Université d'Abomey-Calavi

N° d'ordre : 97

Ecole Doctorale des Sciences de l'Ingénieur (ED-SDI)

Thèse de Doctorat

Présentée pour l'obtention du grade de

Docteur de l'Université d'Abomey-Calavi

Spécialité : Signal et Image

Authentification par la biométrie multimodale sans contact : Stratégie de fusion de scores basée sur l'approche séquentielle adaptée à l'utilisateur

Présentée par :

Abdou-Aziz SOBABE ALI TAHIROU

Soutenue publiquement le 19/02/2021

Devant le jury composé de :

Président :	Eugène C. EZIN	Professeur Titulaire	Université d'Abomey-Calavi (Bénin)
Rapporteurs :	Ahmed Dooguy KORA	Professeur Titulaire	Ecole Supérieure Multinationale des Télécommunications-ESMT (Sénégal)
	Théodore M. Y. TAPSOBA	Professeur Titulaire	Université Nazi BONI (Burkina Faso)
Examinateurs :	Théophile ABALLO	Maître de Conférences	Université d'Abomey-Calavi (Bénin)
	Michel DOSSOU	Maître de Conférences	Université d'Abomey-Calavi (Bénin)
Directeur de thèse :	Antoine VIANOU	Professeur Titulaire	Université d'Abomey-Calavi (Bénin)
Encadreur :	Tahirou DJARA	Maître de Conférences	Université d'Abomey-Calavi (Bénin)
Invité :	Patrick SOTINDJO	Assistant	Université d'Abomey-Calavi (Bénin)

Dédicace

Je dédie cette œuvre à :

- ma mère, Salmata DRAMANE NARO ;
- la mémoire de mon feu père, Tahirou SOBABE ALI.

Remerciements

Je tiens d'abord et avant tout à remercier ALLAH, le Tout Puissant qui m'a facilité la rédaction de cette thèse. Ce travail n'aurait pu connaître un aboutissement heureux sans son assistance et sa protection.

J'exprime ma totale déférence et ma reconnaissance au Professeur Antoine VIANOU, Professeur Titulaire, Membre fondateur de l'Académie Nationale des Sciences, Arts et Lettres du Bénin, Vice-Recteur Honoraire de l'UAC, Directeur de l'Ecole Doctorale des Sciences de l'Ingénieur qui a assumé la lourde charge de Directeur de thèse. Je le remercie du fond du cœur pour ses nombreuses marques de soutien ainsi que les encouragements dont j'ai bénéficié de sa part tout au long de ces travaux. Je lui témoigne ma profonde gratitude.

J'exprime également ma profonde reconnaissance au Dr Tahirou DJARA, Maître de Conférences des Universités du CAMES, qui a assuré mon encadrement, sous la conduite du Professeur Antoine VIANOU. Je ne saurais jamais le remercier assez pour avoir dirigé et orienté mes travaux de recherche. Il est l'architecte hors pair qui a façonné et orienté l'ensemble de ces travaux. Puisse Dieu le lui rendre au centuple.

Je voudrais tout particulièrement remercier le Dr Marc Kokou ASSOGBA, Maître de Conférences des Universités du CAMES, Directeur du LETIA (Laboratoire d'Electrotechnique de Télécommunication et d'Informatique Appliquée) pour sa contribution à l'encadrement de ma thèse. Qu'il soit assuré de ma reconnaissance.

Je voudrais remercier le Professeur Eugène C. EZIN, Professeur Titulaire en Informatique et Intelligence Artificielle, Directeur de l'Institut de Formation et de Recherche en Informatique (IFRI-UAC) pour ses précieux conseils et orientations.

Mes remerciements vont également à l'endroit du Dr Macaire AGBOMAHENA, Maître de Conférences, pour ses précieux conseils et orientations au sein de l'Ecole Doctorale des Sciences de l'Ingénieur.

Le Dr Jean-Marie KABASELE TENDAY (Collaborateur Scientifique à l'Université Catholique de Louvain) trouvera à travers cette œuvre toute ma reconnaissance pour ses sages conseils et sa motivation pour la recherche scientifique.

Je tiens à remercier tous les membres de mon jury pour avoir consacré une partie de leur précieux temps afin de juger mon travail, surtout dans le contexte de la pandémie de COVID-19.

A mon épouse Barikissou AMADOU, mes enfants Abd Rakib, Abdel Hakim et Samiath, mes sœurs et frères Djamilath, Faouziath, Abd Nasser, Abd Rahim et Mohamed, ma nièce Moufidatou DRAMANE NARO, je tiens à dire un grand merci pour leur affection qui m'a été d'une grande utilité.

Mes travaux de recherche se sont déroulés conjointement au LETIA et au LATIP (Laboratoire d'Analyse, de Traitement de l'Image et de la Parole) de l'IITECH (Institut d'Innovation Technologique). Mon passage au niveau de ces deux laboratoires m'a permis d'avoir des échanges fructueux avec les différents membres de leurs équipes de recherche. Je tiens à les remercier pour la fructueuse collaboration que nous avons eue. Je voudrais particulièrement citer Messieurs Blaise BLOCHAOU, Jéhovanis SONON et Eléazar DAOUDOU, membres du LATIP.

Pour finir, j'adresse mes sincères remerciements à tous mes parents, amis, collègues, Responsables Administratifs et tous ceux qui ont contribué de quelque manière à la réalisation de cette œuvre.

Résumé

Dans cette thèse, nous présentons l'architecture de fusion séquentielle adaptée de scores en biométrie. Elle consiste à combiner les scores provenant de modalités de biométrie pure (visage et empreinte digitale sans contact) et les scores provenant de métadonnées (couleur de peau). Dans une première étape, nous effectuons la reconnaissance faciale avec l'algorithme LBPH (Local Binary Patterns Histogram). Nous développons ensuite une méthode originale pour l'authentification par la couleur de peau avec l'algorithme des k-moyennes. Les résultats obtenus montrent une augmentation de l'AUC (Area Under the Curve) qui est passée de 91,6% pour le visage à 94,4% pour le visage combiné à la couleur de peau. Dans une deuxième étape, nous utilisons les réseaux de neurones convolutionnels pour l'authentification par l'empreinte digitale. L'évaluation de la méthode d'intelligence artificielle révèle une confusion à l'échelle d'un individu sur 19, soit un taux de confusion de 5,26%. Les résultats globaux des trois sources biométriques montrent une amélioration des performances de notre algorithme en termes de taux de reconnaissance et de temps d'exécution. Ainsi, nous avons 49% des utilisateurs reconnus à la première étape pour notre méthode séquentielle adaptée contre 41% de reconnaissance pour la méthode séquentielle. De même, notre algorithme a un temps moyen d'exécution par utilisateur de 7,924 secondes contre 8,182 secondes pour la méthode séquentielle.

Mots clés : *Multibiométrie, Fusion séquentielle adaptée, Métadonnée, Biométrie douce, Authentification par la couleur de peau, algorithme des k-moyennes.*

Abstract

In this thesis, we present the adapted sequential fusion architecture of biometric scores. It consists of combining scores from pure biometric modalities (face and contactless fingerprint) and scores from metadata (skin color). Firstly, we perform facial recognition with the LBPH (Local Binary Patterns Histogram) algorithm. We then, develop an original method for skin color authentication with the k-means algorithm. The obtained results show an increase in AUC (Area Under the Curve) from 91.6% for the face to 94.4% for the face combined with skin color. Secondly, we use convolutional neural networks for fingerprint authentication. The Artificial Intelligence assessment revealed confusion at the level of one in 19 individuals, for a confusion rate of 5.26%. The overall results of the three biometric sources show an improvement in the performance of our algorithm in terms of recognition rate and execution time. Thus, we have 49% of users recognized in the first step for our adapted sequential method versus 41% recognition for the sequential method. Similarly, our algorithm has an average execution time per user of 7.924 seconds compared to 8.182 seconds for the sequential method.

Keywords: *Multibiometrics, Adapted sequential fusion, Metadata, Soft biometrics, Skin color authentication, k-means algorithm.*

Sommaire

Table des figures

Liste des tableaux

Liste des algorithmes

Liste des sigles et acronymes

ADN	: Acide Désoxyribonucléique
ECG	: Electrocardiogramme
LBPH	: Local Binary Patterns Histogram
AUC	: Area Under ROC Curve
IRM	: Imagerie par Résonance Magnétique
FAR	: False Acceptance Rate ou Taux de Fausses Acceptations
FRR	: False Rejection Rate ou Taux de Faux Rejets
EMG	: Electromyogramme
FTA	: Failure To Acquire rate ou Taux d'échec à l'acquisition
FTE	: Failure To Enroll rate ou Taux d'échec à l'enrôlement
FNMR	: False Non-Matche Rate ou Taux de fausse non-correspondance
FMR	: False Match Rate ou Taux de fausse correspondance
ROC	: Receiver operating characteristic curve
EER	: Equal Error Rate ou taux d'égale erreur
SVC	: Signature Verification Competition
FVC	: Fingerprint Verification Competition
SPRT	: Sequential Probability Ratio Test ou test séquentiel du rapport de vraissemblance
RGB	: Espace colorimétrique Red, Green, Blue
HSV	: Espace colorimétrique Hue, Saturation, Value
CNN	: Convolutional Neural Network ou réseaux de neurones convolutionnels

Chapitre 1 : Introduction générale

Sommaire

Introduction

Il existe deux méthodes traditionnelles pour vérifier l'identité d'un individu. La première méthode est basée sur la connaissance. Cette connaissance correspond, par exemple, au mot de passe utilisé pour accéder à une application, le code PIN (Personal Identification Number) permettant de déverrouiller un Smartphone. La seconde méthode est basée sur la possession. Il peut s'agir d'une carte d'identité, d'une clé, d'un badge, etc. [1,2]. Ces deux modes d'authentification peuvent être utilisés de manière complémentaire pour obtenir une sécurité accrue (par exemple une carte de crédit). Cependant, ils ont leurs faiblesses respectives. Dans le premier cas, le mot de passe peut être oublié de son utilisateur ou deviné par une autre personne. Dans le second cas, le badge (ou la carte d'identité ou la clé) peut être perdu ou volé. La biométrie est une alternative aux deux modes de vérification précédents. Elle consiste à déterminer ou à vérifier l'identité d'un individu en fonction de ses caractéristiques physiologiques ou comportementales [3,4]. Ainsi, la reconnaissance biométrique est basée sur ce que nous sommes ou sur notre comportement [1,5]. Une description détaillée des limites des systèmes d'authentification traditionnels a été présentée par Djara et al. [6].

1.1. Problématique

Dans le but de faire face aux limites des systèmes biométriques mono modaux (c’est-à-dire basés sur une seule modalité) tels que le problème d’universalité, les limites en termes de performance et de robustesse, etc., les systèmes multi modaux

(qui font appel à l'utilisation combinée de plusieurs modalités) ont été développés. Ainsi, plusieurs techniques de fusion de scores provenant de chaque modalité ont été mises au point. Par ordre de performance croissante, nous avons les techniques de fusion en parallèle, en série puis la fusion séquentielle. Au niveau de l'architecture de fusion séquentielle, il se pose le problème de la non-prise en compte des métadonnées dans le processus de fusion. En clair, l'architecture séquentielle fonctionne exclusivement avec des modalités de biométrie pure (basée sur les caractéristiques physiologiques ou comportementales), sans pouvoir y intégrer les métadonnées. Il n'existe donc pas d'architecture de fusion de scores biométriques permettant de combiner des modalités de biométrie pure et des métadonnées dans une approche séquentielle.

1.2. Objectifs

A partir de la problématique décrite dans la section précédente, nous nous fixons comme objectif dans un premier temps de proposer une nouvelle architecture de fusion de scores qui va améliorer l'architecture séquentielle à travers la prise en compte de métadonnées dans le processus de fusion. Ensuite nous procéderons à une évaluation équitable des deux architectures à travers une implémentation des deux algorithmes correspondants sur les mêmes bases de données. Les valeurs des différentes métriques universelles permettront de connaître le niveau de performance de chaque architecture.

1.3. Organisation

Les travaux effectués dans le cadre de cette thèse ont été structurés en trois parties. La première partie aborde les généralités ainsi que l'état de l'art sur les systèmes biométriques développés à ce jour. De façon plus détaillée, cette partie introductive présente au premier chapitre la définition de la biométrie, son origine, ses différentes formes ainsi que les cas d'application possibles. Le lien entre la biométrie et l'émotion y a été également élucidé. Dans le deuxième chapitre, il a été question de l'évaluation des performances des systèmes biométriques. Les

taux d'erreur caractéristiques, les courbes de performance et les points de fonctionnement ont été présentés. Il en est de même des bases de données et protocoles d'évaluation ainsi que des compétitions internationales qui permettent de jauger les systèmes mis au point par la communauté scientifique.

