarqué que cette probabilité est pratiquement nulle pour les imposteurs.

Comme indiqué ci-dessus, pour implémenter le classificateur MV, les vecteurs moyens et les matrices de covariance de chaque classe doivent être estimés. Le vecteur moyen M_i et la matrice de covariance C_i de la classe i sont estimés comme suit:

$$M_i = \frac{1}{N} \cdot \sum_{j=1}^{N} x_{i,j} \tag{4.10}$$

$$C_i = \sum_{i=1}^{N} (X_{i,j} - M_i)(X_{i,j} - M_i)^T, \tag{4.11}$$

où N représente le nombre d'images par classe et $X_{i,j}$ désigne le j^{th} vecteur caractéristique de la classe i.

La densité de probabilité normalisée Pi(T) d'un vecteur caractéristique T conditionnellement à la classe i est calculée comme suit:

$$P_i(\mathrm{T}) = \frac{\mathrm{p}(\mathrm{T} \mid \mathrm{i})}{\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{p}(\mathrm{T} \mid \mathrm{i})} \tag{4.12}$$

$$\mathrm{p}(\mathrm{T} \mid \mathrm{i}) = \frac{\exp^{-\frac{1}{2}(\mathrm{T}-\mathrm{M_i})^{\mathrm{T}}.\mathrm{Ci}^{-1}.(\mathrm{T}-\mathrm{M_i})}}{2\pi|C_i|^{1/2}} \tag{4.13}$$

Habituellement, dans le classificateur MV, le vecteur T est attribué à la classe avec la probabilité la plus élevée. Cependant, étant donné que le nombre de classes est égal au nombre de sujets (avec des visages réels) à reconnaître, le classificateur décide qu'une image représente un vrai visage uniquement si la probabilité, calculée à l'aide des équations (4.12 et 4.13) et (4.20) est supérieure à un seuil qui assure des taux d'erreur égaux (erreur de rejeter un vrai visage et erreur d'accepter un faux visage). Ce seuil est déterminé en utilisant un ensemble de validation.

IV-Conclusion

Ce chapitre a été consacré à l'étude des contre-mesures aux attaques par masques 3D qui tentent de tromper les systèmes de reconnaissance de visage. La première partie, a été consacrée à un état de l'art sur les méthodes d'extraction de caractéristiques pour l'étape de vérification; les principales techniques de vérification utilisent l'analyse de texture par plusieurs variantes LBP, que nous avons présentées.

Nous avons ensuite présenté deux nouvelles contre-mesures pour faire face au spoofing par masque 3D. La première contremesure utilise une étape de vérification basée sur la décomposition polynomiale de l'image par ART pour l'extraction de caractéristiques et la classification par maximum de vraisemblance.

Pour réduire la complexité, nous avons conçu un nouveau système de reconnaissance qui permet de rejeter les attaques par masques 3D sans la nécessité d'une phase de vérification. Ce système utilise la décomposition polynomiale de l'image par LMI combinée avec la LDA pour l'extraction de caractéristiques et la classification par maximum de vraisemblance.

Chapitre IV :

Implémentation et Résultats

I-Introduction

Dans le chapitres III, nous avons décrit un système de reconnaissance de visage et on a proposé une nouvelle méthode pour l'extraction des caractéristiques qui utilise ART comme alternative aux LMI et PMZ. Dans le chapitre IV on a étudié la vulnérabilité des systèmes de reconnaissance du visage par Rapport aux spoofing par masque 3D et on a proposé deux contre-mesures pour faire face à ce problème.

Toutes ces méthodes ont été implémentées en langage C, en utilisant l'environnement Visual Studio avec l'environnement Matlab10.0a. Pour l'acquisition des images ou vidéos nous avons utilisé la bibliothèque OpenCV, pour les nombreux outils qu'elle offre. Pour la classification par Support Vecteur Machine(SVM), nous avons utilisé la bibliothèque LibSVM [82] qui peut fonctionner avec plusieurs codes sources comme C++, Python et Java.

Pour valider et comparer les différentes méthodes de reconnaissance par (ART, PZM, LMI) proposées dans le chapitre III nous avons utilisé quatre bases de données (ORL, Essex Faces94, Essex Faces96, Yaleface). Pour valider et comparer les deux contremesures par (ART-MV, LMI-LDA) proposées dans le chapitre IV, nous avons utilisé la base de données 3DMAD.

II Bases de données utilisées

Comme pour toute problématique de reconnaissance automatique, les techniques doivent être validées à l'aide d'un ensemble consistant de données. Les bases de données que nous avons utilisées pour valider les différentes méthodes de reconnaissance de visage et d'anti-spoofing par masques 3D sont présentées ci-après.

II-1 Base de données ORL

La base de données ORL (Figure 4.1) a été collectée entre avril 1992 et avril 1994 par un laboratoire AT &T, basé à Cambridge. Elle contient 40 personnes, chacune enregistrée sous 10 vues différentes (Figure 4.2). Cette base est considérée comme une référence pour l'évaluation des algorithmes de reconnaissance de visage.

Pour certaines personnes, les images ont été recueillies à différents moments, avec des variations dans les conditions d'éclairage et les expressions faciales (expression neutre, sourire et yeux fermés) et avec des occlusions partielles par des lunettes. Toutes les images de la base de données sont étiquetées, ce qui permet d'évaluer les performances des méthodes de reconnaissance faciale. La taille de chaque image de la base de données ORL est de 92 x 112

pixels, chaque pixel étant codé en niveaux de gris sur 8 bits..

Figure 4.1. Extrait de la base ORL redimensionnée à 64 x 48 avec les différentes orientations. Les images sont transformées en niveau de gris.

Figure 4.2. Image 1 des dix premiers individus de la base ORL.

II-2 Base de données Essex Face94

La base de données Essex Faces94 (Figure 4.3) a été créée, à l'Université de Cambridge, dans le cadre d'un travail pour la réalisation d'un système de reconnaissance faciale avec les PMZ. Cette base de données a été utilisée pour comparer notre approche, qui utilise l'ART, avec l'approche qui utilise le PMZ. Elle contient 72 classes, enregistrées sans variations d'orientation faciale, et avec 20 changements d'expressions faciales (Figure 4.4). La taille des images dans cette base de données est de 200 × 180 pixels.

Figure 4.3 Extrait de la base Face94 redimensionnée à 65 x 65 montrant les différentes expressions faciales. Les images sont transformées en niveau de gris.

Figure 4.4 Image 1 des 10 premiers individus de la base Face94.

II-3 Base de données Yale face

La base de données de visage de Yale (Figure 4.5) comprend 165 images de 15 personnes, chacune enregistrée sous 11 variantes différentes de l'état de l'éclairage, de l'expression du visage et avec ou sans lunettes (Figure 4.6). La taille des images dans cette base de données est de 92 × 196 pixels.

Figure 4.5 Image n ° = 1 des onze premiers individus dans la base de données Yale.

Figure 4.6 Extraits de la base de données Yale montrant des variations dans l'état de l'éclairage, l'expression du visage et avec ou sans lunettes.

II-4 Base de données Essex Faces96

La base de données Essex Faces96 (Figure 4.7), développée à l'Université de Cambridge, est plus grande par rapport aux bases de données précédentes. Elle contient 152 classes enregistrées sans grande variation d'orientation faciale et avec 20 changements d'expressions faciales (Figure 4.8). La taille des images dans cette base de données est de 196 × 196 pixels.

Figure 4.7 Première image de dix personnes de la base de données Faces96.

Figure 4.8 Extraits de la base de données Faces96 montrant différentes expressions faciales.

II-5 Base de données 3DMAD

La 3DMAD (3D Masque Attack) est principalement composée de 17 différents sujets à visage réel et avec masques 3D d'attaque comme montre la figure 4.9. Elle a été enregistrée par le capteur Kinect de Microsoft sous forme de séquences vidéo. Pour chaque sujet on a 15 séquences vidéo. Chaque séquence vidéo est composée de 300 frames, chacune contenant une image d'un seul visage (réel ou masque), d'une taille égale à 640*480 pixel. Les dix premières séquences pour chaque sujet contiennent les enregistrements avec visage réel et les cinq dernières séquences contiennent les enregistrements avec un masque du même sujet

comme le montre la figure 4.10.

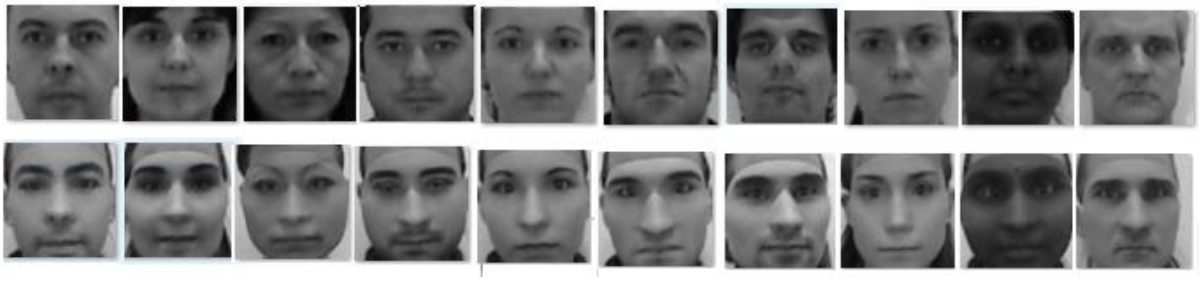


Figure 4.9 Extrait des images de dix personnes de la base 3DMAD à visage réel et avec masque 3D.

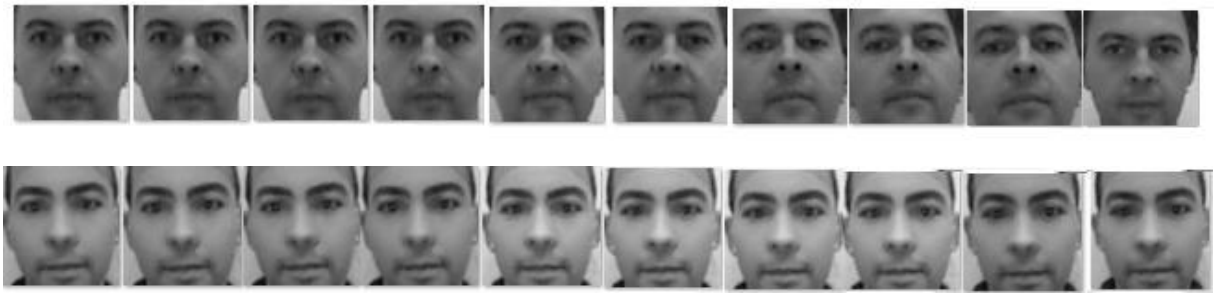


Figure 4.10 Extrait des images d'une personne de la base 3DMAD à visage réel et avec masque 3D.

III- Prétraitement

Avant l'extraction des vecteurs caractéristiques, un prétraitement des images a été effectué. La première étape consiste à convertir chaque image couleur (RGB) dans les bases de données (Faces94, Faces96) en une image en niveaux de gris. Ensuite, les images de la base de données Yale et de la base de données Faces96 ont été recadrées afin de garder la région du visage uniquement. Les images dans la base de données Face94 ont été redimensionnées à 65 x 65, toutes les images dans les bases de données (ORL, Faces96, Yale) ont été redimensionnées à 64 x 48 pixels par échantillonnage [84]. Les étapes de prétraitement concernant la base de données 3DMAD sont montrées dans les étapes suivantes.

-la première étape consiste à convertir l'image couleur RGB (Figure 4.11) de chaque frame en une image en niveau de gris (Figure 4.12), pour calculer le niveau de gris d'un pixel on calcul la moyenne des intensités rouge, vert et bleu de chaque pixel.

Figure 4.11 Image couleur RGB [29]

Figure 4.12 Image convertie en niveau de gris

-La deuxième étape est la détection automatique de la zone du visage dans l'image avec l'algorithme Haar [81], comme indiqué à la Figure 4.13, cet algorithme est déjà réalisé dans notre de magister [83].

Figure 4.13 Détection de visage avec l'algorithme Haar.

-La troisième étape consiste à découper et à redimensionner la zone du visage détectée dans une image de taille 64 x 64, comme le montre la Figure 4.14. Cette résolution a été déterminée expérimentalement. Elle correspond à la plus petite résolution qui donnes des résultats satisfaisant. À la fin de cette étape, nous avons formé notre base de données composée de 3000 images de visages réels et 1500 images de visages avec masques, pour chacun des 17 sujets.

Figure 4.14 Image de visage redimensionnée à 64 x 64.

IV- Résultats et discussions

La présentation des résultats est organisée comme suit. Dans un premier temps, nous allons présenter les résultats de reconnaissance du visage obtenus en utilisant les différents moments pour l'extraction des caractéristiques et le PPV ou les SVM pour la classification. En utilisant la base de données 3DMAD, nous montrerons ensuite que les systèmes de reconnaissance du visage sont généralement vulnérables aux attaques par masques 3D. En fin, nous présenterons les résultats des contre-mesures que nous proposons pour parer à ces attaques et nous les comparerons à ceux des contre-mesures qui utilisent les LBP pour l'extraction des caractéristiques.

IV-1- Reconnaissance du visage

Pour la reconnaissance du visage, nous nous sommes intéressés aux méthodes qui utilisent les moments pour l'extraction des caractéristiques. Avant de comparer ces méthodes entre elles, nous allons examiner l'influence de l'ordre de décomposition sur leurs performances.

IV-1-1 Influence de l'ordre de décomposition sur la performance de chaque méthode

Par les expériences on a trouvé que l'ordre de décomposition « n » affecte le temps de réponse et le taux de reconnaissance des systèmes de reconnaissance faciale, qui sont basés sur les moments polynomiaux. L'augmentation de l'ordre de décomposition augmente le taux de reconnaissance, ainsi que le temps de réponse. Cependant, comme le montrent les résultats présentés dans les Figures 4.15-4.22, l'augmentation de l'ordre de décomposition au-delà d'une certaine valeur, qui dépend de la méthode utilisée, n'améliore pas le taux de reconnaissance. Les Figures 4.15-4.22 illustrent la variation du taux de reconnaissance en fonction de l'ordre de décomposition n, pour les différents moments, avec classification par SVM ou PPV. Les figures 4.15 et 4.16 ont été obtenues, en utilisant la base de données Faces94, avec 20 images par classe (15 images pour l'apprentissage et 5 pour le test), tandis que les figures 4.17 et 4.18 ont été obtenues, en utilisant la base de données ORL, avec 10 images par classe (7 images Pour l'apprentissage et 3 pour le test). Les figures 4.19 et 4.20 ont été obtenues, en utilisant la base de données Yale, avec 11 images par classe (7 images pour l'apprentissage et 4 pour le test).

Les figures 4.21 et 4.22 ont été obtenues, en utilisant une base de données plus grande, c'est-à-dire la base de données Faces96, avec 20 images par classe (15 images pour l'apprentissage et 5 pour le test).

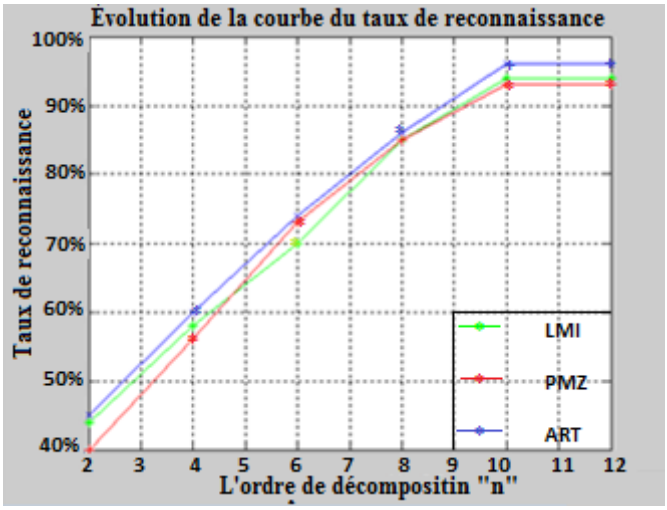


Figure 4.15 Taux de reconnaissance, avec classification par SVM, pour la base de données Faces94.

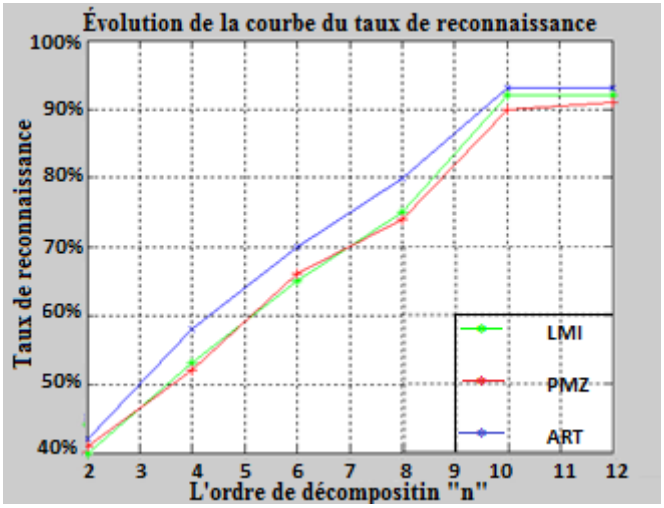


Figure 4.16 Taux de reconnaissance, avec classification par PPV, pour la base de données Faces94.

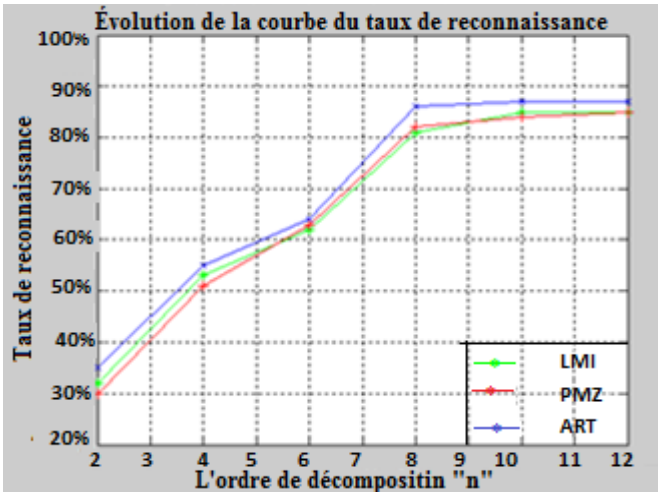


Figure 4.17 Taux de reconnaissance, avec classification par SVM, pour la base de données ORL.

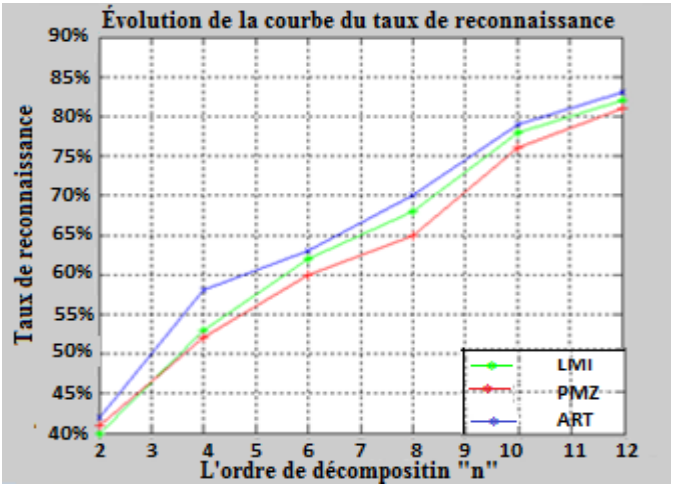


Figure 4.18 Taux de reconnaissance, avec classification par PPV, pour la base de données ORL.

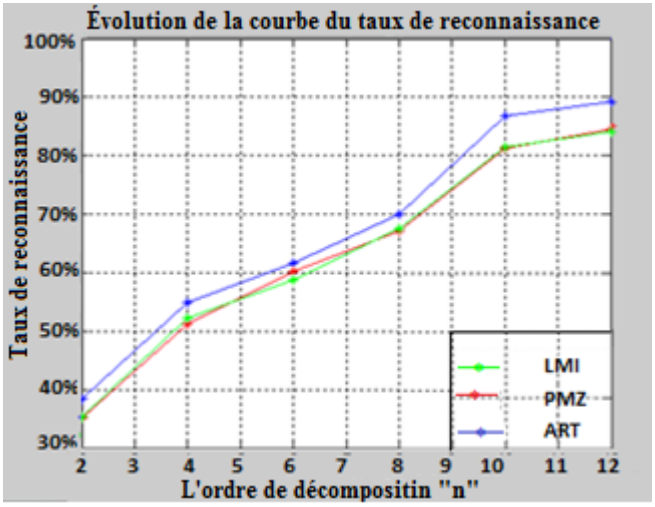


Figure 4.19. Taux de reconnaissance, avec classification par SVM, pour la base de données Yale.

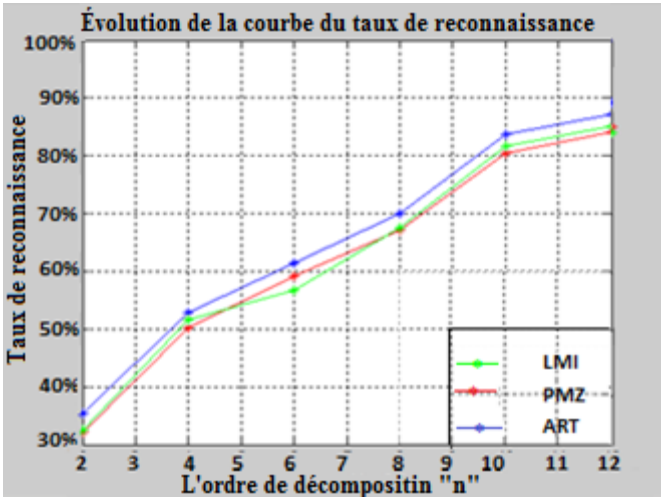


Figure 4.20. Taux de reconnaissance, avec classification par PPV, pour la base de données Yale.

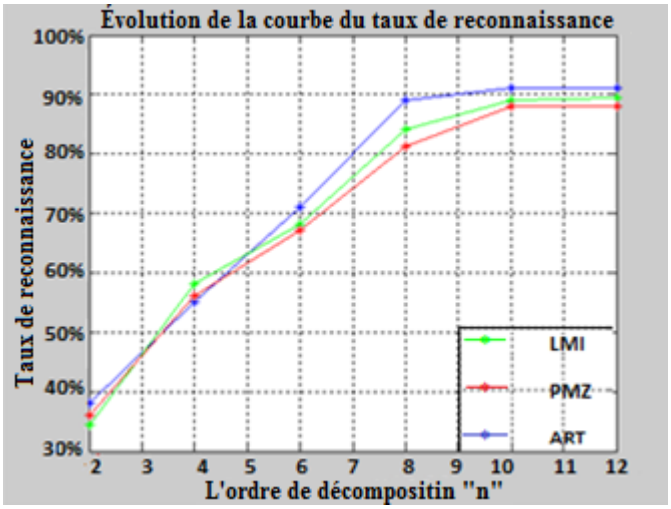


Figure 4.21. Taux de reconnaissance, avec classification par SVM, pour la base de données Faces96.

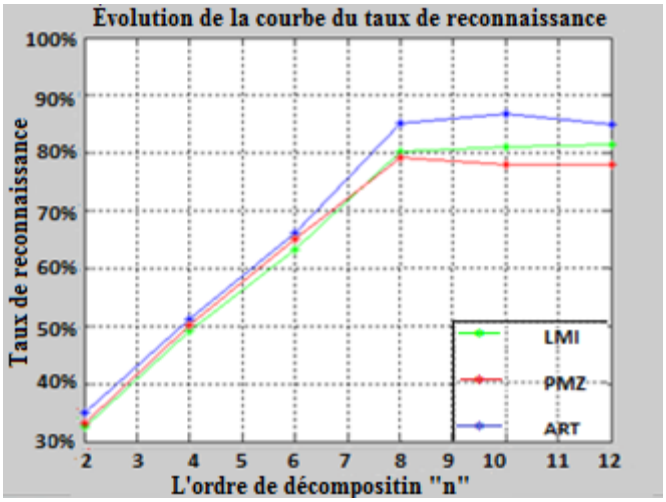


Figure 4.22. Taux de reconnaissance, avec classification par PPV, pour la base de données Faces96.

IV-1-2 Comparaison des différentes méthodes de reconnaissance de visage basées sur les moments

Les meilleurs taux de reconnaissance obtenus par les différentes méthodes de reconnaissance faciale, basées sur les moments, ainsi que les ordres de décomposition avec lesquels ils ont été obtenus sont donnés dans le tableau 4.1. On peut observer à partir de ce tableau que les trois méthodes se classent dans cet ordre : l'ART, suivie pars les LMI, suivis par les PMZ. On peut également constaté que la classification par SVM est meilleure que la classification par PPV.

	ORL (n)	**Faces94 (n)**	**Yale (n)**	**Faces96 (n)**
ART+PPV	83.1% (12)	94.1% (10)	88.0% (12)	87.4% (10)
ART+SVM	88.0% (10)	96.0% (10)	89.4% (12)	90.8% (12)
LMI+ PPV	82.1% (12)	92.2% (10)	85.1% (12)	82.2% (12)
LMI+SVM	84.4% (10)	93.1% (10)	86.3% (12)	89.1% (12)
PMZ+PPV	81.3% (12)	91.3% (12)	84.6% (12)	79.3% (8)
PMZ+SVM	83.1% (10)	92.2% (10)	85.9% (12)	88.2% (12)

Tableau 4.1. Résultats des meilleurs taux de reconnaissance en fonction de l'ordre de répétition 'n' obtenus avec les différentes méthodes basées sur les moments.

Le temps de réponse, ainsi que la taille mémoire et le taux d'identification obtenus dépendent principalement de l'ordre de décomposition et de la méthode de classification utilisée. On a constaté que l'ordre de décomposition optimal est habituellement plus faible lorsque le classificateur SVM est utilisé, en particulier dans le cas de la base de données ORL. Cependant, ce classificateur est plus coûteux en termes de temps de calcul que le PPV, par exemple pour la classification d'une base de données composé de 170 vecteurs d'une taille de 40 éléments, le SVM a nécessité 0.708512 secondes contre 0.000001 secondes pour le PPV tel que donné par le compilateur (expériences réalisées sur un PC hp, CPU i3-3120M CPU @ 2.50 GHz X 2, RAM 4.00 Go).

IV-2- Résultats des contre-mesures aux attaques par masques 3D

Pour l'évaluation des performances des contre-mesures aux attaques par masques 3D, deux expériences ont été réalisées. Le but de la première est de prouver que les systèmes de reconnaissance faciale peuvent en général facilement être piratés. Pour cela on a calculé le SFAR de certains systèmes de reconnaissance, en utilisant la base donnés 3DMAD. La deuxième expérience a été menée pour montrer que la méthode de vérification ART-MV et le système de reconnaissance LMI-LDA-MV proposés sont robustes contre les attaques par masques 3D.