La deuxième partie traite des méthodes originales développées. Dans le troisième chapitre, notre nouvelle architecture de fusion séquentielle adaptée a été présentée. Elle tire sa source d'une part de la méthode d'introduction des métadonnées dans les systèmes multibiométriques et d'autre part de la stratégie de fusion séquentielle de scores. Ensuite nous avons présenté au quatrième chapitre une analyse des causes profondes des vulnérabilités des systèmes biométriques. Partant du mode opératoire de la sécurité informatique, nous avons analysé les métadonnées et produit une typologie de l'adaptation en biométrie. Un accent particulier a été mis sur le rôle que peuvent jouer les métadonnées face aux défis et vulnérabilités biométriques. Au terme de cette analyse, une typologie de l'action des métadonnées sur les vulnérabilités biométriques a été élaborée.

Quant à la troisième partie, elle est relative à l'application des méthodes mises au point ainsi qu'aux résultats obtenus. Ainsi, dans le chapitre 5 nous avons mis en application le scénario d'utilisation de la couleur de peau dans le but d'améliorer les performances de la reconnaissance faciale. Les techniques de détection du visage puis celles de la peau ont été élucidées avant d'aborder la prédiction de l'identité. Les résultats obtenus ont été présentés et discutés. Sur la base de l'ensemble des travaux précédents, le sixième et dernier chapitre a été consacré à la mise en œuvre de la méthode de fusion séquentielle adaptée que nous avons proposée.

Conclusion

Pour les besoins de l'implémentation, nous avons retenu le visage et l'empreinte digitale sans contact comme modalités de biométrie pure tandis que la couleur de

peau du visage a été retenue comme métadonnée. Les scores de comparaison de ces trois modalités ont été fusionnés suivant l'architecture séquentielle d'une part et l'architecture séquentielle adaptée d'autre part. Le temps d'exécution ainsi que la proportion des utilisateurs reconnus à la première étape (étape du visage) ont servi de base à la comparaison des deux architectures. Il ressort que notre nouvelle architecture est plus performante que celle séquentielle.

Première partie : Etat de l’art sur les systèmes biométriques

Chapitre 2 : Etat de l’art sur la biométrie multimodale

Sommaire

Introduction

Etymologiquement, le mot biométrie signifie "mesure + vie" ou "mesure de la vie" et désigne dans un sens très large l’étude quantitative des êtres vivants. Le terme “biométrie” vient des mots grecs bios (vie) et metrikos (mesure) [2]. Les auteurs [7] affirment que la biométrie est à l'origine grecque, “bios” et “métron”, ce qui signifie littéralement « mesure de la vie ». L'émergence de la biométrie date du 19ème siècle [8]. Ce travail vise à présenter une revue de la multibiométrie qui prend en compte des informations auxiliaires (taille, couleur de peau, etc.), en plus des données de biométrie pure (empreinte digitale, visage, etc.). Cette technique consiste à introduire, dans le processus de reconnaissance biométrique, des informations spécifiques à l'utilisateur ou au contexte d'exploitation. La première partie du travail présente les différents systèmes biométriques développés à ce jour, leur architecture et leurs caractéristiques. Nous avons ensuite abordé la multibiométrie à travers ses avantages, sa diversité et les différents niveaux de fusion. Une attention particulière a été portée à la fusion des scores avant

d'aborder la prise en compte des informations auxiliaires en multibiométrie. Nous avons également présenté l'influence de l'émotion sur la performance des systèmes biométriques.

2.1. Panorama des différentes modalités biométriques

Les modalités biométriques peuvent être physiologiques (visage, thermogramme facial, empreinte digitale, ADN (Acide Désoxyribonucléique), rétine, iris, réseau veineux des mains, géométrie de la main, odeur, forme de l'oreille, électrocardiogramme -ECG-, etc.) ou comportementales (voix, signature, démarche, frappe) [3,4,9]. Les modalités physiologiques peuvent être divisées en trois types, à savoir morphologiques (visage, thermogramme facial, empreinte digitale, rétine, iris, réseau veineux de la main, géométrie de la main, forme de l'oreille, etc.), biochimiques (par exemple, ADN, odeur) [6,10] et bioélectriques (électrocardiogramme, électromyogramme, IRM, rayons X) [11]. Certaines modalités sont considérées comme étant au milieu de la morphologie et du comportement. C'est par exemple le cas de la voix [2,12]. Certaines modalités biométriques communes sont illustrées à la figure 2.1. Les modalités morphologiques sont liées à des traits physiques stables et immuables. Elles offrent davantage de garanties de sécurité mais sont généralement mal acceptées par les utilisateurs. D'autre part, les modalités comportementales sont susceptibles de varier en fonction de la volonté ou non de l'individu. En effet, un individu peut décider de changer volontairement de démarche ou de signature, tout comme en cas d'émotion, la voix d'un individu peut changer contre sa volonté. De plus, les modalités comportementales sont plus susceptibles à la fraude que celles physiologiques [12]. Les modalités biochimiques sont reconnues comme les plus fiables en termes d'identification et d'authentification. Cependant, elles sont difficiles à collecter (car elles nécessitent beaucoup de temps et d'équipements spécifiques) et sont détestées par les utilisateurs. Quant aux modalités bioélectriques, elles ont un niveau de précision élevé. Par contre, leur utilisation

n’est pas facile et leur niveau d’acceptation par les utilisateurs est bas. Ces modalités morphologiques, comportementales, biochimiques et bioélectriques identifiant un individu sont aussi appelées indicateurs ou identificateurs [13].

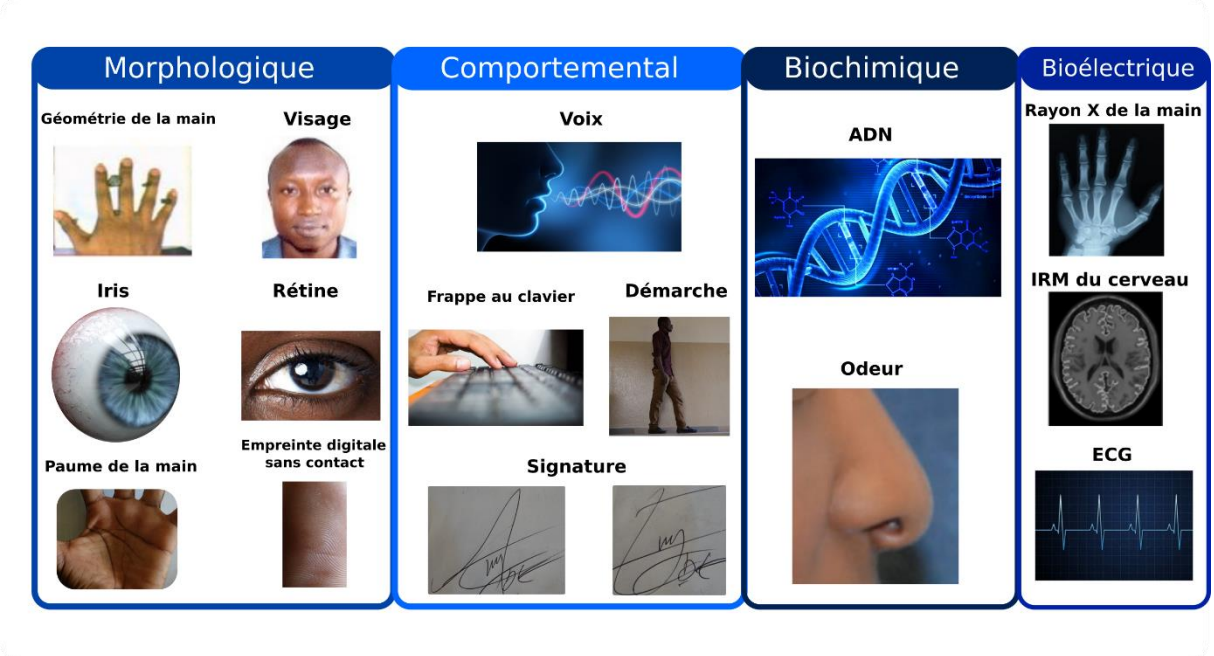

Figure 2.1 : Quelques modalités biométriques courantes [6]

2.1.1. Modalités morphologiques

2.1.1.1. Visage

Naturellement, la première modalité utilisée par les individus pour s'identifier est le visage [12,2]. Ainsi, les techniques de reconnaissance automatique des visages se sont développées et sont parmi les plus utilisées en biométrie aujourd'hui [14]. En 2008, Guerfi [15] a mené une étude sur l'authentification des visages. Elle affirme que la reconnaissance automatique du visage s'effectue en trois étapes principales : (1) détection du visage, (2) extraction et normalisation des traits du visage, (3) identification ou vérification. L'extraction de caractéristiques telles que les yeux, le nez, la bouche est une étape préalable au traitement nécessaire à la reconnaissance faciale [16]. On peut distinguer deux principales pratiques. La première repose sur l'extraction de régions entières du visage, elle est souvent mise en œuvre avec une approche globale de la reconnaissance des visages. La deuxième pratique extrait des points particuliers de différentes zones

caractéristiques du visage, telles que les coins des yeux, la bouche et le nez (voir figure 2.2). Elle est utilisée avec une méthode de reconnaissance locale et également pour l'estimation de la pose du visage [2,17].

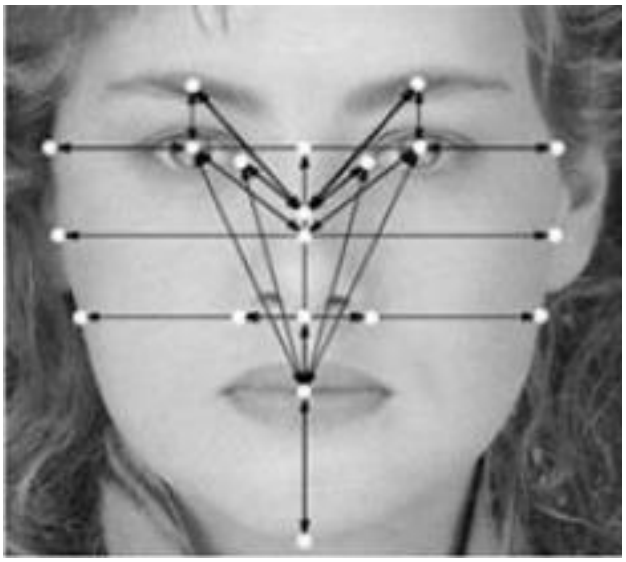

Figure 2.2 : Visage humain avec quelques exemples de caractéristiques recherchées [18]

2.1.1.2. Empreinte digitale

La reconnaissance basée sur les empreintes digitales est la méthode d'identification de personnes la plus ancienne, la plus efficace et la plus utilisée en raison de son caractère unique et de sa cohérence dans le temps, raison pour laquelle elle est utilisée depuis plus d'un siècle [16,18]. Les empreintes digitales consistent en un motif de texture régulière composé de crêtes et de vallées. Ces crêtes sont caractérisées par plusieurs points de repère, appelés minuties, qui se présentent principalement sous la forme de terminaisons de crêtes et de bifurcations de crêtes. On prétend que la distribution spatiale de ces points de minutie est unique pour chaque doigt ; c'est la collection de minuties d'une empreinte digitale qui est principalement utilisée pour faire correspondre deux empreintes digitales [16]. Les techniques d'appariement d'empreintes digitales peuvent être classées en trois catégories, à savoir l'appariement basé sur les

minuties, l'appariement basé sur l'image et la technique d'appariement hybride [19]. L'appariement basé sur les minutes consiste essentiellement à rechercher l'alignement entre le modèle et les jeux de caractéristiques de minuties d'entrée donnant le nombre maximal d'appariements de minuties. Le système de correspondance basé sur les images est utilisé pour comparer deux images d'empreintes digitales en les superposant. Le motif global des crêtes et des creux d'une empreinte est le principal objectif de cet algorithme, où les informations de niveau de gris de l'image sont directement utilisées [20]. Historiquement, l'acquisition d'images d'empreintes digitales était réalisée à l'aide de la technique dite de l'encre. Elle consiste à mouiller le doigt du sujet avec de l'encre noire. Ce doigt est ensuite pressé contre une carte de papier ; pour finir, la carte est numérisée à l'aide d'un scanner papier ordinaire, produisant l'image numérique finale (hors ligne). De nos jours, la plupart des applications utilisent des images numériques en direct captées en détectant directement la surface du doigt avec un scanner d'empreintes digitales (en ligne) [13]. Mais cette méthode d'acquisition de l'image pose plusieurs problèmes. Une pression sur le capteur peut provoquer une distorsion des minuties. Le contact avec le capteur pose également des problèmes d'hygiène (propagation d'épidémies) et de sécurité (risque d'attaques chimiques). Pour toutes ces raisons, les systèmes d'acquisition d'empreintes digitales 2D sans contact [19] et 3D [21] sans contact ont connu un développement dans les années 2000.

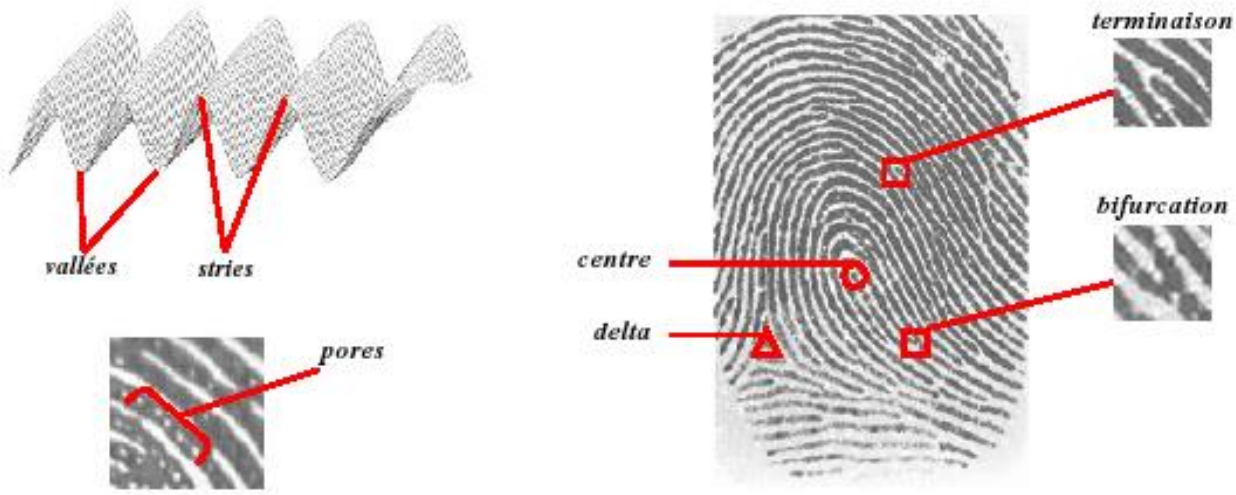

Figure 2.3 : Caractéristiques d'une empreinte digitale [22]

2.1.1.3. Géométrie de la main et des doigts

La main humaine présente des caractéristiques relativement stables (par exemple, la longueur des doigts) qui sont particulières (bien que peu distinctives) à un individu. Elles peuvent donc être utilisées à des fins de reconnaissance biométrique. La "mesure" de la main est composée de plusieurs mesures telles que les dimensions des doigts, les caractéristiques des articulations, la paume et la forme de la main. La première étape de l'authentification est celle de la numérisation. Pour ce faire, la personne doit poser sa main sur une platine. Les doigts doivent être correctement placés. Une caméra infrarouge prend ensuite une image sous deux angles différents pour obtenir une reproduction en trois dimensions de la main. La lecture ne prend normalement pas plus de trois secondes. Les systèmes de géométrie des doigts mesurent la géométrie d'au plus deux doigts par opposition à la main entière. Ils peuvent être préférés en raison de leur taille compacte [14,16,23]. Il est à noter que les membres d'une même famille présentant des similitudes physiques importantes peuvent facilement tromper ces systèmes. En outre, la forme des mains est susceptible de changer en raison de maladies liées au vieillissement, telles que l'arthrite. De plus, la géométrie de la main nécessite un scanner plus grand qu'un lecteur d'empreintes digitales. En

conséquence, son utilisation devient gênante lorsqu'il s'agit de sécuriser un petit objet tel qu'un ordinateur.

2.1.1.4. Iris

La technologie de detection de l'Iris est l'un des systèmes de reconnaissance biométrique les plus fiables mis au point récemment. La reconnaissance de l'iris est une technique biométrique permettant de reconnaître un individu par l'observation de son iris. En effet, la texture de l'iris (c'est-à-dire le motif de l'iris) comprend de nombreuses caractéristiques. Les éléments les plus souvent utilisés en biométrie sont les rayures, les creux et les sillons. Ces éléments de l'iris restent fixes, ils ne varient que très peu au cours de la vie. Chaque motif déterminé par les processus chaotiques au cours du développement embryonnaire est stable et unique (la probabilité de similarité est de 1^{-72}) [14,18]. De plus, le motif de l'iris n'est pas génétique contrairement à la couleur des yeux. Ainsi, deux individus, même s'ils sont parents, peuvent avoir la même couleur d'iris mais jamais le même motif. C'est pourquoi l'iris permet de distinguer les jumeaux identiques. Les systèmes de reconnaissance basés sur l'iris sont désormais moins coûteux et non intrusifs. Cependant, il est important de noter que les systèmes à iris ont un taux de faux rejets (False Rejection Rate ou FRR) élevé [16,24].

2.1.1.5. Rétine

La rétine est la membrane sensorielle qui tapisse la surface interne du dos du globe oculaire. Elle est composée de plusieurs couches, dont une contenant des cellules spécialisées appelées photorécepteurs [25]. Les difficultés rencontrées pour capturer l'image d'une rétine sont autant psychologiques que médicales et techniques. Pour obtenir une image de la rétine, il est nécessaire d'éclairer l'arrière du globe oculaire au moyen d'un faisceau lumineux. Ce faisceau est de très faible intensité pour ne pas gêner l'utilisateur ; il est inoffensif et d'intensité inférieure à celle des dispositifs ophtalmiques. Un système de caméra très précis vient alors récupérer l'image de la rétine. Les capteurs de rétine sont disponibles

et offrent un très haut niveau de sécurité. Après avoir capturé une image de la rétine, le logiciel du dispositif de lecture coupe un anneau autour de la fovéa. Dans cet anneau, il localise l'emplacement des veines et leur orientation. Il crée ensuite une "signature oculaire" utilisée pour la reconnaissance rétinienne. L'opération elle-même est assez simple à décrire mais les algorithmes restent relativement complexes.

2.1.1.6. Paume de la main

Plusieurs auteurs ont étudié la reconnaissance biométrique à base de paume de la main. Parmi ceux-ci, nous avons [2,13,18,24]. C'est une biométrie relativement nouvelle, comparée à d'autres systèmes biométriques tels que les systèmes de reconnaissance de visage, d'empreintes digitales et d'iris. Cette biométrie s'intéresse à la surface interne d'une main. Une paume est recouverte du même type de peau que le bout des doigts et sa taille est supérieure à celle d'un bout de doigt. Cinq caractéristiques d'une empreinte de paume sont généralement utilisées pour identifier une personne de manière unique. Ce sont : (a) les caractéristiques géométriques (telles que la largeur, la longueur et la surface) ; (b) les caractéristiques des lignes principales (emplacement et forme des lignes principales); (c) les traits de rides (les lignes plus minces et plus irrégulières); (d) les caractéristiques du point delta (centre d'une région de type delta de l'empreinte palmaire, généralement située dans la région de la racine du doigt) et (e) les caractéristiques des minuties (voir empreinte digitale). Les systèmes de reconnaissance palmaire utilisent un dispositif de numérisation ou une application basée sur une caméra (sans contact), ainsi qu'un logiciel associé qui traite les données d'image d'une photographie de la paume de la personne et les compare à un enregistrement stocké pour cette personne.

D'autres traits morphologiques sont souvent utilisés dans la reconnaissance biométrique. Ces systèmes comprennent le réseau veineux de la paume, l'oreille

et le thermogramme facial. Les références [18,9] fournissent de plus amples informations sur ces autres modalités biométriques liées à la morphologie.

2.1.2. Modalités comportementales

2.1.2.1. Voix

L'identification vocale est considérée par les utilisateurs comme l'une des formes les plus normales de technologie biométrique, car elle n'est pas intrusive et ne nécessite aucun contact physique avec le capteur. La technologie d'analyse vocale (également appelée analyse de locuteur) est appliquée avec succès lorsque d'autres technologies sont difficiles à utiliser. Elle est utilisée dans des domaines tels que les centres d'appels, les banques, l'accès aux comptes, sur les ordinateurs domestiques, pour accéder à un réseau ou pour des applications légales. L'analyse de la voix n'est pas seulement liée à des caractéristiques personnelles, mais également à de nombreuses variables environnementales et sociolinguistiques, la génération de voix étant le résultat d'un processus extrêmement complexe. Ainsi, la voix transmise intégrera une version dégradée des spécificités du locuteur et sera influencée par de nombreuses variables contextuelles difficiles à gérer [2,12,13]. La plupart des systèmes d'identification vocale utilisent l'affichage d'un texte ; certains mots doivent être lus, puis prononcés pour vérifier que la personne à authentifier est présente et qu'il ne s'agit pas d'une question d'enregistrement. Le fait qu'une voix authentique puisse être enregistrée et utilisée par un imposteur pour une identification non autorisée constitue un inconvénient majeur du système de reconnaissance vocale. Malgré toutes ces difficultés, la voix reste un moyen biométrique intéressant à exploiter car pratique et disponible via le réseau téléphonique, contrairement aux autres modalités biométriques [14,18].

2.1.2.2. Signature

Une signature est un nom manuscrit ou un surnom, un dessin ou toute marque stylisée de manière à être propre à un individu. En fonction des informations de signature introduites, nous avons deux classes de méthodes de vérification de

signature : les méthodes “en ligne” et “hors ligne”. Le procédé en ligne fait référence à l'utilisation des fonctions temporelles du processus de signature dynamique (par exemple, les trajectoires de position ou la pression en fonction du temps), qui sont obtenues à l'aide de dispositifs d'acquisition tels que des écrans tactiles ou des tablettes à numériser. La méthode hors ligne fait référence à l'utilisation de l'image statique de la signature. Un avantage de la signature manuscrite est qu’elle peut être facilement acquise avec un stylo encreur sur une feuille de papier ou par des moyens électroniques en utilisant des dispositifs à pointeur existants, tels que des tablettes à stylet, des tablettes PC, des écrans tactiles, etc. Malgré les avantages de la modalité de signature manuscrite, son déploiement pratique est très lent et la biométrie de signature représente un important défi en matière de recherche. Cela est principalement dû aux grandes variations intra-classe et, en considérant les contrefaçons, également aux petites variations inter-classe. Parmi les autres défis de la biométrie de signature, on peut citer la faible universalité, la faible permanence et la vulnérabilité aux attaques par usurpation directe [2,13].

2.1.2.3. Démarche

La démarche fait référence à la façon dont une personne marche, et est l'un des rares traits biométriques qui peuvent être utilisés pour reconnaître des personnes à distance. Par conséquent, ce trait est très approprié dans les scénarios de surveillance où l'identité d'un individu peut être établie subrepticement [2]. Les systèmes basés sur la démarche offrent également la possibilité de suivre une personne sur une longue période. Cependant, la démarche d'un individu est influencée par plusieurs facteurs, notamment le choix de la chaussure, la nature du vêtement, l'affliction des jambes, la surface de marche, etc. [24]. La démarche d'un individu peut être facilement connue dans les lieux publics grâce à une instrumentation simple et ne nécessite pas l'attention active de l'individu, ce qui constitue un avantage par rapport aux autres systèmes biométriques. Le cadre du

système comprend la détection du sujet, l'extraction de la silhouette, l'extraction de caractéristiques, la sélection de caractéristiques et la classification. Les trajectoires de certaines parties du corps humain telles que les pieds et la tête ont ainsi été normalisées pour apparaître tels qu'ils sont vus d'en-face et d'un point de vue parallèle [18].

2.1.2.4. Dynamique de frappe au clavier

La dynamique de frappe au clavier correspond à la manière dont chaque individu manipule le clavier selon une forme particulière, et on pense que cette biométrie peut fournir suffisamment d'informations pour faciliter l'identification et la vérification. Cependant, il n'est pas perçu comme un trait unique chez chaque individu. C'est une solution biométrique "Software Only", c'est-à-dire uniquement logicielle. Elle est appliquée au mot de passe, qui devient beaucoup plus difficile à "imiter". Lors de la mise en œuvre de cette technique, l'utilisateur est invité à entrer son mot de passe une dizaine de fois à la suite. En utilisant un algorithme qui exploite le temps de support sur chaque touche et le temps entre les différentes touches, la "moyenne" des dix entrées est générée pour créer un "profil de frappe" de l'utilisateur qui servira de référence. Aux accès suivants, l'entrée du mot de passe sera couplée à un profil de saisie qui sera comparé au profil de référence. Le droit d'accès est alors accordé en fonction du niveau de similarité de ce profil avec la référence. Comme inconvénient majeur, vous devez être en bonne condition avant d'utiliser le système, au risque de voir votre propre mot de passe refusé [2,18].