Dans les deux expériences, trois ensembles ont été formés: un ensemble d'apprentissage, un ensemble de test et un ensemble de validation. Pour la formation de ces ensembles 10 images (d'un visage réel ou avec masque) ont été prises pour chaque sujet.

- Pour les deux expériences, l'ensemble d'apprentissage est composé de 12 sujets sur les 17 sujets de la base de données. Pour chaque sujet, nous avons sélectionné aléatoirement 10 images de son visage réel. Cet ensemble est utilisé pour calculer les vecteurs caractéristiques de référence.
- L'ensemble de validation est utilisé pour calculer le EER (Equal Error Rate) pour établir le seuil de décision. Il est composé de deux sous-ensembles. Pour la première expérience, le premier sous-ensemble est composé de différentes images des visages réels des 12 sujets utilisés dans la phase d'apprentissage, pour calculer le taux de faux rejets (FRR), ç-à-dire le taux de vrais visages rejetés. Le deuxième sous-ensemble est composé des images des visages réels des cinq sujets non utilisés dans la phase d'apprentissage, pour calculer le taux de fausses acceptations (FAR), ç-à-dire le taux de visages réels acceptés qui n'appartiennent pas à l'ensemble d'apprentissage (imposteurs).

Pour la deuxième expérience, le premier sous-ensemble se compose de la même manière que dans la première expérience, pour calculer le FRR, tandis que le deuxième sous-ensemble est composé de masques 3D des 12 sujets utilisés dans la phase d'apprentissage, pour calculer le taux d'acceptation des falsifications (SFAR).

-L'ensemble de test est composé de deux sous-ensembles, pour les deux expériences. Le premier sous-ensemble est composé d'images des visages réels des 12 sujets utilisés dans la phase d'apprentissage et des 5 sujets non utilisés dans la phase d'apprentissage pour calculer le FRR et le FAR. Le deuxième sous-ensemble est composé d'images faciales avec des masques 3D des 12 sujets utilisés dans la phase d'apprentissage pour calculer SFAR.

IV-2-1 Test de vulnérabilité des systèmes de RF aux attaques par masques 3D

Nous allons maintenant présenter les résultats qui montrent que les systèmes de RF sont en général vulnérables aux attaques par masques 3D. Nous avons pour cela choisi les systèmes (LMI-PPV, LMI-SVM, ART-PPV, ART-LMI). Pour les systèmes utilisant la classification par PPV, le seuil de décision optimal doit être déterminé. Les Figure 4.23 et 4.24 montrent comment le EER (entre FRR et FAR) est utilisé pour déterminer expérimentalement le seuil de décision optimal, dans les systèmes LMI-PPV et ART-PPV, à l'aide de la base de données 3DMAD.

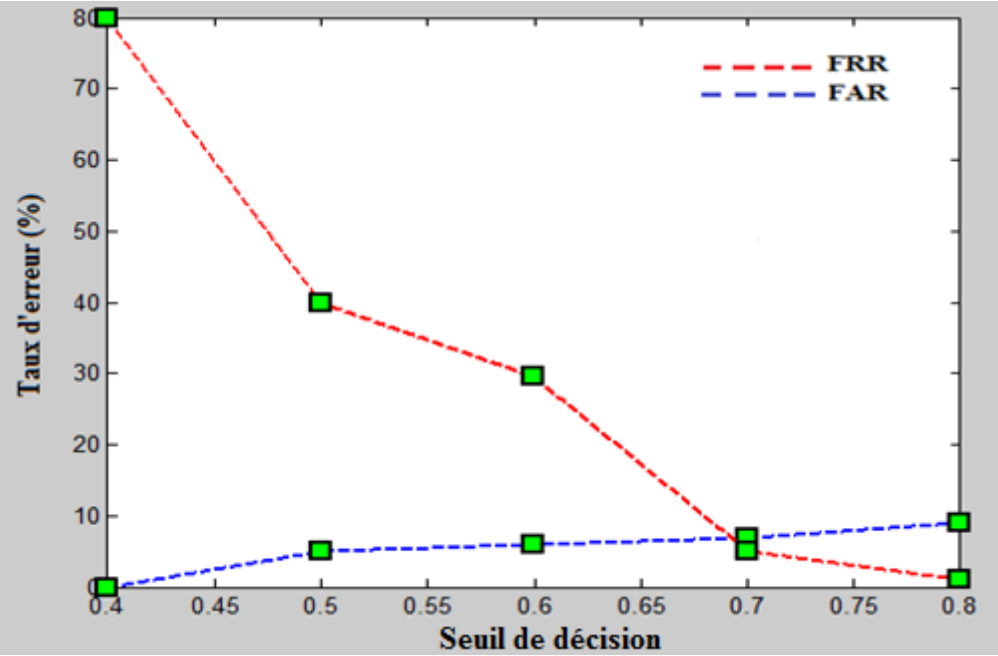


Figure 4.23. EER pour fixer le seuil de décision dans la méthode LMI-PPV.

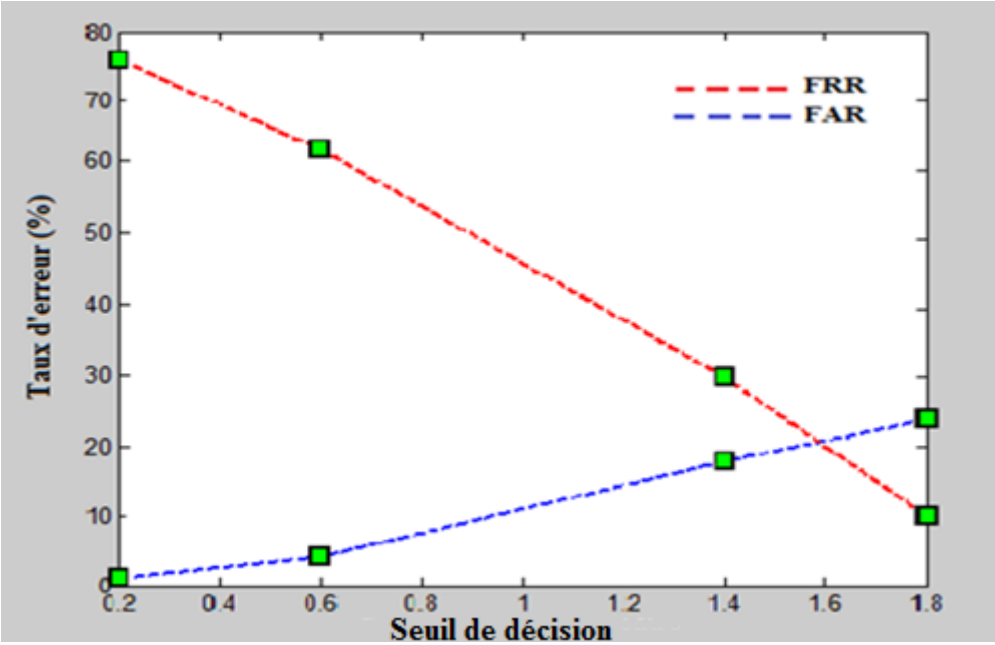


Figure 4.24. EER pour fixer le seuil de décision dans la méthode ART-PPV.

Pour tester la vulnérabilité aux attaques par masques 3D des différents systèmes de RF choisis, nous estimons leurs SFARs en utilisant la base de données 3DMAD. Mais, avant cela nous montrons dans le tableau 4.2 que, comme avec les autres bases de données, ces systèmes permettent d'obtenir des taux de reconnaissance très élevés avec la base de données 3DMAD. Ils peuvent donc être considérés comme de très bons systèmes de RF. Cependant, comme le montre le tableau 4.3 ces systèmes, à l'instar d'autres systèmes, comme la LDA+PPV, la LDA+SVM, et l'ISV (Inter Session Variability) [26], ont des SFARs supérieurs à 57%. Ils sont donc très vulnérables aux attaques par masques 3D, ce qui les rend impraticables dans des applications sensibles, où la perméabilité aux attaques doit être quasiment nulle. Pour de telles applications des contre-mesures à ces attaques doivent être trouvées.

Méthode	Taux de reconnaissance
LMI+SVM	97.20%
LMI+PPV	97.40%
ART+SVM	97.91%
ART+PPV	97.68%

Tableau 4.2. Taux de reconnaissance obtenus par différentes méthodes en utilisant la base de données 3DMAD.

Méthode	SFAR
LDA+PPV	57.24%
LDA+SVM	64.21%
LMI+PPV	60.0%
LMI+SVM	66.80%
ART+PPV	57.24%
ART+SVM	67.2%
ISV [26]	65.7%

Tableau 4.3. Taux SFAR de certains systèmes de reconnaissance faciale.

IV-2-2 Résultats des contre-mesures aux attaques par masques 3D

Dans ce qui suit, nous allons présenter les résultats des deux approches que nous proposons pour contrecarrer les tentatives de spoofing des systèmes de RF par masques 3D. La première approche repose sur une méthode de vérification, tandis que la seconde est un système de RF intrinsèquement immunisé contre ces attaques.

IV-2-2-1 Contre-mesure par méthode de vérification

Les performances de la méthode de vérification ART-MV proposée sont mesurées à l'aide du HTER (Half Total Equal Rate), qui est la moyenne des taux d'erreurs FRR et SFAR, calculés sur l'ensemble de test, pour le seuil EER obtenu avec l'ensemble de validation.

Le tableau 4.4 ci-dessous donne le HTER obtenu avec cette méthode et le compare avec celui obtenu par la méthode LBP-LDA [26] en utilisant la même base de données.

Méthode	HTER
ART-MV	0.91%
LBP-LDA	0.95%

Tableau 4.4. Comparaison entre les HTER des méthodes ART-MV et LBA-LDA[26].

D'après ce tableau, on peut dire qu'en comparaison avec la méthode LBP-LDA, la méthode proposée ART-MV donne des résultats légèrement meilleurs avec la base de données 3DMAD.

IV-2-2-2 Système de RF intrinsèquement immunisé contre les attaques par masques 3D

A la Figure 4.25, le FRR et le SFAR obtenus par la méthode LMI-LDA-MV sont tracés en fonction du seuil de décision. On peut déduire, à partir de cette figure, que le seuil correspondant à l'EER est égal à 0,4.

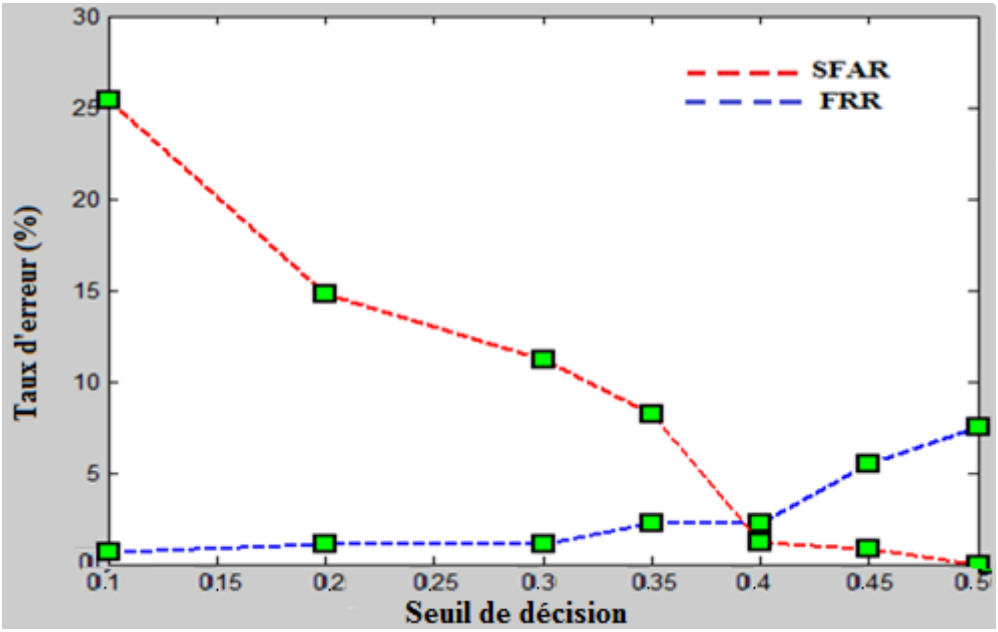


Figure 4.25. EER pour fixer le seuil pour la décision dans la méthode LMI-LDA-MV.

Par ailleurs, la Figure 4.26 affiche la caractéristique de fonctionnement du récepteur (ROC) correspondante, qui représente la variation du taux de reconnaissance par rapport au SFAR lorsque le seuil de décision diminue.

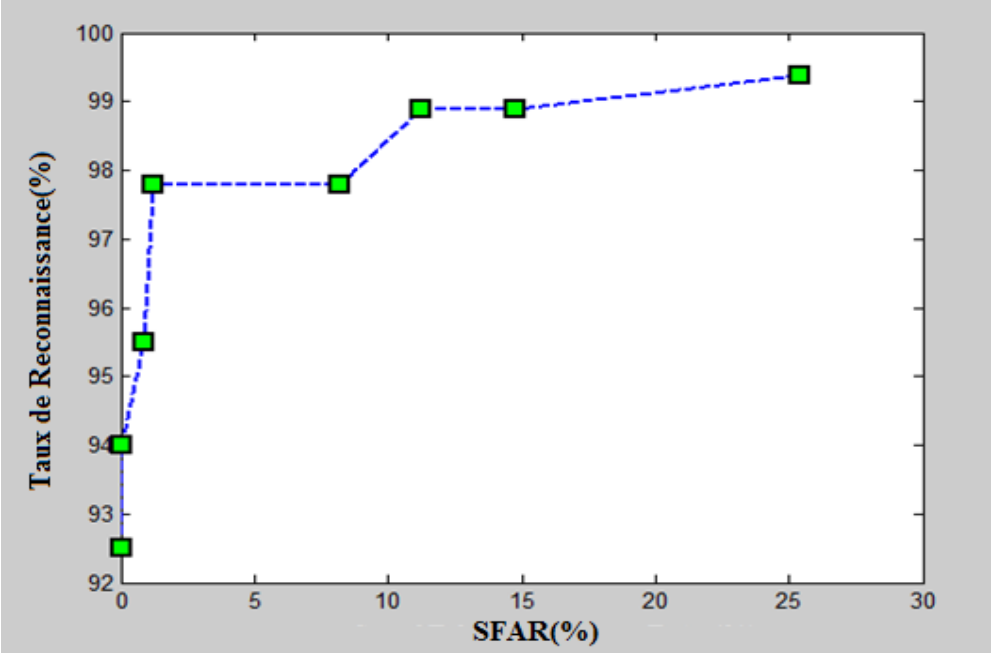


Figure 4.26. Courbe ROC de la méthode LMI-LDA-MV.

Le tableau 4.5 énumère les taux de reconnaissance et les FAR, obtenus par la méthode LMI-LDA-MV, à l'aide de la base de données 3DMAD et des bases de données ORL, Faces94, Yale et Faces96. Pour toutes les bases de données, sauf la Yale, ces taux de reconnaissance sont meilleurs que ceux obtenus avec la méthode LMI-PPV et présentés dans les tableaux 4.1 et 4.2.

Base de données	Taux de reconnaissance	FAR
3DMAD	97.6%	0.0%
ORL	90.6%	3.4%
Faces94	94.4%	3.1%
Yale	84.6%	4.5%
Faces96	88.2%	3.2%

Tableau 4.5. Taux de reconnaissance et FARs obtenus par la méthode LMI-LDA-MV, avec différentes bases de données.

Il est à noter que les FAR obtenus avec les différentes bases de données sont très faibles, celui obtenu avec la BD 3DMAD est carrément nul. La Figure 4.27 présente la probabilité moyenne qu'un visage réel de test appartienne à sa classe, ainsi que l'incertitude qui y attachée. Les barres représentent les probabilités, et les lignes verticales les écarts-types.

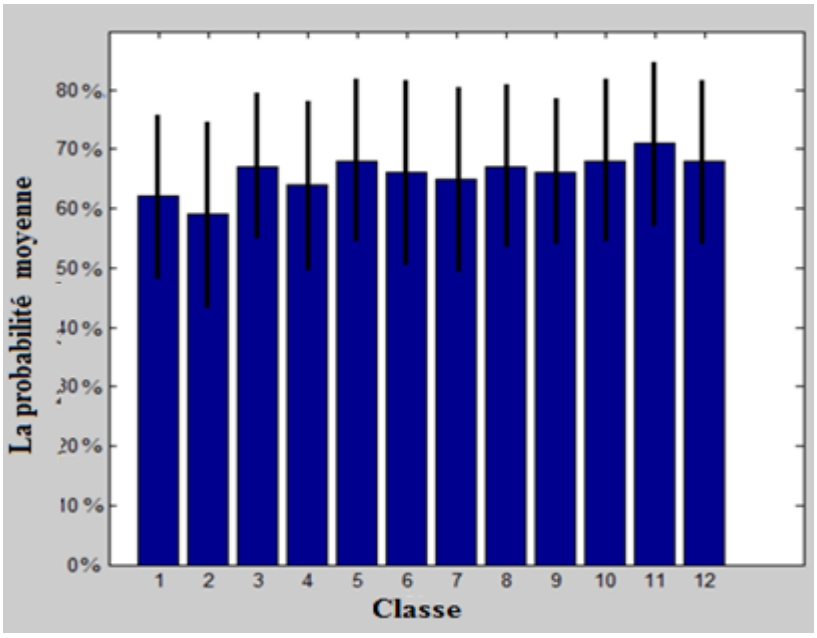


Figure 4.27. Probabilité moyenne qu'un visage réel appartienne à sa classe.

A titre de comparaison, la Figure 4.28 montre la probabilité moyenne, avec l'incertitude correspondante, qu'un masque d'un visage appartienne à la classe de ce visage.

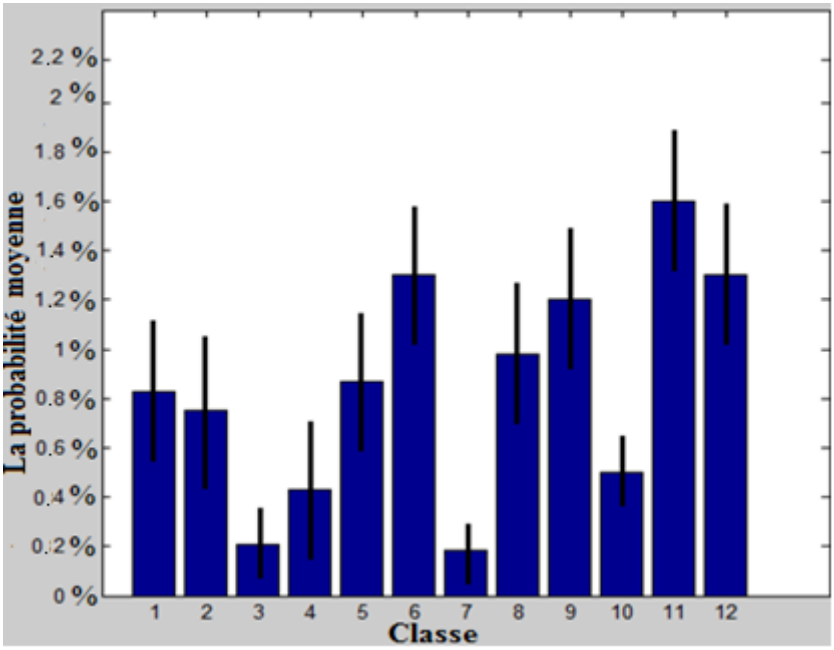


Figure 4.28. Probabilités moyennes qu'un masque d'un visage appartient à la classe de ce visage.

A partir de ces deux figures, on constate que la densité de probabilité moyenne qu'un visage réel appartienne à sa propre classe varie autour de 0.6 et que cette probabilité est inférieure à 0.016 pour un masque. Sur la base de ces résultats, on peut affirmer que la LMI-LDA-MV est très capable de discriminer entre un visage réel et un masque 3D.

Le tableau 4.6 ci-dessous compare le SFAR obtenu par la méthode LMI-LDA-MV proposée avec ceux obtenus par les méthodes LMI-PPV, LMI-SVM, ART-PPV, ART-SVM et ISV [26]. On peut observer à partir de ce tableau que contrairement aux cinq dernières méthodes, la méthode proposée permet de détecter presque toutes les attaques par masques 3D.

Méthode	SFAR
LMI-LDA-MV	0.87%
LMI+PPV	60.0%
LMI+SVM	66.80%
ART+PPV	57.24%
ART+SVM	67.2%
ISV	65.7%

Tableau 4.6. SFARs obtenus par différentes méthodes.

Pour évaluer d'avantages la méthode LMI-LDA-MV proposée, nous comparons sa performance avec celles de deux méthodes qui utilisent une étape de vérification: la méthode ART-MV que nous avons proposée et la méthode proposée dans [26], basée sur l'extraction des caractéristiques par LBP et la classification par LDA. Le tableau 7 présente le SFAR obtenu avec la première méthode et les (HTER) des deux dernières.

Méthode	Taux
LMI-LDA-MV	0.87%(SFAR)
ART-MV	0.91%(HTER)
LBP-LDA	0.95%(HTER)

Tableau 5.7. Performances des méthodes proposées et de la méthode proposée dans [26].

On peut observer que les performances des trois méthodes sont comparables et qu'elles sont toutes efficaces pour protéger un système de reconnaissance contre les attaques de masque 3D. L'avantage de la première méthode « LMI-LDA-ML » est qu'elle n'utilise pas une étape de vérification; elle est donc plus rapide.

V- Conclusion

Dans ce chapitre nous avons présenté les résultats d'évaluation des différents systèmes de reconnaissance de visage réalisés. Outre la comparaison des taux de reconnaissance de ces systèmes, nous avons évalué l'influence de certains facteurs sur les performances de ces systèmes. Ensuite on a présenté les résultats d'investigation concernant le spoofing des systèmes de RF par masque 3D pour montrer l'ampleur du danger auquel sont exposés ces systèmes et la nécessité de recourir à des contre-mesures pour les protéger. Enfin nous avons présenté les résultats des contremesures proposées pour cela.

A partir des résultats obtenus, on peut dire qu'un système de reconnaissance du visage par ART-SVM avec un système de vérification par ART-MV ou un système de reconnaissance simplement par LMI-LDA-MV nous permettent d'avoir un bon système de reconnaissance du visage sécurisé contre les attaques par masques 3D avec un taux de spoofing très faible qui vaut 0.87% et un taux de reconnaissance égal à 97.6%.

Conclusion générale

Conclusion générale

Les travaux effectués dans le cadre de cette thèse portent sur la réalisation d'un système biométrique basé la reconnaissance du visage qui soit sécurisé contre les attaques par masques 3D.

Les systèmes reconnaissance de visage ont fait l'objet d'études approfondies ces dernières années. Ces études avec le développement des technologies d'acquisition 2D, qui sont devenues de plus en plus précises et moins coûteuses, ont permis d'améliorer les performances de ces systèmes et les ont rendus attractifs pour plusieurs applications. Cependant, l'apparition d'attaques (spoofing en anglais), notamment celles menées à l'aide de masques 3D, qui tentent de tromper ces systèmes constitue un frein à cette généralisation.

L'anti-spoofing est un nouveau créneau de recherche dans le domaine de la biométrie faciale et les études dans ce domaine sont limitées, notamment celles relatives à l'anti-spoofing par masques 3D, qui ont été freinées à cause de l'indisponibilité de bases de données de masques 3D publiques.

Le travail présenté dans cette thèse comporte deux parties. Dans la première partie, on a proposé d'utiliser la Transformation Angulaire Radiale (ART), qui est une décomposition polynomiale par moments, pour extraire les caractéristiques du visage qui sont ensuite fournies à un classifieur à base de SVM ou voisin le plus proche pour la reconnaissance de visage. Les images faciales des bases de données ORL, Essex Faces94, Essex Faces96 et Yale ont été utilisées pour tester l'approche proposée. Les résultats expérimentaux obtenus montrent que la méthode proposée est plus efficace, en termes de taux de reconnaissance, que les méthodes basées sur les moments de Zernike et Legendre. Il a été également constaté que sa performance est comparable à celles des meilleures méthodes de l'état de l'art.

Ayant constaté la vulnérabilité de ce système de RF, à l'instar d'autres systèmes, aux attaques par masques 3D, nous avons proposé dans la seconde partie deux systèmes de RF sécurisés contre ces attaques. Le premier système utilise deux étapes : une étape de reconnaissance suivie d'une étape de vérification. Pour celle-ci, nous avons utilisé la méthode ART pour l'extraction de caractéristiques et le MV pour la classification. Le second système RF proposé peut être qualifié comme étant intrinsèquement immunisé contre les attaques par masques 3D. En effet, sans recourir à une phase de vérification, il permet de rejeter pratiquement toutes ces attaques. Il est de ce fait plus simple et donc plus rapide. Les performances des deux systèmes ont été évaluées à l'aide de la base de données 3DMAD. Elles peuvent être qualifiées comme étant très satisfaisantes et se comparent favorablement à celle du seul système proposé dans la littérature qui utilise la même base de données et qui utilise une phase de reconnaissance et une phase de vérification, basée sur les LBP pour l'extraction des caractéristiques.

En perspective, il est envisagé de réaliser un système de reconnaissance 3D qui puisse détecter les personnes portant un masque 3D et les signaler.

Annexes

Les SVM (Support Vector Macines) sont utilisées pour résoudre des problèmes de discrimination, c'est-à-dire décider à quelle classe appartient un échantillon « x » et lui affecter une valeur numérique en fonction de sa classe, par exemple +1 pour la classe 2 et -1 pour la classe 1, voir la figure ci-dessous.

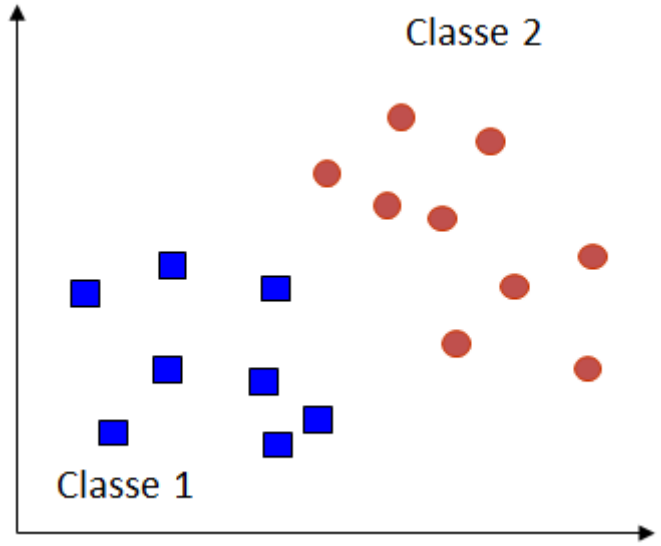


Figure. A.1. Problème de discrimination entre la classe 1 et la classe 2.

La première idée pour classifier des données linéairement séparable est de trouver une frontière séparatrice optimale à partir d'un ensemble d'apprentissage. La résolution de ce problème de discrimination par SVM passe par la construction d'une fonction h de type hyperplan qui a un vecteur d'entrée « x » fait correspondre une sortie *y*:

$$y = h(x), \tag{A.1}$$

avec:

$$h(x) = w^T . x + w_0 , \tag{A.2}$$

où $x = (X_1,\ldots,X_N)^T$, $w = (W_1,\ldots,W_N)^T$ est un vecteur poids et w_0 est une constante.