2.1.3. Modalités biochimiques

2.1.3.1. ADN

L'acide désoxyribonucléique (ADN) est une chaîne de nucléotides contenus dans le noyau de nos cellules. Il peut être utilisé comme un outil biométrique très stable pour classifier et guider l'identification d'individus inconnus à partir d'échantillons biologiques qu'ils ont laissés. Jain et Kumar [16] affirment que la

structure de l'ADN de chaque être humain est unique, à l'exception des vrais jumeaux, et est composée de gènes qui déterminent des caractéristiques physiques telles que la couleur des yeux ou des cheveux. Les profils peuvent être générés à partir d'un prélèvement buccal, de dépôts ou de cellules biologiques, généralement de taches de sang, de salive, d'urine ou de sperme, de poils (avec racines) et de cellules cutanées (laissées par simple contact, par exemple une marque de doigt). La correspondance d'ADN est très populaire pour les applications de police scientifique et d'application de la loi. Cependant, cela nécessite des échantillons concrets et ne peut pas encore être fait en temps réel. Actuellement, toutes les étapes de la comparaison de l'ADN ne sont pas automatisées et les résultats peuvent donc être faussés si le processus n'est pas correctement conduit ou si les échantillons d'ADN sont eux-mêmes contaminés.

2.1.3.2. Odeur

L'odeur des individus contient des compositions chimiques uniques pouvant être utilisées pour l'identification [26]. L'odeur peut être détectée en utilisant un "capteur chimique" tel que ceux basés sur la technologie à oxyde métallique. L'odeur humaine a l'avantage d'être impossible à reproduire [14,24]. En 2014, Inbavali et Nandhini [27] ont créé un système qui authentifie les personnes en fonction de leur odeur corporelle. Leurs résultats expérimentaux montrent que cet identifiant biométrique présente le taux d'erreur le plus faible (15%) par rapport à d'autres identificateurs biométriques tels que l'iris, les empreintes digitales et la reconnaissance faciale. Il convient de noter que l'effet des odeurs ambiantes et auxiliaires sur l'odeur humaine est diversement apprécié par les auteurs. En effet, Oloyede et Hancke [18] affirment qu'il n'est pas certain que l'invariance de l'odeur corporelle puisse être détectée s'il existe la présence de déodorant ou de composition chimique du milieu environnant. D'autre part, les auteurs [27] concluent, après une étude, que même les déodorants et les parfums ne peuvent

pas masquer l'odeur humaine de base. Ces parfums artificiels n'éliminent pas les composés organiques présents dans l'odeur humaine.

2.1.4. Modalités bioélectriques (modalités cachées)

2.1.4.1. Electrocardiogramme

Un électrocardiogramme (ECG) est un test qui étudie le fonctionnement du cœur en mesurant son activité électrique. À chaque battement de cœur, une impulsion électrique (ou "onde") traverse le cœur. Cette onde provoque la contraction du muscle cardiaque de sorte qu'il expulse le sang du cœur [28]. L'ECG est principalement utilisé dans les applications cliniques pour diagnostiquer les maladies cardiovasculaires. Le signal ECG est caractérisé par la forme de ses battements composés de cinq ondes typiques, à savoir P, Q, R, S et T ou parfois l'onde U [29]. La biométrie ECG a fait l'objet d'un certain nombre d'études [30,31,32]. L'utilisation de l'ECG en biométrie est relativement nouvelle. En fait, il existe plusieurs méthodes biométriques basées sur l'ECG. Il existe des approches basées sur l'analyse ECG [33]. D'autres méthodes s'appuient sur l'intégration de fonctionnalités analytiques et d'apparence extraites des signaux ECG [34]. Comparé à d'autres modalités biométriques, l'ECG est plus universel et difficile à forger [35].

2.1.4.2. Electromyogramme

Les signaux d'électromyogramme (EMG) sont des signaux bioélectriques enregistrés dans les muscles. Leur utilisation comme modalité biométrique cachée peut être particulièrement intéressante. Dans ce contexte, quelques expériences récentes ont été réalisées [34,36,37]. En particulier, ce travail s'est concentré sur l'analyse des signaux d'électromyographie de surface (SEMG) [38]. Lors de l'acquisition de ces signaux, les individus sont encouragés à appliquer une pression manuelle d'une intensité constante sur une sonde de force pendant plusieurs secondes [29]. Le signal ainsi obtenu est analysé dans le domaine spectral. Ensuite, des paramètres sont extraits tels que la force du signal, la

fréquence moyenne, le coefficient d'aplatissement et le coefficient d'asymétrie. En effet, ces paramètres fournissent un vecteur périphérique qui peut être utilisé pour caractériser les individus.

2.1.4.3. IRM du cerveau

L'imagerie par résonance magnétique (IRM) est l'une des techniques d'imagerie médicale les plus récentes. Elle permet de visualiser avec une grande précision les organes et tissus mous, dans différents plans de l'espace. Il est ainsi possible de déterminer la position exacte de lésions autrement invisibles. Réalisé sous la direction d'un médecin radiologue, cet examen ne provoque aucune irradiation. Il ne fait appel en effet, qu'aux propriétés des champs magnétiques. Contrairement au scanner ou à la radiographie l'IRM n'utilise pas les rayons X. Fiable et précise, elle permet d'obtenir des vues en deux et trois dimensions de l'intérieur du corps. Elle s'avère ainsi particulièrement utile pour détecter des maladies ou troubles internes que des examens classiques n'ont pas réussi à identifier. Les images obtenues par résonance magnétique sont le résultat de l'interaction entre un champ magnétique créé par la machine et les atomes d'hydrogène contenus dans l'organisme du patient. L'appareil est en effet équipé d'un aimant très puissant qui permet d'envoyer des ondes qui vont faire vibrer les noyaux d'hydrogène contenus dans les tissus organiques. Le retour de ces noyaux d'hydrogène à leur état d'équilibre va engendrer la formation d'un signal dans une antenne réceptrice. Ce sont ces modifications qui vont être utilisées pour créer les images sur l'écran [39,40]. On peut aussi définir ce qu'on appelle brain code ou code du cerveau à travers une segmentation de la zone d'intérêt du cerveau [41]. L'avantage principal de ce type de modalité cachée est le fait que le cerveau est totalement protégé contre toutes sortes de changements. Cependant, l'inconvénient principal de cette modalité est la non-disponibilité de systèmes d'IRM robuste consacrés à la biométrie.

2.1.4.4. Le rayon X

Les rayons X sont des ondes électromagnétiques de hautes fréquences. Ils possèdent une forte énergie et pénètrent facilement la matière. C'est une forme de lumière que l'œil ne peut pas percevoir et qui peut passer à travers les objets et les êtres vivants. Les médecins les utilisent en imagerie médicale dans deux domaines, la radiographie et le scanner pour ainsi pouvoir visualiser l'intérieur du corps et diagnostiquer des maladies qu'on ne pourrait pas découvrir autrement.
L'application de ce type de technologie dans la biométrie est envisageable en exploitant des images radiographiques de la main par exemple où le but est de caractériser les phalanges à l'aide de quelques outils de traitement d'images [42].

Le scanner, également appelé tomodensitomètre ou scanographe, est un appareil qui « balaie » la région à explorer de façon à réaliser des images en coupes fines (ou « en tranches ») de l'organisme. Il permet ainsi de déterminer très précisément la localisation et l'étendue d'une lésion sur un organe ou un tissu. Le scanner combine des rayons X (comme la radiographie) à des supports informatiques, ce qui permet de reconstruire des images dans les différentes dimensions : on peut ainsi obtenir des vues en coupe longitudinale ou horizontale. Les indications du scanner sont larges, dans la mesure où cet examen permet d'explorer la plupart des organes. Il est d'ailleurs souvent prescrit pour préciser les résultats d'une radiographie ou d'une échographie. Contrairement à la radiographie qui cible plus précisément les os, le scanner permet de visualiser la plupart des organes, et ce même au travers des os.

Divers systèmes utilisant les rayons X sont déjà utilisés pour la surveillance aux frontières et dans les aéroports, sur les objets et les véhicules. D'autres sont en test ou à l'étude concernant l'humain.

2.2. Propriétés des modalités biométriques

Pour être pratique et fiable, un système biométrique devrait répondre à certaines

propriétés spécifiques [2,7,9,10]. Dans ce travail, nous mettons l'accent sur huit propriétés souhaitables pour une modalité biométrique. Le tableau 2.1 présente les performances des différents systèmes de reconnaissance biométrique.

- **Universalité ou disponibilité**: signifie que toute la population doit posséder cette modalité;
- **Caractère distinctif ou unique**: signifie que deux personnes doivent avoir suffisamment de caractéristiques différentes;
- **Permanence ou stabilité**: la caractéristique biométrique doit être invariante dans le temps;
- **Possibilité de collecte**: signifie que la caractéristique biométrique doit être mesurable avec un dispositif de détection (pratique);
- **Précision**: La précision biométrique est basée sur plusieurs critères de vérification, notamment le taux d'identification, les taux d'erreur et les normes supplémentaires du système biométrique;
- **Acceptabilité**: signifie qu'un utilisateur et le public en général ne devraient avoir aucune objection (forte) à la mesure / collecte du trait biométrique;
- **Résistance au contournement**: désigne le degré de difficulté requis pour vaincre ou contourner le système;
- **Coût**: Il s'agit des charges générées par la mise en œuvre du système.

Tableau 2.1: Performances des différents systèmes de reconnaissance biométrique (adapté de [18,9,7,43,29])

Légende : U: Universalité; D: Caractère distinctif; P: Permanence; T: Possibilité de collecte; A: Précision; E: Acceptabilité; V: Résistance au contournement; C: Coût; H: Haut; M: Moyen; L: Bas

Modalités biométriques	***U***	***D***	***P***	***T***	***A***	***E***	***V***	***C***
Modalités Morphologiques								
Visage	H	L	M	H	L	H	H	M
Empreinte digitale	M	H	H	M	H	M	M	L

Modalités biométriques	***U***	***D***	***P***	***T***	***A***	***E***	***V***	***C***
Géométrie de la main	M	M	M	H	M	M	M	M
Iris	H	H	H	M	H	L	L	H
Rétine	H	H	M	L	H	L	L	H
Paume de la main	M	H	H	M	H	M	M	H
Thermogramme facial	H	H	L	H	M	H	L	H
Oreille	M	M	H	M	H	H	M	H
Réseau veineux de la paume	M	M	M	M	H	M	L	H
Modalités comportementales								
Voix	M	L	L	M	M	H	H	H
Signature	L	L	L	H	M	H	H	L
Démarche	M	L	L	H	H	H	M	H
Dynamique de frappe au clavier	L	L	L	M	L	M	M	L
Modalités biochimiques								
ADN	H	H	H	L	H	L	L	H
Odeur	H	H	H	L	L	M	L	H
Modalités bioélectriques (Thèse Hafs)								
Electrocardiogramme	H	H	H	L	H	L	H	H
Electromyogramme	H	H	H	L	H	L	H	H
IRM	H	H	H	L	H	L	H	H
Rayon X	H	H	H	L	H	L	H	H

2.3. Architecture d'un système biométrique

L'architecture de base d'un système biométrique intègre principalement cinq modules tels qu'illustrés sur la figure 2.4 [2,7,13].

- **Module de capture**: Il consiste à capturer les données biométriques brutes afin d'en extraire une représentation numérique qui servira de modèle de référence.
- **Module de traitement du signal**: Il permet de réduire la représentation numérique extraite afin d'optimiser la quantité de données à stocker

pendant la phase d'enregistrement ou de faciliter le temps de traitement pendant les phases de vérification et d'identification.

- **Module de stockage**: Il est utilisé pour stocker les modèles biométriques des individus après traitement.
- **Module de correspondance (ou appariement)**: il sert à comparer les données brutes biométriques extraites à un ou plusieurs modèles biométriques précédemment stockés. Le module détermine donc le degré de similarité (ou de divergence) entre deux vecteurs biométriques.
- **Module de décision**: il est utilisé pour déterminer si l'indice de similarité renvoyé est suffisant pour déterminer l'identité d'un individu.