Pour un problème de discrimination à deux classes (discrimination binaire), le but d'un algorithme d'apprentissage supervisé est d'apprendre la fonction h(x) par le biais d'un ensemble d'apprentissage :

$$(\{(x_0, y_0), (x_1, y_1), \ldots\ldots\ldots, (x_k, y_k)\},\ x \subset \Re^N,\ y_i \in \{-1,1\}) \tag{A.3}$$

Si le problème est linéairement séparable, $h(x_i)$ doit vérifier:

$$y_i . h(x_i) \geq 0 \tag{A.4}$$

Il est alors décidé que « x » appartient à la classe 1 si h (x) est positif et « x » appartient à la classe -1 si h (x_i) est négatif.

Les SVM sont un classifieur linéaire qui peut s'appliquer pour des vecteurs d'entrée de grande dimension dans un espace X.

Le choix de l'hyperplan séparateur n'est pas facile même dans le cas où les données sont linéairement séparable, Il existe une infinité d'hyperplans séparateurs qui peuvent classifier correctement ces données (voir Figure A.1).

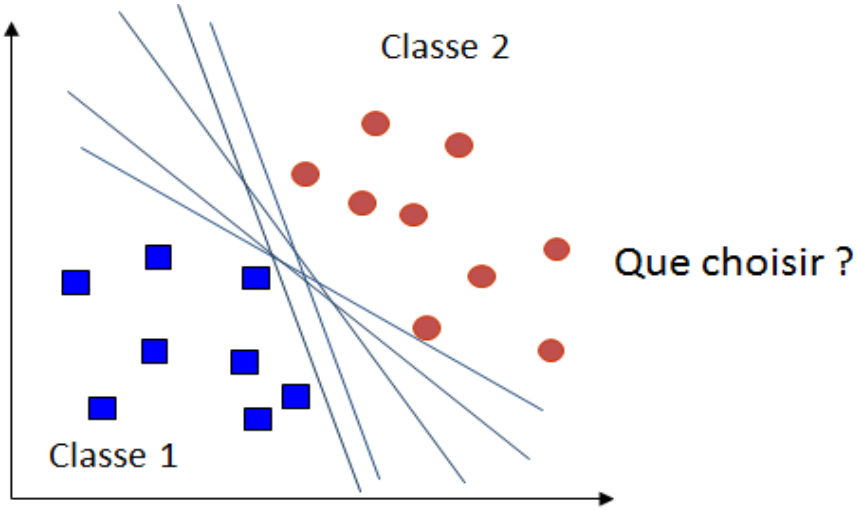


Figure. A.2. Infinité d'hyperplans séparateurs pour un ensemble de points linéairement séparables.

Dans [71], il a été proposé de choisir comme hyperplan séparateur optimal celui qui maximise ses marges par rapport aux échantillons des deux classes, d'où l'appellation Séparateurs à Vaste Marge. En utilisant la théorie statistique de l'apprentissage de Vapnik-Chervonenkis, ce choix a été justifié dans [72], où il a été montré que le taux de classification exacte par hyperplans séparateurs augmente lorsque la marge maximale augmente.

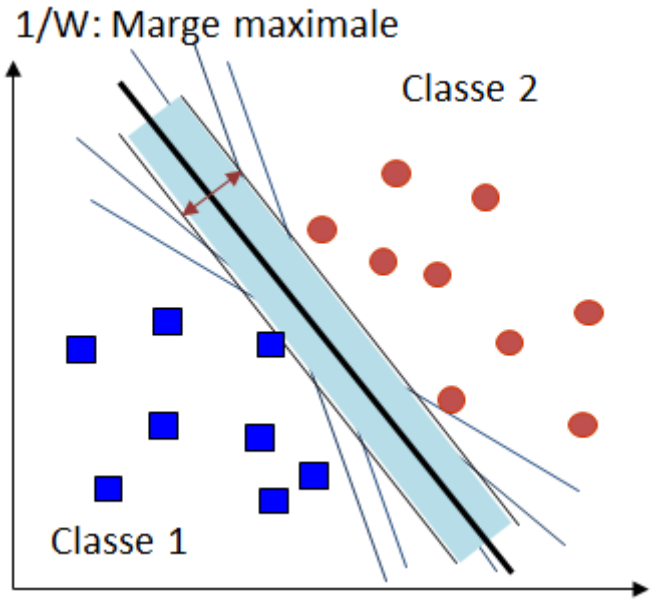


Figure. A.3. Hyperplan séparateur optimal avec une marge maximale.

La notion de marge maximale et la procédure de recherche de l'hyperplan séparateur ne permettent de résoudre que des problèmes de discrimination linéairement séparables. La marge est la plus petite distance entre les échantillons d'apprentissage et l'hyperplan séparateur qui satisfait la condition de séparabilité (y_k .(w^T. x_k+w_0) ≥0). Ces derniers sont appelés vecteurs supports.

La distance d'un échantillon x_k à l'hyperplan est donnée par sa projection orthogonale sur l'hyperplan :

$$d(x_k) = \frac{|y_k(w^T.x_k+w_0)|}{\|W\|} \tag{A.5}$$

L'hyperplan séparateur (w, w_0) de marge maximale est donné par l'équation suivante:

$$\arg\max_{W,W_0}\{\frac{1}{\|W\|}\min_k[\, y_k(w^T.x + w_0)]\} \tag{A.6}$$

Pour que l'optimisation soit facile il faut normaliser les deux coefficients w et w_0, pour que les échantillons x^+_{marge} pour les vecteurs supports sur la frontière positive et les échantillons x^-_{marge} pour les vecteurs supports sur la frontière négative satisfassent :

$$\begin{cases} w^T.x^+_{marge} + w_0 = 1 \\ w^T.x^-_{marge} + w_0 = -1 \end{cases} \tag{A.7}$$

On a alors pour tous les échantillons:

$$y_k(w^T.\ x_k + w_0) \geq 1 \qquad \text{(A.8)}$$

L'hyperplan obtenu avec cette normalisation s'appelle l'hyperplan canonique. Avec cette mise à l'échelle, la marge est égale à $\frac{1}{\|W\|}$ et la formulation dite primale des SVM revient à maximiser $\|W\|^{-1}$ sous contraintes. Ce qui peut être exprimé mathématiquement par :

$$\text{Minimiser } \frac{1}{2}\|w\|^2 \text{ sous les contraintes } y_k(w^T.x_k + w_0) \geq 1 \qquad \text{(A.9)}$$

Ce problème peut être résolu par la méthode classique des multiplicateurs de Lagrange. Le Lagrangien est donné dans ce cas par:

$$L(w, w_0, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{k=1}^{p} \alpha_k.(y_k(w^T.x_k + w_0) - 1) \qquad \text{(A.10)}$$

Les conditions dites de Karush-Cuhen-Tucker (KKT) pour la minimsation de (A.9) soit possible sont:

$$\begin{cases} \frac{d(L(w,w_0,\alpha))}{d(w)} = w_{opt} - \sum_{k=}^{p} \alpha_k.y_k.x_k = 0, \text{ d'où : } w_{opt} = \sum_{k=}^{p} \alpha_k.y_k.x_k \\ \frac{d(L(w,w_0,\alpha))}{d(w_0)} = \sum_{k=1}^{p} \alpha_k.\ x_k = 0 \\ \frac{d(L(w,w_0,\alpha))}{d(\alpha_k)} > 0,\ avec\ \alpha_k \geq 0 \end{cases} \qquad \text{(A.11)}$$

Le lagrangien est minimisé par rapport à w et w_0 et maximisé par rapport à α.

Pour passer du problème primal au problème dual il faut introduire les multiplicateurs de Lagrange pour chaque contrainte (exemple d'apprentissage) :

$$\begin{cases} \text{Maximiser } L(\alpha) = \sum_{k=1}^{p} \alpha_k - \frac{1}{2}\sum_{i,j}^{p} \alpha_i.\alpha_j.y_i.y_j.x_i^T.x_j\ , \\ \text{sous les contraintes : } \alpha_k \geq 0 \text{ , et } \sum_{k=1}^{p} \alpha_k.\ x_k = 0 \end{cases} \qquad \text{(A.12)}$$

La résolution de ce problème de programmation quadratique de dimension p (nombre d'échantillons) donne les multiplicateurs de Lagrange optimaux α_k^*; ces multiplicateurs sont non-nuls et s'appellent vecteurs de support. Pour obtenir la fonction de décision il faut simplement remplacer w par la valeur optimale w_{opt} dans l'équation de h(x).

$$w_{opt} = \sum_{k=}^{p} \alpha_k^*.y_k.x_k \qquad \text{(A.13)}$$

On trouve alors:

$$h(x) = \sum_{k=}^{p} \alpha_k^*.y_k.x_k.\ x + w_0 \qquad \text{(A.14)}$$

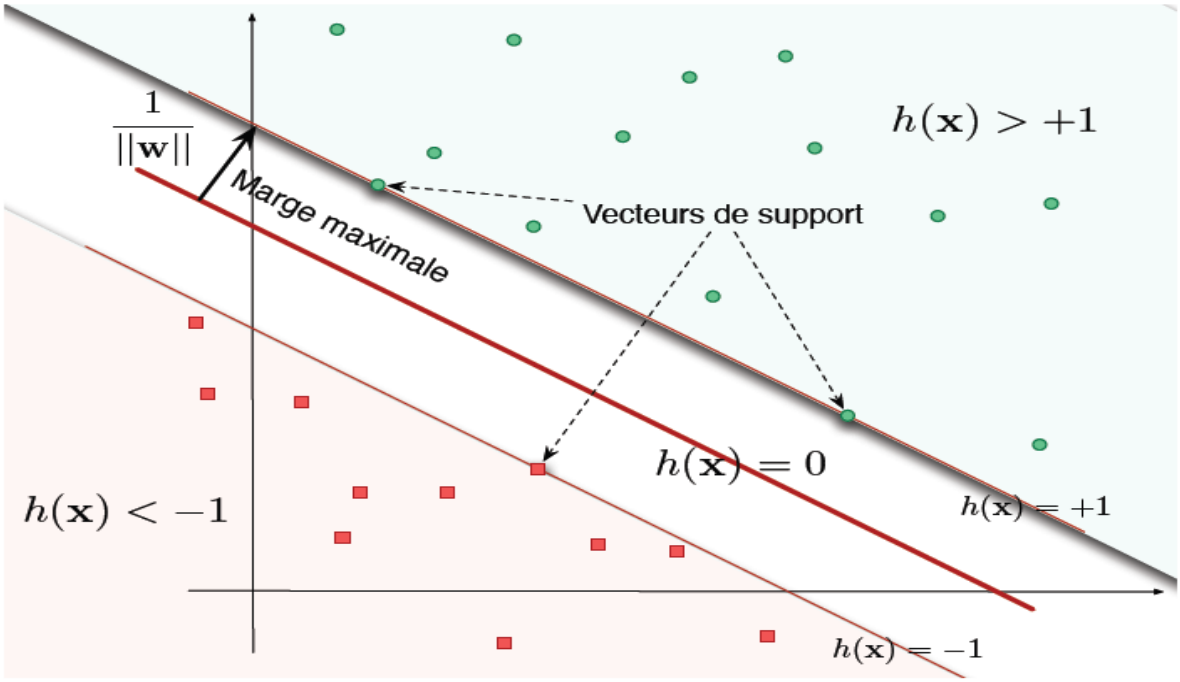


Figure. A.4. L'hyperplan optimal (en rouge) avec la marge maximale

Pour assouplir les contraintes (autoriser les erreurs de classification), les chercheurs ont introduit les variables 'ressort' (figure A.5). La formulation primale des SVM devient:

Minimiser $(\frac{1}{2}\|w\|^2 + C\sum_{k=1}^{p}\xi_k)$ sous les contraintes $y_k(w^T.x_k + w_0) \geq 1 - \xi_k$ (A.15)

Le problème dual reste le même sauf pour α qui est limité par la borne supérieur C.

$$\begin{cases} \text{Maximiser } L(\alpha) = \sum_{k=1}^{p}\alpha_k - \frac{1}{2}\sum_{i,j}^{p}\alpha_i.\alpha_j.y_i.y_j.x_i^T.x_j\ , \\ \\ \text{sous les contraintes : } \alpha_k \geq 0\ , \text{ et } \sum_{k=1}^{p}\alpha_k.\ x_k = 0 \end{cases} \quad \text{(A.16)}$$

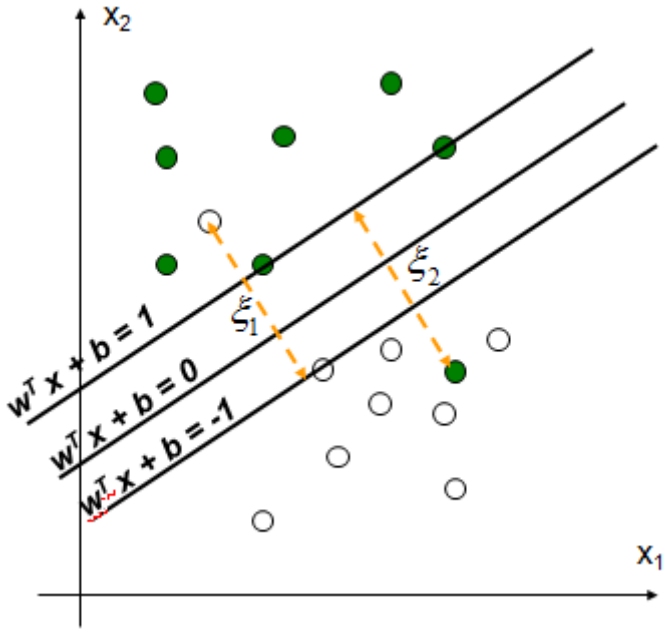


Figure. A.5. L'hyperplan optimal calculé avec des variables ξ_i 'ressort'

Lorsque les données ne sont pas linéairement séparables comme le montre la figure ci-dessous, le problème peut être résolu en utilisant les fonctions noyaux. Les méthodes à noyaux étendent les algorithmes de classification par SVM.