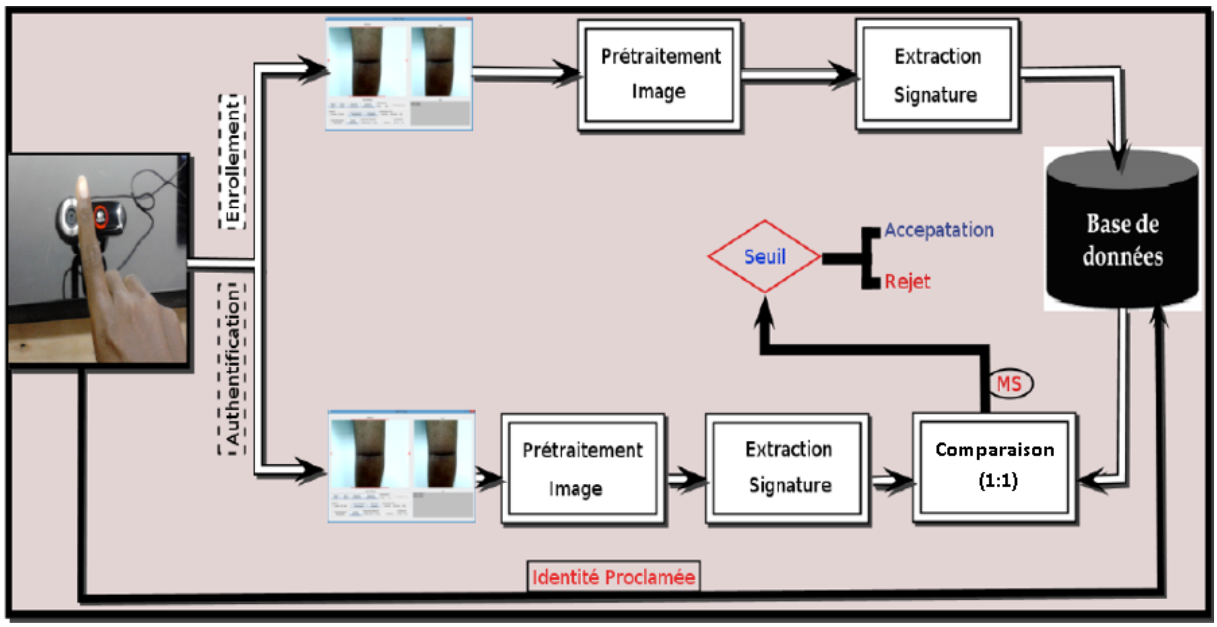

Figure 2.4 : Architecture générique d'un système biométrique [22]

2.4. Types de correspondance ou d'appariement

Il existe deux types d'appariement en biométrie: l'authentification (ou la vérification d'identité) et l'identification. Les deux types sont connus sous le nom de reconnaissance. Dans le cas de l'authentification, l'identité (partie non biométrique) d'une personne est proclamée et une comparaison est effectuée avec l'échantillon de la base de données correspondant à cette identité, à savoir une comparaison unique (un à un ou 1:1). Cependant, dans le cas de l'identification,

aucune identité n'est proclamée. Cela conduit à une comparaison entre l'échantillon biométrique prélevé et tous les échantillons de la base de données (un à plusieurs ou 1: n) [2,13,14]. De plus, dans le cas d'applications spécifiques telles que le contrôle d'accès, il n'y a pas de distinction claire entre authentification et identification, ce qui signifie que les techniques développées dans un scénario d'application peuvent être appliquées à un autre [44]. L'authentification consiste à répondre à la question: *êtes-vous celui que vous prétendez être?* Par contre, identifier revient à répondre à la question: *qui êtes-vous?* Dans les utilisations les plus courantes de la biométrie, le mode d'authentification est utilisé. L'identification est souvent utilisée dans les opérations de police scientifique telles que les enquêtes criminelles et les autopsies.

2.5. Mode de reconnaissance

Selon l'application considérée, un système biométrique peut fonctionner en mode de reconnaissance positive ou négative [43,45].

- **Reconnaissance positive**: Dans ce mode, le système détermine si la personne est bien celle qu'elle prétend être. L'objectif de la reconnaissance positive est d'empêcher de nombreuses personnes d'utiliser la même identité. C'est par exemple le cas où une seule personne est autorisée à accéder à une ressource sécurisée.
- **Reconnaissance négative**: Ce type de système est utilisé pour déterminer si la personne est celle qu'elle nie être. Dans ce cas, le but de la reconnaissance est d'empêcher une seule personne d'utiliser plusieurs identités. Cela correspond, par exemple, à une application de prestations sociales dans laquelle le système enregistre dans sa base de données les personnes ayant déjà reçu des prestations.

2.6. Les défis dans les systèmes biométriques

Les systèmes biométriques font face à trois défis principaux [7,12,18]. Il s'agit des limites en termes de performances, d'acceptabilité et d'architecture:

- **Limites de performances**: les caractéristiques humaines sont sujettes à variation et ces modifications ont un impact négatif sur les performances de reconnaissance biométrique. Les erreurs de vérification sont dues à de nombreuses raisons telles que les occlusions, les facteurs environnementaux (par exemple, les éclairages, le bruit) et l'adaptation croisée des dispositifs;
- **Limites d'acceptabilité**: Trois facteurs principaux contribuent à la complexité du système biométrique en termes d'acceptabilité: 1) l'exactitude en termes d'erreurs, 2) l'échelle ou la taille de la base de données et 3) l'utilisabilité en termes de facilité d'utilisation, de sécurité et de confidentialité;
- **Limites d'architecture**: Ces limites concernent huit points de vulnérabilités qui ont été identifiés dans un système biométrique générique [7]. Ces huits points se situent respectivement (1) au niveau du capteur, (2) entre le capteur et le module d'extraction des caractéristiques, (3) au niveau du module d'extraction des caractéristiques, (4) entre le module d'extraction des caractéristiques et le module de comparaison, (5) au niveau du module de comparaison, (6) au niveau de la base de données, (7) enre le module de comparaison et la base de données puis (8) au niveau du module de décision.

Des failles de sécurité ont été identifiées dans tous les systèmes biométriques. Toutes ces limites feront l'objet d'une analyse détaillée au niveau du chapitre 4, en lien avec le rôle des métadonnées.

2.7. Modes d'identification en biométrie

Il existe deux modes d'identification biométrique, à savoir l'identification d'ensemble ouvert et l'identification d'ensemble fermé. En mode ouvert, le sujet présenté au système biométrique n'est pas encore inscrit dans la base de données.

C'est le cas d'une identification de liste de surveillance à partir de caméras de surveillance qui implique un contrôle continu d'une liste de personnes à partir d'images vidéo. D'autre part, l'identification en mode fermé concerne la situation dans laquelle la personne concernée est susceptible d'être présente dans la base de données biométrique et dans ce cas, le système biométrique ne renvoie pas une liste vide de candidats [18].

2.8. Classification des systèmes biométriques

Il existe deux types de biométries, à savoir la biométrie unimodale (ou monomodale) et la multibiométrie.

2.8.1. La biométrie unimodale

Ces systèmes reposent sur la preuve d'une source d'information unique pour l'authentification (par exemple, une empreinte digitale unique, un visage). Ils font face à une variété de problèmes tels que : (i) le bruit dans les données numérisées; (ii) les variations intra-classe ; (iii) les similitudes inter-classes ; (iv) la non-universalité et (v) les attaques par usurpation d'identité [2].

2.8.2. Taxonomie de la multibiométrie

La multibiométrie désigne la fusion de différents types d'informations. Ce type de biométrie réduit les contraintes de la biométrie unimodale en combinant plusieurs systèmes. On distingue six types de systèmes multibiométriques en fonction des systèmes qu'ils combinent [6,9,24,46,47].

On parle des systèmes:

- *Multi-capteurs* lorsqu'ils associent plusieurs capteurs pour acquérir la même modalité, par exemple un capteur optique et un capteur capacitif pour acquérir l'empreinte digitale.
- *Multi-instances* lorsqu'ils associent plusieurs instances de la même biométrie, par exemple l'acquisition de plusieurs images de visage avec des modifications de la pose, de l'expression ou de l'éclairage.

- *Multi-algorithmes* lorsque plusieurs algorithmes traitent la même image acquise, cette multiplicité d'algorithmes peut intervenir dans le module d'extraction en considérant plusieurs ensembles de caractéristiques et / ou dans le module de comparaison utilisant plusieurs algorithmes de comparaison.
- *Multi-échantillons* lorsqu'ils combinent plusieurs échantillons différents de la même modalité, par exemple deux empreintes digitales différentes ou deux iris.
- *Multi-modal* lorsque plusieurs modalités sont prises en compte, par exemple le visage et les empreintes digitales.
- *Multi-origine* lorsque les modalités en présence proviennent d'origines différentes. C'est le cas par exemple de la combinaison de modalités de biométrie pure (géométrie de la main, voix, iris, rétine, etc.) et celles de biométrie douce (couleur de la peau, sexe, taille, poids, couleur des yeux, etc.) ou des métadonnées en général.

En dehors des six formes de multibiométrie pure, on peut citer les systèmes hybrides qui combinent différents types d'associations, par exemple l'utilisation du visage et de l'empreinte digitale, mais en utilisant plusieurs doigts. Il s'agit dans ce cas d'un système multi-modal et multi-échantillons.

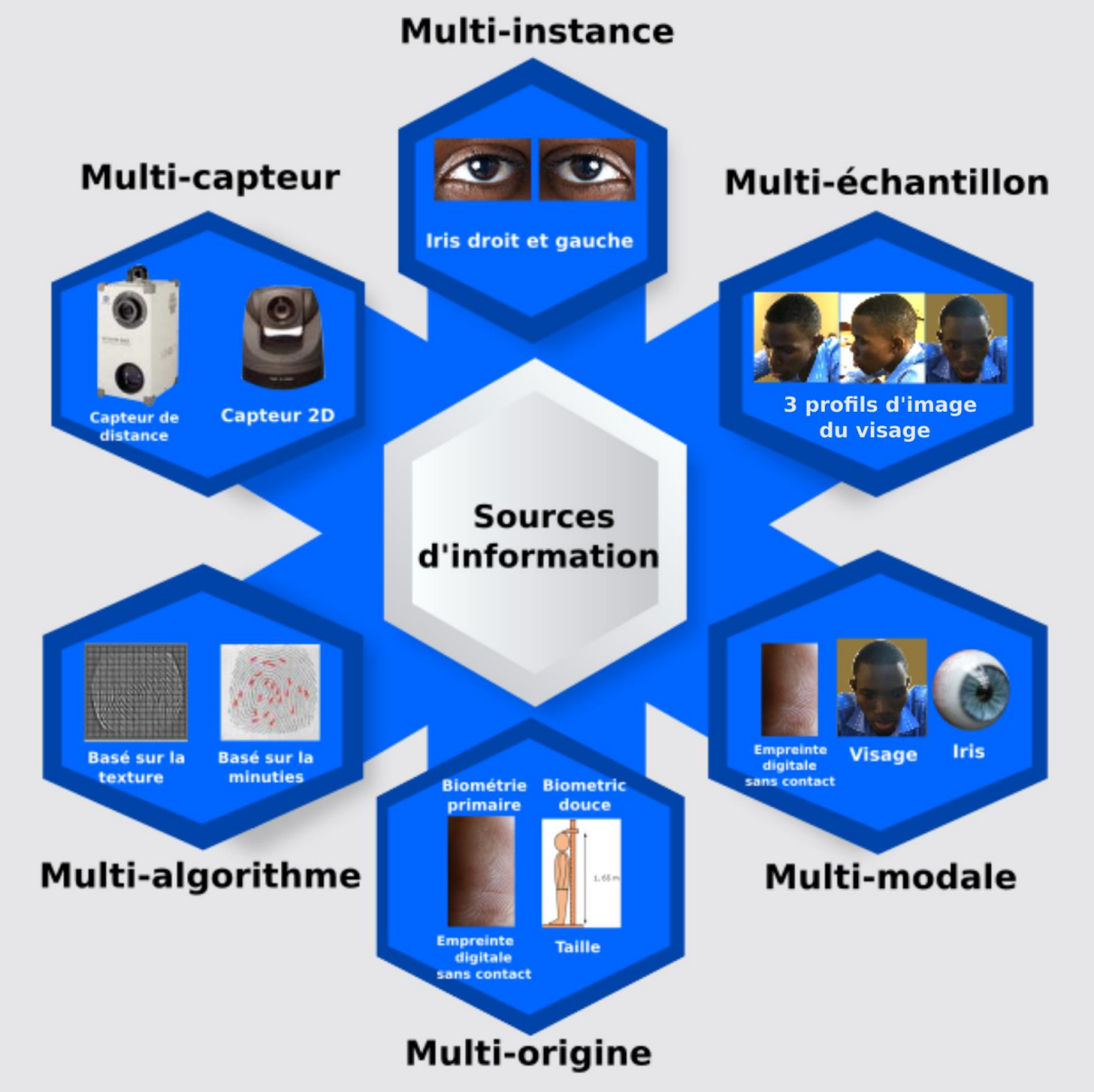

Figure 2.5 : Sources d'information pour la fusion en biométrie [6]

2.8.3. Niveaux de fusion

Selon Sanderson et Paliwal [48], la combinaison de plusieurs systèmes biométriques peut être réalisée à cinq différents niveaux: niveau des données brutes (ou capteur), niveau des caractéristiques, niveau des scores, niveau du classement et niveau de la décision. Ces cinq niveaux de fusion peuvent être classés en deux sous-ensembles: la fusion pré-classification (avant l'appariement) et la fusion post-classification (après l'appariement). La fusion post-classification

peut s'effectuer au niveau des scores issus des modules de correspondance, du classement [49] ou au niveau de la décision. Gad et al. [9] ont présenté en 2015 la fusion de type hybride qui vise à accélérer le système et à réduire considérablement le taux d'erreur.

2.8.4. Avantages, limites et solutions aux défis de la biométrie multimodale

2.8.4.1. Avantages

La biométrie multimodale permet de résoudre certains problèmes liés à la monomodalité, tels que les distinctions intra-classe, l'inflexibilité, le bruit, la non-universalité, les attaques par usurpation et les taux d'erreur élevés [50,18]. En termes de distinctions intra-classe, un système multimodal utilise plus d'un facteur biométrique et la fusion permet ainsi d'initialiser plus de points de données, ce qui permet une meilleure classification des points de données. La fusion de plusieurs facteurs offre également une flexibilité au système et empêche les données bruitées d'avoir des effets substantiels sur la décision.

2.8.4.2. Limites

L'implémentation d'un système biométrique multimodal nécessite de relever de nombreux défis au préalable. Kumar et Farik [50] ont identifié en 2016 les quatre défis suivants : (1) les systèmes multimodaux sont difficiles à concevoir ; (2) le niveau d'acceptation par l'utilisateur est assez faible ; (3) ils nécessitent un niveau d'investissement plus élevé et (4) le compromis sur la performance. La généralisation des systèmes multi-biométriques dépend en partie du niveau d'acceptation de l'utilisateur. Généralement, les utilisateurs n'aiment pas faire face à plusieurs capteurs pour des raisons de convenance ou de gêne. Les systèmes multimodaux nécessitent des investissements importants en matériel, logiciels, besoins humains et en temps. Enfin, compte tenu des investissements initiaux importants dans le système multimodal, le système doit fonctionner à un niveau d'acceptation standard.

2.8.4.3. Solutions aux défis

Comparés à la biométrie monomodale, les systèmes multimodaux offrent de nombreux avantages, mais ils sont également confrontés à des défis. Pour bénéficier de ces avantages, il est nécessaire de proposer des solutions à ces défis. Une des solutions proposées par les auteurs [50] réside dans le développement d'environnements de développement intégré (IDE). Les applications doivent actuellement être développées en important différents SDK (software developpement kit) pour différentes modalités biométriques. Cette tâche devient fastidieuse et prend du temps. Un IDE spécifique conçu pour développer une application biométrique allégerait la charge des programmeurs en leur permettant de se concentrer sur de meilleures structures de données et algorithmes pour l'implémentation. Nous suggérons d'autres solutions telles que (1) l'utilisation/conception d'un capteur capable d'acquérir des données provenant de plusieurs modalités différentes et (2) la conception de systèmes capables d'extraire plusieurs caractéristiques différentes de plusieurs modalités à partir d'une seule acquisition. Ces solutions peuvent réduire considérablement les coûts de mise en œuvre des systèmes multibiométriques.

2.8.5. Focus sur la fusion de scores des systèmes biométriques

2.8.5.1. Architectures de fusion des systèmes multimodaux

Les systèmes multimodaux combinent plusieurs systèmes biométriques et nécessitent donc l'acquisition et le traitement de plusieurs données suivant trois modes, à savoir (i) le mode en série (ou en cascade), (ii) le mode en parallèle et (ii) celui hiérarchique [51,52]. Dans le mode en série, l'acquisition et le traitement sont effectués successivement et un score de similarité est obtenu à la fin de chaque acquisition tandis que dans le mode en parallèle, le traitement est effectué simultanément et le score final est obtenu après tous les traitements. En mode hiérarchique, les classificateurs individuels sont joints dans une structure semblable à une arborescence hiérarchique. Ce mode est préférable lorsqu'on est

en face de plusieurs classificateurs différents. En fait, l'acquisition des données biométriques est généralement séquentielle pour des raisons pratiques. L'architecture est donc généralement liée au traitement et en particulier à la décision. L'architecture en parallèle (figure 1.6) est la plus utilisée car elle permet d'utiliser toutes les informations disponibles et donc d'améliorer les performances du système. Par ailleurs, l'acquisition et le traitement de nombreuses données biométriques sont coûteux en temps et en matériel, et réduisent la facilité d'utilisation. Par conséquent, l'architecture en série (figure 1.7) peut être préférée dans certaines applications ; par exemple, si la multimodalité est utilisée pour offrir une alternative aux personnes incapables d'utiliser l'une des modalités en présence.

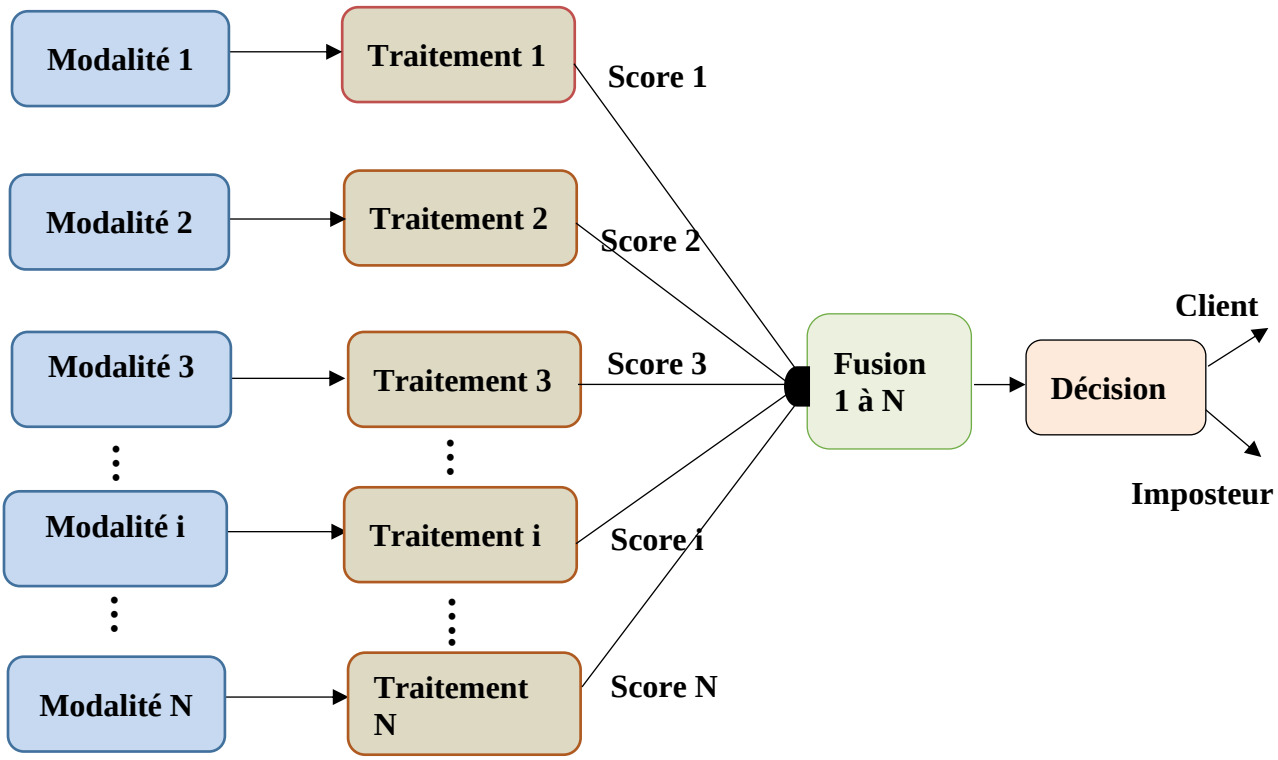

Figure 2.6 : Architecture de fusion en parallèle (adaptée de [52])

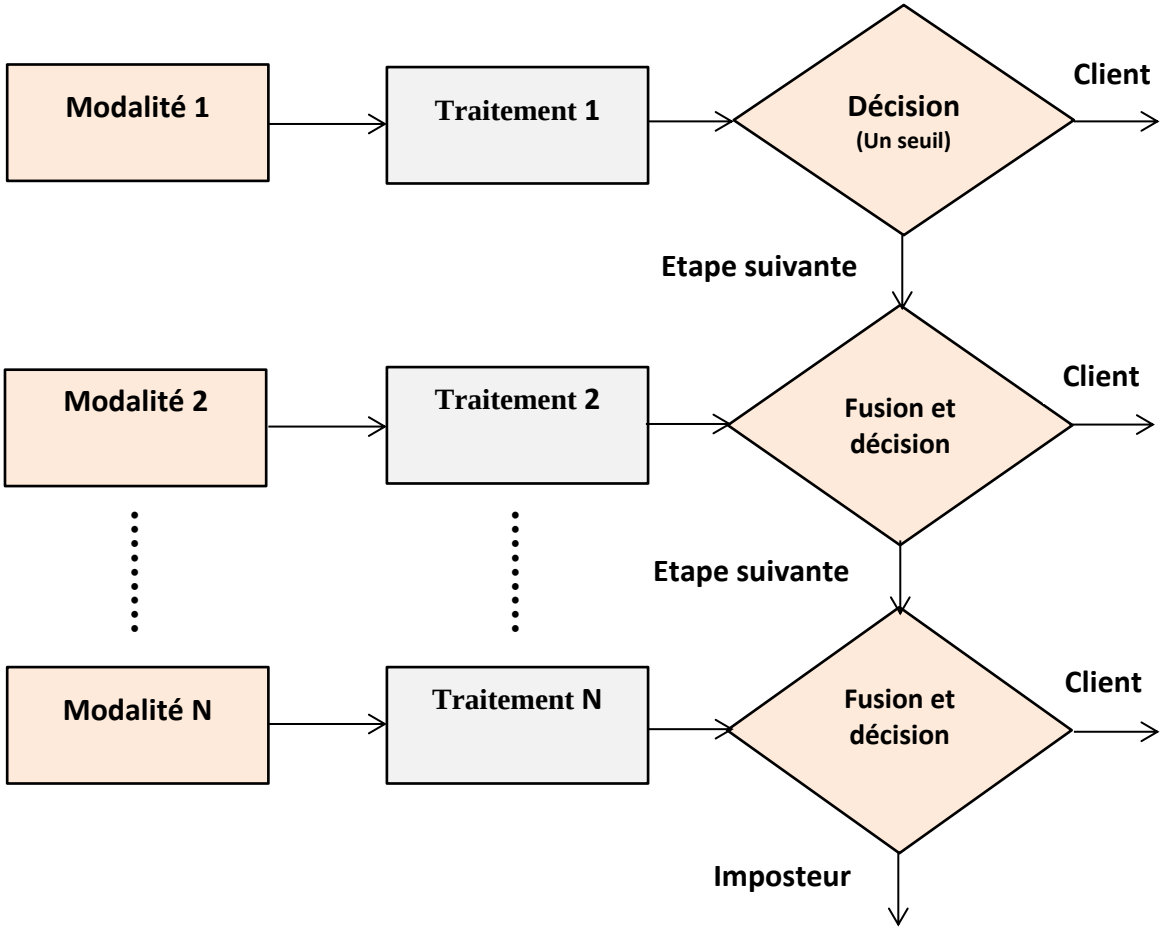

Figure 2.7 : Architecture de fusion en série (adaptée de [52])

Pour améliorer les performances de fusion, Allano et al. [53] ont proposé le mode de fusion séquentielle ou incrémentale. Ce mode est dérivé de celui en série avec la particularité d'avoir deux seuils de décision au lieu d'un seul seuil dans le mode en série.