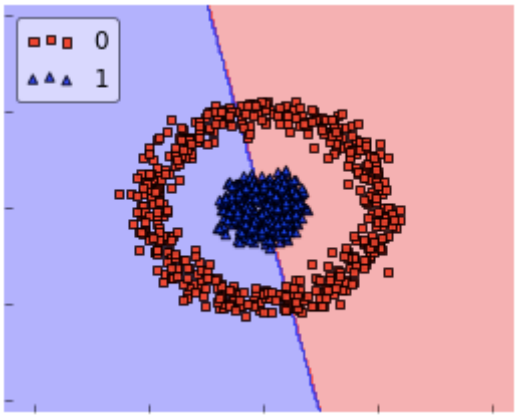


Figure. A.6. Le classificateur SVM avec un noyau linéaire ne parvient pas à séparer des données non linéairement séparables.

Au lieu de chercher un hyperplan dans l'espace des entrées, on passe dans un espace de représentation intermédiaire où les données peuvent être séparées linéairement. Le passage dans ce nouvel espace se fait par une transformation:

$$\Phi : \Re^d \longrightarrow F$$

$$X \longrightarrow \Phi(X) \tag{A.17}$$

La formulation primale des SVM devient:

$$\begin{cases} \text{Maximiser } L(\alpha)= \sum_{k=1}^{p} \alpha_k - \frac{1}{2}\sum_{i,j}^{p} \alpha_i . \alpha_j . y_i . y_j . \Phi(x_i)^T . \Phi(x_j) \\ \\ \text{Sous les contraintes : } 0 \leq \alpha_k \leq C, \text{ et } \sum_{k=1}^{p} \alpha_k . x_k = 0 \end{cases} \quad (A.18)$$

Pour résoudre ce problème, on utilise une fonction noyau connue sous le nom de Kernel trick qui vérifie :

$$K(x_i, x_j) = \Phi(x_i)^T . \Phi(x_j) \quad (A.19)$$

L'hyperplan séparateur en fonction de la kernel est donné par l'équation suivante:

$$h(x)= \sum_{k=}^{p} \alpha_k^* . y_k . k(x_k . x)+w_0 \quad (A.20)$$

Comme exemples de fonctions noyaux, on trouve:

• Linéaire : $k (x, x ') = x \times x '$

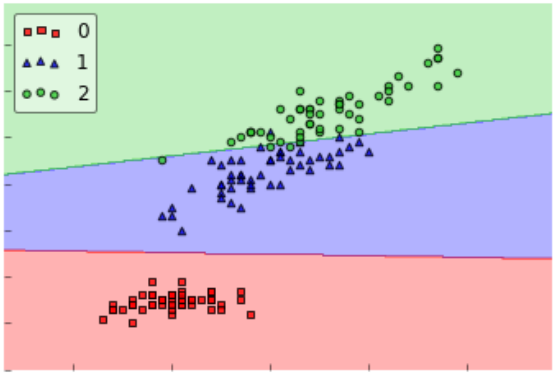


Figure. A.7. Le classificateur SVM avec noyau linéaire pour des données linéairement séparables.

• Gaussien : $k (x, x') = e^{-|x-x'|^2/2.\sigma^2}$

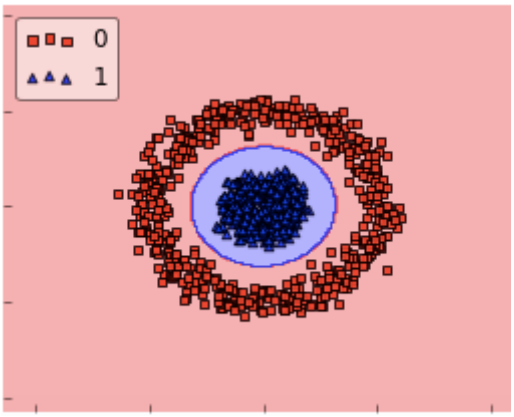


Figure. A.8. Le classificateur SVM avec un noyau RBF non linéaire.

• Polynomial : $k(x, x') = (x \times x')^d$ ou $k(x, x') = (c + x \times x')^d$

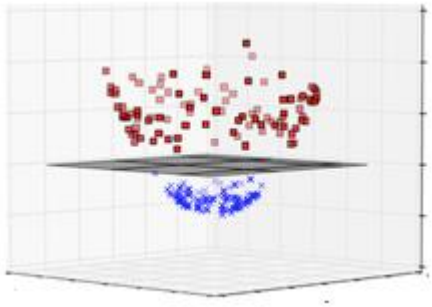


Figure. A.9. Le classificateur SVM avec un noyau polynomial.

Le noyau polynomial est moins utilisé, en pratique, pour des raisons d'efficacité (calculs et prédictions). Dans le cas où la taille des données est trop grande, il n'y a pas de règle pour choisir le noyau SVM. La seule façon est d'effectuer des simulations et de choisir le noyau qui donne les meilleurs résultats. La figure ci-dessous montre les procédures nécessaires pour déterminer l'hyperplan séparateur des SVM combinés avec les fonctions noyau (Kernel).

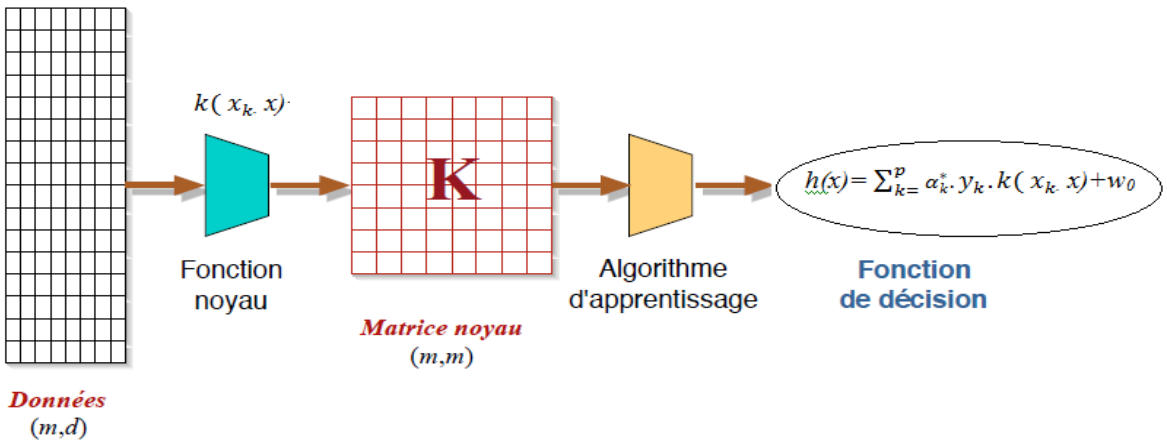


Figure. A.10. Chaîne de traitement des SVM à noyaux.

Pour étendre le SVM au problème de classification multi-classes (classes M> 2) plusieurs méthodes peuvent être utilisées. Les plus utilisées sont un contre tous (OVA : One Versus All) et un contre un (OVO : One Versus One). Pour la méthode « un contre tous » il faut construire M classifieurs binaires en indexant par 1 les échantillons d'une des classes et par -1 les échantillons de toutes les autres classes. La classe choisie est celle qui obtient le meilleur score. Pour la méthode « un contre un » il faut construire M (M -1)/2 classifieurs binaires,

chaque classifieur confrontant deux classes parmi les M classes. Pour le choix de la classe, on procède par vote majoritaire. La figure suivante illustre les principes des deux méthodes.

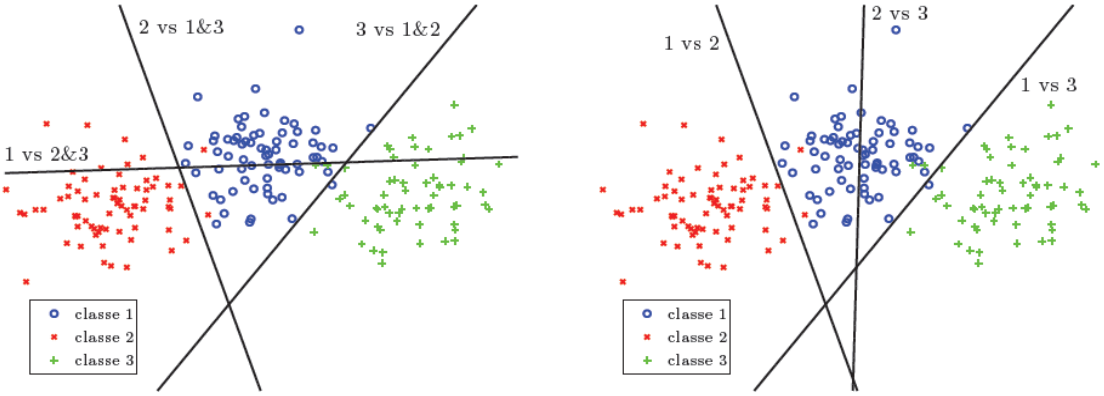


Figure. A.11. Exemple d'hyperplans de séparation SVM OVA et OVO.

B.1 LBP Standard

$LBP_{P,R}$, P étant le nombre de voisins et R est le rayon du voisinage, est calculé de telle sorte que pour un pixel central donné dans une image, un motif est calculé en comparant la valeur du pixel central avec celle de ses voisins.

Les équations (B.1) et (B.2) montrent comment le $LBP_{P,R}$ est calculé.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c).2^p \quad \text{(B.1)}$$

$$s(x) = \begin{cases} 1 & \text{pour} \quad x \geq 0 \\ 0 & \text{sinon} \end{cases} \quad \text{(B.2)}$$

Dans l'équation (B.1) et sur la figure B.1, g_c est le niveau du gris du pixel central en rouge, et g_p est le niveau du gris de ses voisins en vert. Selon l'application voulue, le rayon du voisinage peut être différent.

Dans la définition générale donnée ci-dessus, si l'on suppose que les coordonnées de g_c sont (X_c, Y_c), alors les coordonnées de g_p du voisinage sont données par : $X = X_c - R.Sin(2\pi p/P)$ et $Y = (Y_c + R.Cos(2\pi p/P))$.

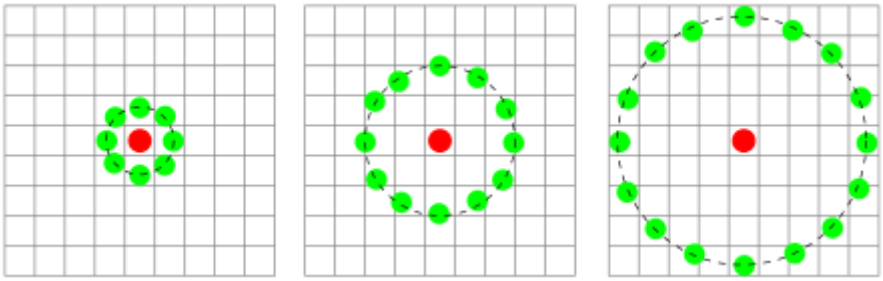


Figure. B.1. Opérateur LBP étendu pour utiliser des voisinages de différentes tailles.

(1) P=8, R=1 (2) P=12, R=2.5 (3) P=16, R=4.

B.2 Modèle LBP uniforme uLBP

Un ULBP est appelé uniforme si le motif binaire contient au plus deux transitions bits (de 0 à 1 ou vice versa) lorsque le motif de bits est parcouru de manière circulaire.

La valeur U pour LBP est le nombre de transitions spatiales (changements de bits 0/1) dans le motif.

La notation du LBP uniforme d'ordre 2 est $LBP_{P,R}^{U2}$ [78]. U2 signifie l'utilisation uniquement des motifs uniformes qui satisfont $U \leq 2$.

Pour calculer un LBP uniforme on utilise l'équation suivante:

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_{P-1} - g_0)| + \sum_{p=0}^{p-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (B.3)$$

Le nombre de motifs uniformes est égal à P × (P-1) + 2.

Pour obtenir l'histogramme $LBP_{P,R}^{U2}$ d'une image I(X × Y), le motif $LBP_{P,R}^{U2}$ de chaque pixel (i, j) est utilisé dans le calcul comme suit:

$$H(k) = \sum_{i=2}^{X-1} \sum_{j=2}^{Y-1} f(U(LBP_{P,R})(i,j), k), \quad k = [0, K-1] \quad (B.4)$$

Dans le cas d'un $LBP_{P,R}^{U2}$ avec P égal à 8, on aboutit à un histogramme comme vecteur caractéristique à 59 motifs (58 motifs sur 256 sont uniformes et tous les motifs non uniformes sont étiquetés avec une seule étiquette) comme le montre la Figure B.3.