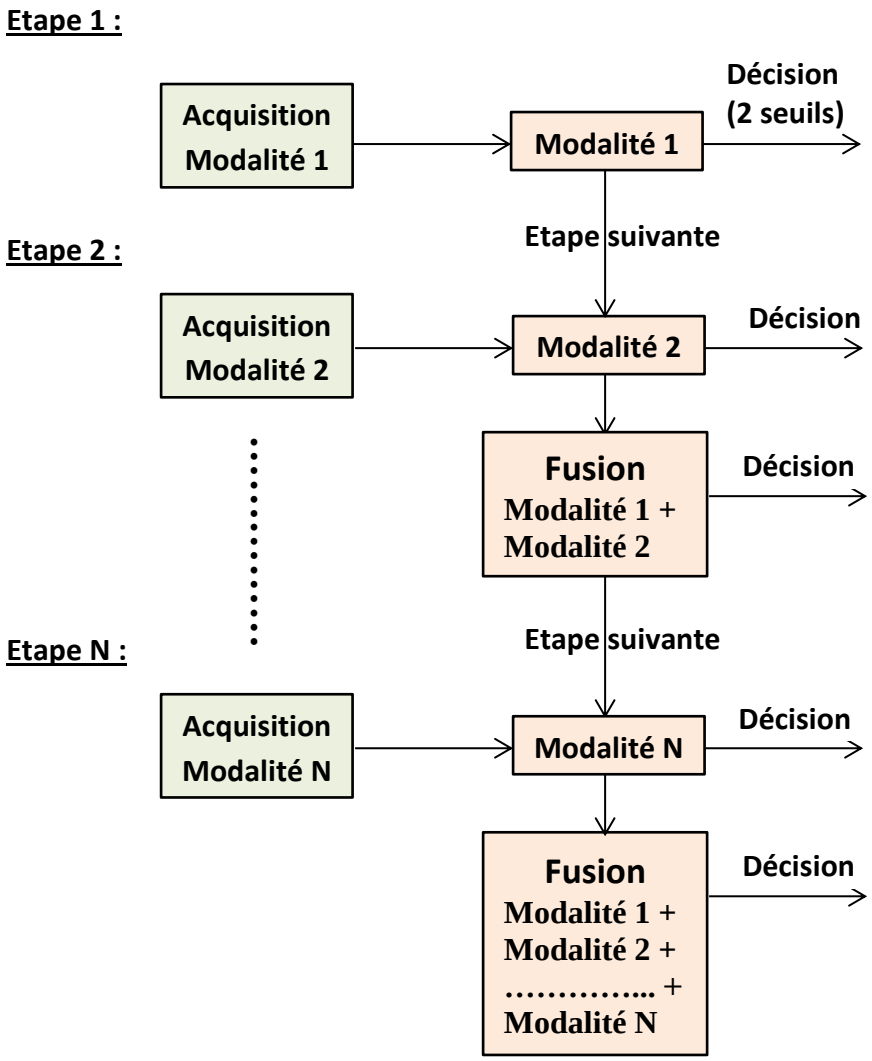

Figure 2.8 : Architecture de fusion séquentielle [52]

2.8.5.2. Méthodes de combinaison de scores

Les méthodes de combinaison de scores sont des méthodes très simples dont l'objectif est d'obtenir un score final S à partir des N scores disponibles S_i pour $i = 1$ à N à partir de N systèmes. Les méthodes les plus couramment utilisées sont la moyenne, le produit, le minimum, le maximum ou la médiane.

La combinaison des scores par la moyenne consiste à calculer S tel que:

$$S = \frac{1}{N}\sum_{i=1}^{N} S_i \qquad (1.1)$$

La combinaison des scores par le produit consiste à calculer S tel que:

$$S = \prod_{i=1}^{N} S_i \tag{1.2}$$

La combinaison des scores par le minimum consiste à calculer S tel que:

$$S = \min (S_i) \tag{1.3}$$

La combinaison des scores par le maximum consiste à calculer S tel que:

$$S = \max (S_i) \tag{1.4}$$

La combinaison des scores par la médiane consiste à calculer S tel que:

$$S = \text{med} (S_i) \tag{1.5}$$

Toutes ces méthodes sont simples et ne nécessitent aucune adaptation. Il existe également des méthodes de combinaison plus avancées qui nécessitent des paramètres tels que la somme pondérée:

$$S = \sum_{i=1}^{N} \omega_i S_i \tag{1.6}$$

La somme pondérée permet d'attribuer des poids différents ω_i à chacun des sous-systèmes en fonction de leurs performances individuelles ou de leur intérêt pour le système multimodal.

Cependant, toutes ces méthodes de combinaison ne peuvent être utilisées que si tous les scores issus des sous-systèmes sont homogènes. Pour cela, les méthodes de combinaison des scores nécessitent une étape préalable de normalisation des scores.

2.9. Les applications de la biométrie

Les exigences liées au niveau de développement de l'humanité et aux contraintes de sécurité nécessitent une authentification rapide et fiable de l'utilisateur. La biométrie est un domaine où la technologie améliore notre capacité à identifier

une personne [54]. Les applications de la biométrie peuvent être divisées en trois groupes principaux [13,15,43]:

- Les applications commerciales concernent l'accès au réseau informatique, la sécurité des données électroniques, le commerce électronique, l'accès Internet, la carte de crédit, le contrôle de l'accès physique, le téléphone portable, la gestion des dossiers médicaux, les études en ligne, etc. C'est le cas de la Royal Bank of Scotland (RBS) et NatWest qui ont été les premiers à utiliser la technologie des empreintes digitales via les téléphones mobiles pour authentifier les utilisateurs.
- Les applications gouvernementales comprennent la carte d'identité nationale, le permis de conduire, la sécurité sociale, le contrôle au niveau des frontières, le contrôle des passeports, l'enregistrement des électeurs, etc. Par exemple, le gouvernement du Bénin a créé une liste électorale permanente informatisée, dont la première version a été mise en ligne en janvier 2015. Depuis lors, elle a servi de base aux différentes élections tenues dans le pays. Le dispositif Schiphol Privium de l'aéroport d'Amsterdam utilise un capteur d'iris pour accélérer le processus de vérification des passeports et des visas. Le système US-VISIT est un autre projet lancé par le Département américain de la sécurité intérieure qui collecte, conserve et partage des informations, notamment des identifiants biométriques, sur certains ressortissants étrangers demandant un visa ou une entrée dans le pays.
- Les applications médico-légales comprennent l'identification du corps, les enquêtes criminelles, le terrorisme, la détermination du lien parental, les cas d'enfants disparus, etc. Le FBI (Federal Bureau of Investigation) du ministère de la Justice américain utilise le système intégré automatisé d'identification d'empreintes digitales, un système d'authentification utilisant les empreintes digitales des dix doigts [50].

2.10. Biométrie et vie privée

Les données biométriques sont des données personnelles car elles permettent d'identifier une personne. Leur utilisation non encadrée pose une question très délicate relative aux droits et libertés fondamentaux. Une compagnie d'assurance maladie peut par exemple utiliser des informations supplémentaires sur les antécédents d'un individu fournies par un système biométrique pour des gains économiques en refusant des avantages sociaux à une personne considérée comme présentant un risque élevé.

La généralisation des systèmes de traçabilité des hommes soulève de nombreuses questions. Avec la vidéosurveillance et le développement des nanotechnologies, l'utilisation de la biométrie appliquée à l'homme soulève des questions de bioéthique. Le terme « ***nanotechnologie*** » recouvre d'une manière générale la recherche sur les principes et propriétés existant à l'échelle nanométrique, c'est-à-dire au niveau des atomes et des molécules. En termes de libertés et de droits fondamentaux, la biométrie oppose clairement le droit de l'individu à la protection des données et de la vie privée à l'exigence de sécurité collective. Il apparaît donc la nécessité de rechercher un équilibre entre ces droits et les intérêts légitimes de sécurité publique.

Du côté positif, la biométrie peut être utilisée comme un moyen de protéger la vie privée des personnes en préservant leur identité et leur intégrité. Par exemple, l'utilisation d'une carte de crédit et l'accès aux dossiers médicaux peuvent être sécurisés par la vérification d'une modalité biométrique. Cependant, une législation spécifique est nécessaire pour garantir que ces informations restent confidentielles et que leur utilisation abusive soit punie de manière appropriée [43].

Pour protéger la vie privée des individus, les systèmes biométriques doivent être sécurisés. Il faut trouver des solutions aux nombreuses vulnérabilités mentionnées

à la section 7. L'une des solutions généralement mises en œuvre est liée à la protection du modèle. Les auteurs [55] ont mené une étude qui illustre un aperçu complet de la recherche en multibiométrie, de la protection de leurs modèles et des problèmes de confidentialité qui se posent.

2.11. Influence de l'émotion sur la performance des systèmes biométriques

L'informatique affective est l'étude et le développement de systèmes et de dispositifs capables de reconnaître, d'interpréter, de traiter et de simuler les effets d'une activité humaine. C'est un domaine interdisciplinaire englobant l'informatique, la psychologie et les sciences cognitives [56]. L'informatique affective et l'analyse des sentiments sont des facteurs décisifs pour le développement de l'intelligence artificielle et de tous les domaines de recherche qui en découlent [57]. De nombreux concepts liés à l'informatique affective sont souvent utilisés de manière interchangeable dans la littérature, à savoir affect, sensation, émotion, sentiment et opinion. Dans le reste du document, nous nous concentrons sur le concept d'émotion.

Les émotions peuvent être générées par plusieurs sources, telles que a) les choses auxquelles nous pensons, b) les actes que nous menons, c) la façon dont nous réagissons aux stimulis [58,59]. Selon Tayari et al. [60], l'émotion est une notion floue et difficile à définir. Plusieurs définitions et rôles ont été donnés à l'émotion. Ces définitions diffèrent selon les différentes approches proposées. Cependant, malgré ces divergences, la plupart des auteurs contemporains conservent une définition consensuelle des états émotionnels. Ils décrivent l'émotion comme un système de réponse complexe intégrant trois aspects :

- l'aspect physiologique / biologique qui couvre les réactions physiologiques (rythme cardiaque, rythme respiratoire, ...),
- l'aspect comportemental qui couvre les réactions comportementales et expressives, fortement influencées par la personnalité du sujet et

- l’aspect cognitif couvrant les réactions cognitives et expérientielles (état ou sentiment interne).

Il convient de noter que l’humeur et la personnalité ont une influence sur l’émotion. Djara et al. [61] ont présenté les concepts et outils liés à ce domaine d’études. La biométrie intervient principalement dans la phase de détection et de reconnaissance des émotions. L'émotion a une influence négative sur la performance des systèmes biométriques. Cependant, cette influence est limitée essentiellement au niveau des modalités comportementales. Il a été prouvé que les états émotionnels sont accompagnés de réactions physiques. Cela rend possible la caractérisation de l'émotion à travers la mesure de caractéristiques physiques. La mesure biométrique des émotions la plus couramment explorée est la reconnaissance de l'expression faciale [62]. La référence [45] a réalisé un travail sur l'influence de l'émotion sur le visage. Il estime que l'expression faciale de l'émotion, combinée à la parole, peut produire des changements significatifs dans l'apparence des visages. Dans la revue de littérature présentée par la référence [45], certains auteurs affirment que les expressions faciales n’ont pas d’influence notable sur les algorithmes de reconnaissance, si elles restent raisonnables.

Etant donné que nous pouvons mesurer l'émotion exprimée ou ressentie par un individu, nous pensons qu'il est alors possible d'utiliser les valeurs obtenues comme métadonnées dans l'authentification biométrique. Cette hypothèse fera l’objet de travaux futurs.

Conclusion

Les systèmes multibiométriques sont censés améliorer la précision de la reconnaissance d’un système d’authentification des personnes en combinant des données provenant de plusieurs sources d’information. La précision de la reconnaissance peut être encore améliorée en ajoutant au système des informations auxiliaires. Ce chapitre présente un aperçu des différentes méthodes

biométriques, les systèmes multibiométriques avant de traiter des schémas de fusion. Nous nous sommes intéressés ensuite à la fusion de scores dans les systèmes multibiométriques après quoi nous avons traité de l'influence de l'émotion sur la performance des systèmes biométriques. Sur la base de la multitude de modalités présentées, des systèmes biométriques sont mis au point et améliorés au fur et à mesure. Ces systèmes font l’objet d’évaluation et de comparaison en fonction de leur niveau de performance. Nous verrons dans le chapitre suivant les outils utilisés à cette fin.

Chapitre 3 : Evaluation des performances des systèmes biométriques

Sommaire

Introduction

La biométrie est une technique globale visant à établir l'identité d'une personne en mesurant une de ses caractéristiques physiologiques ou comportementales. Les techniques usuelles de contrôle d'accès sont basées sur ce que l'on sait (mot de passe, code PIN, . . .) et sur ce que l'on possède (carte d'identité, badge, …). Mais ces méthodes posent des problèmes de fiabilité (falsification de document, oubli de son code, décryptage du mot). Contrairement à ce que l'on sait ou ce que l'on possède, la biométrie est basée sur ce que l'on est et permet ainsi d'éviter la duplication, le vol, l'oubli ou la perte. Selon le livre blanc sur "la compréhension de l'évaluation des performances en biométrie" [63], deux questions très fondamentales se posent souvent lorsqu'on traite des systèmes ou composants biométriques : Comment peut-on mesurer la précision d'un système biométrique (ou de ses composants) et comment comparer différents systèmes les uns avec les autres ? La réponse à ces deux questions passe par la détermination d'une sixième modalité qui est celle de la performance. Ce facteur de performance a un double avantage pour le concepteur mais aussi pour l'utilisateur du système. Les systèmes biométriques sont conçus et mis au point dans des laboratoires avec pour vocation d'être utilisés dans la vie courante. Mais avant leur déploiement en

situation réelle, il est nécessaire de les évaluer afin de connaître leurs performances et leurs limites. Selon les applications, cette évaluation peut prendre en compte plusieurs paramètres tels que : la facilité d'usage pour les utilisateurs, la sécurité, le coût, les problèmes de protection de données, la fiabilité du système ou des capteurs, les nécessités de maintenance, les besoins humains de contrôle en mode opérationnel et bien sûr les taux d'erreur de reconnaissance [52].