Dans l'équation (B.2), K est la valeur maximale du motif $LBP_{P,R}^{U2}$. Dans cet histogramme, chaque motif LBP a un facteur de pondération égal à 1, avec :

$$f(U(LBP_{P,R})(i,j),k) = \begin{cases} 1 & U(LBP_{P,R})=k \\ 0 & \text{sinon} \end{cases} \quad (B.5)$$

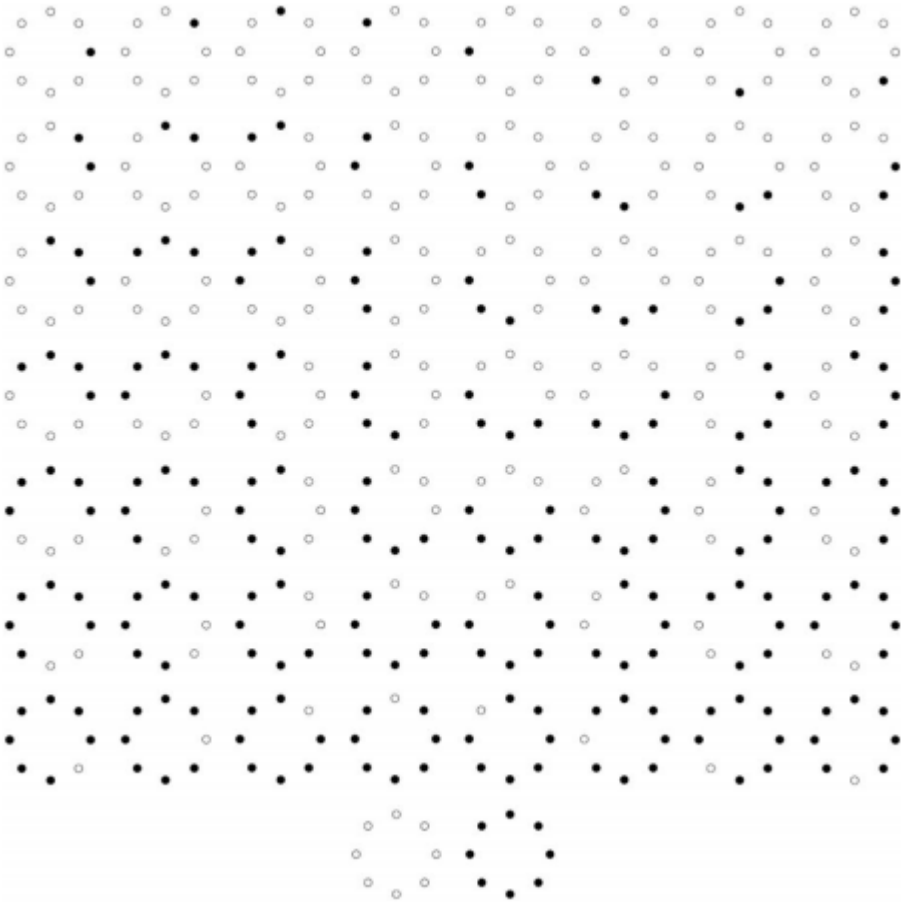


Figure. B.2. 58 motifs LBP uniformes lorsque P = 8. Les points noir et blanc représentent les valeurs de bits de 1 et 0 dans la sortie 8 bits de l'opérateur LBP.

B.3 Modèle LBP codé par transition tLBP

La règle de codage utilisée dans le LBP classique compare les valeurs des niveaux de gris des pixels voisins à celle du pixel central. Cela donne une connaissance approximative du pixel par rapport au centre, mais les relations entre les pixels ayant la même valeur binaire sont perdues.

Dans le LBP codé par transition, la valeur binaire est composée de comparaisons entre les pixels voisins dans le sens des aiguilles d'une montre pour tous les pixels à l'exception du pixel central (voir Figure B.3). On peut trouver la valeur décimale du pixel central dans le cas du tLBP en utilisant l'équation suivante :

$$\text{tLBP}_{\text{P,R}} = s(g_0 - g_{P-1}) + \sum_{p=1}^{P-1} s(g_p - g_{p-1}).2^p \qquad \text{(B.6)}$$

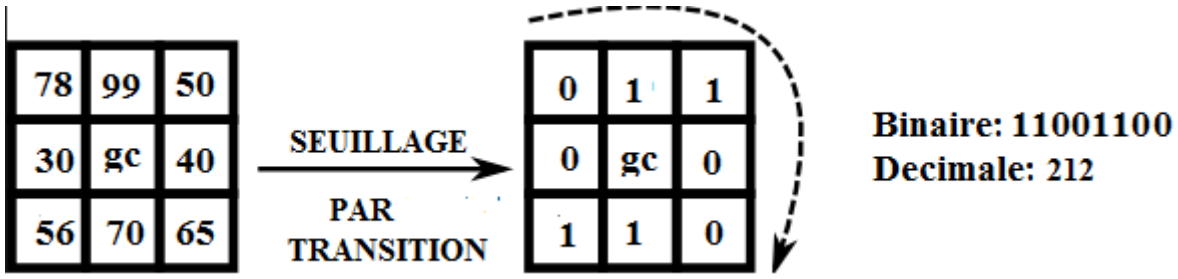


Figure. B.3. LBP codé par transition (tLBP)

B.4 Modèle LBP codé par direction (dLBP)

Le dLBP compare seulement quatre pixels adjacents, mais comprend également les informations de direction dans un extra bit. La motivation du dLBP est de fournir une meilleure information concernant le motif local dans le sens des fonctions de direction. Il existe quatre directions de base à travers le pixel central dans un LBP (voir Figure B.4). Le dLBP code la variation d'intensité le long de ces directions avec un bit (0/1). On peut trouver la valeur décimale du pixel central pour le dLBP en utilisant l'équation suivante :

$$\mathrm{dLBP}_{\mathrm{P,R}} = \sum_{p'=0}^{P'-1}\left(s(g_{p'} - g_c)\,.\,s(g_{p'+P'} - g_c)2^{2.p'} + s(\left|(g_{p'} - g_c) - (g_{p'+P'} - g_c)\right| - \left|(g_{p'} - g_c) - (g_{p'+P'} - g_c)\right|)\,2^{2.p'+1}\right), \quad \text{(B.6)}$$

avec P = 2P'

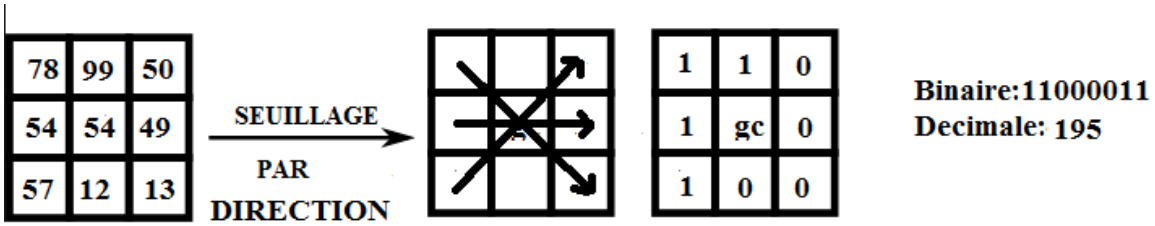


Figure. B.4. LBP codé par direction (dLBP)

B.5 Modèle LBP modifié (mLBP)

mLBP compare les pixels dans le voisinage 3 × 3 à leur moyenne au lieu du pixel central.

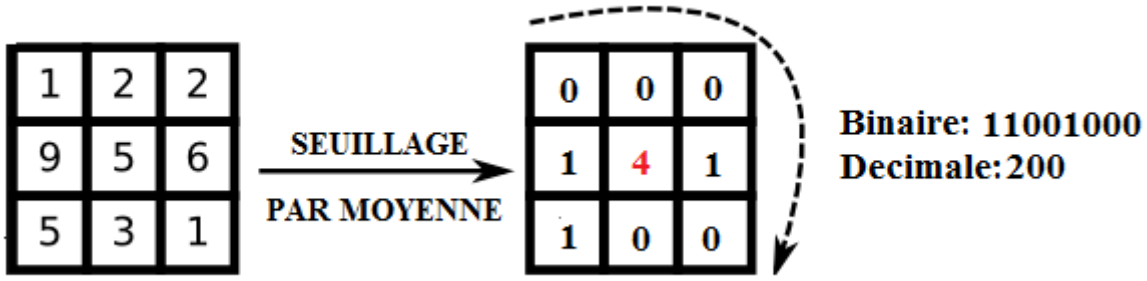


Figure. B.5. LBP modifié codé sur 8 bits (mLBP).

C.1 ART Algorithme

La fonction qui permet de décomposer une image en utilisant la transformation angulaire radial.

Table art(image[N][M])

```
//déclaration des vecteurs IART + RART + MART
Table:IART, RART, ART ;
// coordonnées polaire (r,θ) de l'image
float angle,r;
float X,Y;
//taille de l'image NxM
integer N,M;

//boucles de décomposition de l'image.
For n<=8; //L'ordre de décomposition n.

  For m<=4; //L'ordre de répétition m.

    For i<N :

      For j<M:

        X=((2.*j/M)-1);Y=(1-(2.*i/N));
                r=sqrt(pow(X,2)+pow(Y,2));
                if(r<=1)
                angle=atan(Y,X);
               im= image[i][j]/255;
                  if(n>0)
                    RART +=(1/PI)*im*cos(m*atan(Y,X))*cos(n*r*PI);
                  IART +=(1/PI)*im*sin(m*atan(Y,X))*cos(n*r*PI);
                   elseif(m==0)
                     RART +=im;
                     IART +=0;
                   else
                    RART +=(1/PI)*im*cos(m*atan(Y,X));
                    IART +=(1/PI)*im*sin(m*atan(Y,X));}
                    end
                  end
                  end
                  end
                end
ART = RART+IART;
return ART;
end
```

Un vecteur, issu d'une décomposition d'une image par ART et pour un ordre de décomposition n=8, est composé de 64 éléments quelle que soit la taille de l'image.

C.2 LMI Algorithme

La fonction qui permet de décomposer une image en utilisant les moments de legendre invariant (LMI).

Table lmi (image[N][M])

```
//déclaration du vecteur LMI
Table:LMI ;
// coordonnées polaire (r,θ) de l'image
float X,Y;
//taille de l'image NxM
integer N,M;

//boucles de décomposition de l'image.
For n<=8; //L'ordre de décomposition n.
  For m<=4; //L'ordre de répétition m.
    For i<N :
      For j<M:
          X=((sqrt(2.)*j/N)-(sqrt(2.)/2));
          Y=((sqrt(2.)/2)-(sqrt(2.)*i/M));
          im= image[i][j]/255;
          LMI=pol(Y,m)*pol(X,n)*im;
        end
     end
    end
   end
 return LMI;
end

float pol(float x,integer i)
for i<n
  if i==0
    P=1;
  elseif i==1
    P=x;
   Else
    Pn+1=((2.n+1)/(n+1).x.Pn )-(n/(n+1).x.Pn-1)
   End
 End
 End
 return P;
end
```

Un vecteur, issu d'une décomposition d'une image par LMI et pour un ordre de décomposition n=8, est composé de 32 éléments quelle que soit la taille de l'image.

C.3 PMZ Algorithme

La fonction qui permet de décomposer une image en utilisant les moments de legendre invariant (PMZ).

```
Table pmz (image[N][M])

//déclaration des vecteurs IPMZ + RPMZ + MPMZ
Table:IPMZ, RPMZ, PMZ ;
// coordonnées polaire (r,θ) de l'image
float angle,r;
float X,Y;
//taille de l'image NxM
integer N,M;
//boucles de décomposition de l'image.
For n<=8; //L'ordre de décomposition n.
  m=n ;
   For m>=0; //L'ordre de répétition m.
    m-- ;
      For i<N :

        For j<M:

          X=((2.*j/M)-1);Y=(1-(2.*i/N));
                    r=sqrt(pow(X,2)+pow(Y,2));
                    if(r<=1)
                     angle=atan(Y,X);
                    im= image[i][j]/255;
                      RPMZ +=(n+1/PI) mz(n,m,r)*im*cos(m*atan(Y,X));
                    IPMZ +=(n+1/PI)  mz(n,m,r)*im*sin(m*atan(Y,X));;
                    end
        end
       end
    end
  end
PMZ = RPMZ+IPMZ;
return PMZ
end

float mz(integer p,integer q,float r)
 for k <= (p-q)
   M1=(pow(-1.0,k) )*( factorial(2*p-k+1) );
   M2=(factorial(k)*factorial(p-abs(q)-k)*factorial(p+abs(q)+1-k));
   M=M1/M2;
   N=p-k;
   sum +=   (pow (r , N ))*M ;
 end
 Return sum;
end
```

Un vecteur, issu d'une décomposition d'une image par PMZ et pour un ordre de décomposition n=8, est composé de 72 éléments quelle que soit la taille de l'image.

Références

[1] https://da.keylemon.com/

[2] H.-K. Jee, S.-U. Jung, J.-H. Yoo, "Liveness detection for embedded face recognition system.", International Journal of Biological and Medical Sciences, 1(4):235–238, 2006.

[3] G. Pan, L. Sun, Z. Wu, "Liveness detection for face recognition.", INTECH Open Access Publisher, 2008.

[4] J. Maatta, A. Hadid, M. Pietikainen, "Face spoofing detection from single images using micro-texture analysis ", IEEE international joint conference In Biometrics (IJCB), pp. 1–7, 2011.

[5] P.Gautam, K. S. Jayash, "Face Liveness Detection using Local Diffused Patterns ", International Journal of Computer Applications, Vol. 149 – No.4, pp.0975 – 8887, 2016.

[6] W. Bao, H. Li, N. Li, W. Jiang, "A liveness detection method for face recognition based on optical flow field", IEEE International Conference In Image Analysis and Signal Processing, pp. 233–236., 2009.

[7] T. Wang, J. Yang, Z. Lei, S. Liao, S. Z. Li., "Face liveness detection using 3d structure recovered from a single camera", In Biometrics (ICB), 2013

[8] S. Milborrow, F. Nicolls., "Active shape models with sift descriptors and mars", In VISAPP, Vol. 2, pp. 380–387, 2014.

[9] V. Kazemi, J. Sullivan., "One millisecond face alignment with an ensemble of regression trees.", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.1867-1874, June, 2014

[10] M. De Marsico, M. Nappi, D. Riccio, J.-L. Dugelay., "Moving face spoofing detection via 3d projective invariants.", 5th IAPR International Conference In Biometrics (ICB), pp. 73–78., 2012.

[11] J Fokkema, " Using a Challenge to Improve Face Spoofing Detection.", essay.utwente.nl. 2016.

[12]W. Di, H. Hu, Anil K. Jain, "Face Spoof Detection with Image Distortion Analysis",IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

[13]http://biometrics.cse.msu.edu/pubs/databases.html

[14] A. Anjos, S. Marcel, "Counter-measures to photo attacks in face recognition: A public database and a baseline," in Proc. IJCB, pp. 1–7, 2011.

[15] X. Tan, Y. Li, J. Liu, L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model." in Proc. ECCV, pp. 504–517, 2010.

[16] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S. Z. Li, "A face antispoofing database with diverse attacks." in Proc. ICB, pp. 26–31, 2012.

[17] I. Chingovska, A. Anjos, S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing." in Proc. IEEE BIOSIG, pp. 1–7, 2012.

[18] K. Kollreider, H. Fronthaler, J. Bigun, "Verifying liveness by multiple 730 experts in face biometrics", IEEE Computer Society Conference on Com- 731 puter Vision and Pattern Recognition Workshops, pp. 1–6, 2008.

[19] Y.Kim, J.Na, S.Yoon, J.Yi, "Masked fake face detection using radiance measurements", Journal of the Optical Society of America A, Vol. 2, issue 4, pp. 760–766, 2009.

[20] Z. Zhang, D. Yi, Z. Lei, S. Li, "Face liveness detection by learning multispectral reflectance distributions", In IEEE International Conference on Automatic Face Gesture Recognition and Workshops, pages 436 –441, March 2011.

[21] R. Basri, D.W. Jacobs, "Lambertian Reflectance and Linear Subspaces", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.25, issue.2, 2003.

[22] N. Kose, J.-L. Dugelay, "Countermeasure for the protection of face recognition systems against mask attacks. In IEEE International Conference on Automatic Face and Gesture Recognition, April 2013.

[23] TABULA RASA Project, http://www.tabularasa-euproject.org/, 2010.

[24] MORPHO, http://www.morpho.com/, 2010.

[25] P. M. Ojala T., M. T. Multiresolution, "gray-scale and rotation invariant texture classification with local binary patterns.", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, issue 7, pp. 971–987, 2002.

[26] N. Erdogmus, S. Marcel., “Spoofing 2d face recognition systems with 3d masks.”, In Conference of the Biometrics Special Interest Group (BIOSIG), 2013.

[27] M. C. Wallace R., McLaren, “Inter-session variability modelling and joint factor analysis for face authentication.”, In International Joint Conference on Biometrics, pp. 1–8, 2011.

[28] N. Kose, J.L. Dugelay, “Shape and texture based countermeasure to protect face recognition systems against mask attacks”, IEEE Computer Vision and Pattern Recognition Workshop, on Biometrics (CVPRW), Vol. 747, pp. 111–116, 2013.

[29] https://www.idiap.ch/dataset/3dmad

[30] M. Turk, A. Pentland., “Eigenfaces for Recognition.”, J. Cognitive Neuroscience, Vol. 3, issue.1, pp. 71-86, 1991.

[31] B. Moghaddam, A. Pentland, "Probabilistic visual learning for object representation", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.19, pp.696–710, 1997.

[32] J. Wu, Z.-H. Zhou., “Face recognition with one training image per person.”, Pattern Recognition Letter,Vol. 23, issue.14, pp. 1711–1719, 2002.

[33] J.Yang, D. Zhang, A.F. Frangi, J.Yang, “Two-dimensional PCA: a new approach to appearance-based face representation and recognition.”, IEEE Trans. Pattern Anal. Mach. Intell., pp.131–137,2004.

[34] B. Moghaddam, A. Pentland, “Probabilistic visual learning for object Representation.”, IEEE Trans. Pattern Anal. Mach. Intell., Vol.19, issue7, pp. 696–710, 1997.

[35] X. He, X. Yan, Y. Hu, p. Niyogi, H. Zhang, “Face recognition using Laplacianfaces,”, IEEE Trans. Pattern Anal. Mach. Intell., Vol.27, issue.3, pp.328–340, 2005.

[36] D. L. Swets, J. Weng, "Using discriminant eigenfeatures for image retrieval", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.18, pp. 831– 836, 1996.

[37] P. Belhumeur, J. Hespanha, D. Kriegman, “Eigenfaces vs. fisherfaces: recognition using class specific linear projection.”, IEEE Trans. Pattern Anal. Mach. Intell., Vol.19, isuue7, pp. 711–720, 1997.

[38] M.S. Bartlett, H.M. Lades, T. SEJNOWSKI, "Independent component

representation for face recognition", Proceedings SPIE Symposium on Electronic Imaging:Science and Technology, Vol. 3299, pp. 528-539, 1998.
[39] B. Draper, K. Baek, M.S. Bartlett, R. Beveridge, "Recognizing Faces with PCA and ICA", Computer Vision and Image Understanding, Vol. 91, pp. 115-137, 2003.
[40] B. Moghaddam, "Principal manifolds and probabilistic subspace for visual recognition", IEEE Trans. Pattern Anal. Mach. Intell., Vol.24, issue6, pp. 780–788, 2002.
[41] R. Brunelli, T. Poggio. "Face recognition: features versus templates", IEEE Trans. Pattern Anal. Mach. Intell, pp-1042–1062, 1993.
[42] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, Malsburg von der, R. Wurtz. "Distortion invariant object recognition in the dynamic link Architecture", IEEE Trans. Comput., Vol.42, issue. 3, pp. 300–311, 1993.
[43] T.S. Lee, "Image representation using 2-d Gabor wavelets", IEEE Trans. Pattern Anal. Mach. Intell., Vol.18, issue.10, pp. 959–971, 1996.
[44] L. Wiskott, R. Fellous, N. Kruger, C. von Malsburg. "Face recognition by elastic bunch graph matching", IEEE Trans. Pattern Anal. Mach. Intell., pp.775–779, July 1997.
[45] N. Ahmed, T. Natarajan, K. Rao, "Discrete cosine transfom ", IEEE Transactions on Computers, Vol.23, issue1, pp. 90–93, janvier 1974.
[45] K. R. Rao, P. Yip. "Discrete Cosine Transform : Algorithms, Advantages, Applications". Academic Press, Boston, 1990.
[47] Z. Hafed, M. D. Levine. "Face recognition using the discrete cosine transform". International Journal of Computer Vision, Vol. 43, issue. 3, pp. 167–188, juillet 2001.
[48] A. Samra, S. E. T. G. Allah, R. Ibrahim., "Face recognition using wavelet transform, fast Fourier transform and discrete cosine transform.", IEEE International Symposium on Micro-NanoMechatronics and Human Science, Vol. 1, pp. 272–275, décembre 2005.
[49] A. Pentland, B. Moghaddam , T. Starner., "View-based and modular eigenspaces for face recognition.", In Proceedings, IEEE Conference, 1994.
[50] F. Samaria, "Face segmentation for identification using hidden Markov Models ", in: British Machine Vision Conference, BMVA Press, pp. 399–408, 1993.
[51] 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.

[52] M.K.Hu, "Visual Pattern Recognition by Moment Invariants", IRE Transactions on Information Theory, Vol. 49, pp.179-187, 1961.

[53] S. Annadurai, A. Saradha, "Face Recognition Using Legendre Moments", Proceedings of the Fourth Indian Conference on Computer Vision, Graphics & Image Processing, 2004.

[54] M. R. Teague," Image analysis via the general theory of moments", J. Opt. Soc. Am., Vol.70, pp.920–930, 1980.

[55] F. Zernike: "Diffraction theory of the cut procedure and its improved form, the phase contrast method.", Physica, Vol.1, pp. 689–704, 1934.

[56] R. Mukundan. , K.R. Ramakrishnan, "Fast computation of Legendre and Zernike moments", Pattern Recognition, Vol. 28, issue. 9, pp. 1433-1442, 1995.

[57] A. Nabatchian, I. Makaremi, "Pseudo-Zernike Moment Invariants for Recognition of Faces Using Different Classifiers in FERET Database", Third International Conference on Convergence and Hybrid Information Technology, Vol. 1, pp. 933- 936, 2008.

[58] S.H. Lee,S.Sharma, L. Sang, J. Park, Y.G.Park,"An Intelligent Video Security System using Object Tracking and Shape Recognition", ACVIS LNCS Springer-Verlag, Berlin Heidelberg, Vol. 6915, pp. 471-482, 2011.

[59] O.Wahdan, K. Omar, F. Nasrudin, "Logo Recognition System Using Angular Radial Transform Descriptors", Journal of Computer Science, Vol. 7, pp. 1416-1422, 2011.

[60] Fang, J., Qui, G.,"Human Face Detection using Angular Radial Transform and Support Vector Machines", International Conference on Image Processing, Vol. 1, pp. 69-72, 2003.

[61]M. Bober,"MPEG-7 Visual Shape Descriptors", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, issue 6, pp. 716-719, 2001.

[62] N. Morizet, F. Amiel, I. D. Hamed, T. Ea." A comparative implementation of PCA face recognition algorithm. 14th IEEE International Conference on Electronics, Circuits and Systems, Vol. 2, pages 865 –868, décembre 2007.

[63] K. Etemad , R. Chellappa, "Discriminant analysis for recognition of human face images", Journal of the Optical Society of America A, Vol. 14, issue8, pp.1724–1733, Octobre 1997.

[64] W. Zhao, R. Chellappa, A. Krishnaswamy., "Discriminant analysis of principal components for face recognition", 3rd IEEE International Conference on Face and Gesture Recognition, pp. 36–341, avril 1998.

[65] Z. Mu-chun., "Face recognition based on fastICA and RBF neural networks",

International Symposium on Information Science and Engineering,Vol. 1, pp. 588 – 592, decembre 2008.

[66] M. J. Er, W. Chen, S. Wu ,"High-speed face recognition based on discrete cosine transform and RBF neural networks", IEEE Transactions on Neural Networks, Vol.16, issue. 3, 676–691, mai 2005.

[67] G. Guo, S. Z. Li, K. Chan, "Face recognition by support vector machines", Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 10.1109, pages 196 – 201, mars 2000.

[68] E. Osuna, R. Freund, F. Girosit., 'Training support vector machines : an application to face detection.", Proceedings of the Conference on Computer Vision and Pattern Recognition, CVPR'97, San juan, PR, USA, pp. 130 – 136, juin 1997.

[69] J. Huang, X. Shao, H. Wechsler., "Face pose discrimination using support vector machines (SVM)", Proceedings of the 14th International Conference on Pattern Recognition, ICPR'98, Brisbane, Australia, Vol. 1, pp. 154 – 156, août 1998.

[70] B. Heisele, P. Ho, T. Poggio, "Face recognition with support vector machines : Global versus component-based approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 688–694, avril 2001.

[71] Marti A. Hearst, "Support Vector Machines.", IEEE Intelligent Systems, Vol. 13, issue. 4, pp. 18-28, Jul/Aug, 1998.

[72] V. Vapnik., S. Kotz, "Estimation of Dependences Based on Empirical Data", Springer Series in Statistics, ISBN: 978-0387907338, 1982.

[73] T.T. M aenpaa,T. Ojala., M. Pietikainen, M.Soriano," robust texture classification by subsets Local Binary Pattern", Proc.15th int.Conf. on Pattern Recognition, 2000.

[74] Z. Guo, L. Zhang, D. Zhang, "Rotation invariant texture classification using lbp variance (lbpv) with global matching.", Elsevier Pattern Recognition, pp. 706–719, 2010.

[75] N. Kose , J.-L. Dugelay, "Reflectance analysis based countermeasure technique to detect face mask attacks", Int. Conf. on Digital Signal Processing (DSP).", pp. 1–6, 2013.

[76] N. Kose, J.-L. Dugelay," Countermeasure for the protection of face recognition systems against mask attacks," IEEE Automatic Face and Gesture Recognition (FG), pp. 1–6, 2013.

[77] N. Kose, J.-L. Dugelay , " Mask spoofing on face recognition and countermeasures", Elsevier - Image and Vision Computing Journal, under review, Vol 15, 2014.

[78] Z. Guo, L. Zhang, D. Zhang, “Rotation invariant texture classification using LBP variance (LBPV) with global matching”, Pattern Recognition, Vol. 43, issue 3, pp. 706-719 2010.

[79] J. Trefn`y , J. Matas., “Extended set of local binary patterns for rapid object detection “, Proceedings of the Computer Vision Winter Workshop, 2010.

[80] R. Kimmel, M. Elad, D. Shaked, R. Keshet, I. Sobel., ” A variational framework for retinex.”, International Journal of Computer Vision, Vol. 46, pp. 7–23, 2003.

[81] A. Mohan, C. Papageorgiou, and T. Poggio, “Example-based object detection in images by components”, PAMI, Vol. 23, issue 4,pp. 349-361, 2001.

[82] C.C. Chang, C.J. Lin, "LIBSVM: a library of support vector machines", Software available at http://www.csie.ntu.edu.tw/˜cjlin/libsvm2001.

[83] H.Bensenane , M. Keche , ” Réalisation d’un système biométrique basé sur la reconnaissance du visage”, Magister – Technique de communications modernes, 2013.

[84] S.Annadurai, A.Saradha, “Face Recognition Using LegendreMoments”, http://citeseerx.ist.psu.edu