Dans un autre travail consacré à l'estimation des performances des systèmes biométriques, Phillips et al. [64] ont spécifié trois types d'évaluation en fonction de l'application. Il s'agit de : l'évaluation technologique, l'évaluation de scénario et l'évaluation opérationnelle. L'évaluation technologique se charge de tester uniquement les performances des parties algorithmiques du système (extraction de caractéristiques, comparaison et décision) en utilisant une base de données pré-acquise. L'évaluation de scénario couvre un champ d'action plus vaste qui comprend en plus les capteurs, l'environnement et la population spécifique à l'application (scénario) testée. L'évaluation opérationnelle quant à elle prend en compte un système biométrique global en condition réelle d'utilisation. Nous présentons dans ce chapitre les principaux outils permettant d'évaluer les performances d'un système biométrique.

3.1. Les mesures des taux d'erreur

L'Organisation Internationale de Normalisation ISO/IEC 19795-1 [66] subdivise les mesures des taux d'erreur en trois classes à savoir : les taux d'erreur fondamentale, les taux d'erreur de systèmes d'authentification et les taux d'erreur de systèmes d'identification [52,65].

3.1.1. Taux d'erreur fondamentale

Nous abordons dans cette section quatre types d'erreur fondamentale à savoir :

- Taux d'échec à l'acquisition (*failure-to-acquire rate, FTA*) : proportion des tentatives de vérification ou d'identification pour lesquelles le système biométrique n'a pas pu acquérir l'information biométrique requise ;
- Taux d'échec à l'enrôlement (*failure-to-enroll rate, FTE*) : proportion des individus pour lesquels le système n'a pas pu générer le modèle biométrique durant la phase d'enrôlement. Prenons par exemple le cas des empreintes, il existe certaines personnes qui n'ont pas d'empreintes pour des raisons génétiques, ou des empreintes quasi-inexistantes pour des raisons médicales ;
- Taux de fausse non-correspondance (*false non-match rate, FNMR*) : proportion de fausse non-correspondance, par l'algorithme de comparaison, entre la donnée biométrique acquise et le modèle correspondant ;
- Taux de fausse correspondance (*false match rate, FMR*) : proportion de fausse correspondance, par l'algorithme de comparaison, entre la donnée biométrique acquise et le modèle correspondant à un autre individu.

3.1.2. Taux d'erreur de systèmes d'authentification

Les erreurs de classification correspondent aux erreurs de décision des systèmes. Ces erreurs de décision sont de deux types : Fausses Acceptations (FA) : si le système déclare l'individu comme étant le client alors que c'est un imposteur. Faux Rejets (FR) : si le système rejette l'individu alors que c'est le client. Lors de l'évaluation d'un système de vérification sur une base de données, on mesure des taux d'erreur sur cette base. Le taux de Fausses Acceptations (FAR : "False Acceptance Rate") est égal au nombre de Fausses Acceptations divisé par le nombre de tests Imposteur dans la base (Ni). Le taux de Faux Rejets (FRR : "False Rejection Rate") est égal au nombre de Faux Rejets divisé par le nombre de tests Client dans la base (Nc). Les taux d'erreur de décision des systèmes de vérication

biométriques (FAR et FRR) sont dépendants du seuil de décision fixé dans le module de décision et sont en général donnés en fonction du seuil (τ) :

$$FAR(\tau) = FA(\tau)/Ni \quad (2.1)$$

et

$$FRR(\tau) = FR(\tau)/Nc \quad (2.2)$$

En pratique, ces deux taux se calculent de la manière suivante :

- Taux de faux rejets (*false rejection rate, FRR*) : proportion des transactions des utilisateurs légitimes rejetées par erreur. Ces transactions sont rejetées, par l'algorithme de correspondance, en raison de non-correspondance à tort ainsi que ceux rejetées en raison d'un échec à l'acquisition.

Exemple : pour une transaction de vérification à une seule tentative et un seuil fixé τ, le taux de faux rejets est calculé comme suit :

$$FRR(\tau) = FTA + FNMR(\tau) \times (1 - FTA) \quad (2.3)$$

- Taux de fausses acceptations (*false acceptance rate, FAR*) : proportion des transactions des imposteurs acceptées par erreur.

Exemple : pour une transaction de vérification à une seule tentative et un seuil fixé τ, le taux de fausses acceptations est calculé comme suit :

$$FAR(\tau) = FMR(\tau) \times (1 - FTA) \quad (2.4)$$

3.2. Les courbes de performance

La performance d'un système biométrique pour différents paramétrages (seuil de décision) est illustrée graphiquement à l'aide de courbes spécifiques. Dans le cadre de ce travail, nous allons présenter deux courbes à savoir :

- La courbe ROC ou *Receiver operating characteristic curve* [67] : cette courbe constitue l'une des méthodes les plus couramment utilisées afin d'évaluer la performance globale d'un système d'authentification biométrique. La courbe ROC représente la variation du taux de faux rejets ou FRR (en ordonnées) en fonction du taux de fausses acceptations ou FAR (en abscisses) lorsque le seuil de décision varie ;

- La courbe RC ou *Robustness curve* : cette courbe illustre la robustesse du système en terme de performance contre les divers types d'altérations (*i.e.*, altérations dues au bruit pendant l'acquisition de données biométriques). La performance du système est illustrée par le point de fonction *taux d'égale erreur* (EER) qui fera l'objet d'une présentation détaillée dans la section 2.4. Cette courbe a été présentée dans [68] avec de plus amples détails.

3.3. Les points de fonctionnement ou points de performance

On utilise les points de performance pour illustrer la performance des systèmes biométriques. Parmi les métriques les plus utilisées dans la littérature, nous avons le taux d'égale erreur (EER), le taux d'erreur pondérée (WER), le taux d'erreur moyenne (HTER), l'aire sous la courbe ROC (AUC) et la capacité [66,69].

- Taux d'égale erreur (*Equal Error Rate*, EER) : l'EER est obtenu à l'intersection de la courbe ROC et de la droite *d* : FAR = FRR. Cette valeur n'a presque pas d'utilité pratique car on ne souhaite généralement pas que le FAR et le FRR soient les mêmes, mais elle constitue un indicateur de la précision du dispositif biométrique. En d'autres termes, plus l'EER est faible, plus le système est performant. A noter que ce taux d'erreur est le plus couramment utilisé dans la littérature pour illustrer la performance des systèmes biométriques ;

- Taux d'erreur pondérée (*Weighted Error Rate*, WER) : ce taux d'erreur correspond au seuil tel que le FRR soit proportionnel au FAR avec un coefficient qui dépend de l'application. Pour un coefficient égal à 1, le seuil du WER est égal au seuil de l'EER ;

- Taux d'erreur moyenne (*Half Total Error Rate*, HTER) : c'est une métrique qui correspond à la moyenne entre le FAR et FRR pour un seuil fixé τ :

$$HTE(\tau) = \frac{FAR(\tau)+FRR(\tau)}{2} \tag{2.5}$$

- Aire sous la courbe ROC (*Area Under ROC Curve*, AUC) : c'est une métrique qui permet de quantifier la diversification de la distribution de scores des

utilisateurs légitimes et d'imposteurs. En d'autres termes, étant donnés deux utilisateurs choisis au hasard, un parmi les utilisateurs légitimes et l'autre parmi les imposteurs, l'AUC représente la probabilité $P(S^{gen} > S^{imp})$ (*i.e.*, probabilité de bon classement). Plusieurs méthodes sont proposées dans la référence (2.6) pour estimer l'AUC. Tronci *et al.* [71] ont suggéré une estimation de l'AUC basée sur le test statistique de Wilcoxon-Mann-Whitney (WMW) [72]. L'AUC est ainsi définie par :

$$AUC = \frac{\sum_{p=1}^{n_g}\sum_{q=1}^{n_i} I(s_p^{gen}, s_q^{imp})}{n_g\, n_i} \tag{2.6}$$

Où n_g et n_i représentent le nombre des utilisateurs légitimes et d'imposteurs respectivement, $\{s_p^{gen}\}$ et $\{s_q^{imp}\}$ correspondent aux scores des utilisateurs légitimes et d'imposteurs respectivement, et la fonction $I(s_p^{gen}, s_q^{imp})$ est définie par :

$$I(s_p^{gen}, s_q^{imp}) = \begin{cases} 1 \; si \; s_p^{gen} > s_q^{imp}) \\ 0 \; sinon \end{cases} \tag{2.7}$$

L'AUC constitue également un bon indicateur pour évaluer et comparer les systèmes biométriques. Plus l'AUC est grande, plus l'algorithme est performant.

- Capacité d'un système biométrique (*Constrained capacity of biometric system*) : la capacité permet de quantifier la performance de systèmes biométriques, en utilisant la base de données des utilisateurs et la fonction de similarité. Bhatnagar et Kumar [69] ont caractérisé la distribution de scores des utilisateurs légitimes et d'imposteurs par la distribution gaussienne. Les indices de performances proposés sont :

  - La capacité d'un utilisateur m, notée par C_m, permet d'illustrer le degré de distinction de l'utilisateur m par rapport aux autres utilisateurs de la base. Elle est donnée par :

$$C_m = \frac{1}{2} log_2\left(1 + \frac{d_m^2}{4\max(\sigma_{g(m)}^2, \sigma_{i(m)}^2)}\right) \tag{2.8}$$

Où d_m est la distance entre les médianes $\hat{g}_m$ et $\hat{\imath}_m$, $\sigma^2_{g(m)}$ et $\sigma^2_{i(m)}$ sont les variances de la distribution de scores intraclasses (échantillons du même utilisateur) et interclasses (échantillons d'utilisateurs différents), respectivement.

- La capacité d'un système biométrique, notée par C_s, permet d'illustrer la fiabilité du système en terme de nombre d'utilisateurs correctement authentifiés (*i.e.*, peuvent être authentifiés d'une manière sûre). Elle est donnée par :

$$C_s = \frac{1}{2} log_2\left(1 + \frac{\overline{d_m^2}}{4\max\left(\overline{\sigma_g^2}, \overline{\sigma_i^2}\right)}\right) \quad (2.9)$$

Où $\overline{d_m}$ est la moyenne des distances d_m pour chaque utilisateur de la base, σ_g^2 et σ_i^2 sont les moyennes des variances des distributions intraclasses et interclasses, respectivement.

Après avoir offert un aperçu sur les points de fonctionnement, nous abordons les complémentarités qui les caractérisent. En effet, le point particulier EER est le plus couramment utilisé dans la littérature pour évaluer et comparer les systèmes biométriques de façon traditionnelle. Dans la section précédente, nous avons vu la faiblesse d'utiliser seulement l'EER pour comparer les systèmes biométriques. Pour des systèmes biométriques ayant des taux d'erreur différents, l'utilisation de l'EER peut être suffisante pour affirmer qu'un système est meilleur qu'un autre. Tandis que dans le cas où les systèmes à comparer présentent des taux d'erreurs similaires (lors des compétitions), utiliser une métrique complémentaire devient indispensable. Nous avons vu qu'il existe dans la littérature d'autres métriques complémentaires à l'EER (l'AUC, la courbe RC et la capacité) que nous pouvons utiliser afin de comparer les systèmes biométriques dans un cadre précis. L'AUC présentée dans la référence [71], qui permet de quantifier la diversification de scores des utilisateurs légitimes et d'imposteurs, est un bon indicateur de performance complémentaire à l'EER. Il permet de bien représenter les

performances globales de l'algorithme. Enfin, nous pouvons conclure que les métriques AUC, capacité et RC sont complémentaires à l'EER pour avoir une meilleure précision sur la performance du système testé.

3.4. Les bases de données d'évaluation

L'évaluation des systèmes biométriques nécessite le recours à une base de données dédiée à cette évaluation. Cette base assure que les systèmes seront testés selon les mêmes conditions d'acquisition. Les bases de données biométriques peuvent servir également à régler les paramètres d'un système monomodal (paramétrage du seuil de décision) et multimodal (paramétrage des poids pour la fusion). Les bases de données biométriques collectées sont généralement divisées en trois types : bases de données réelles, synthétiques et virtuelles.

3.4.1. Bases de données réelles

Ces bases contiennent des données biométriques réelles acquises grâce à la participation de volontaires. Dans la littérature, il existe deux ensembles des bases : les bases monomodales et celles multimodales. Dans la catégorie des bases monomodales, nous pouvons citer entre autres : *FACES94*, *AR*, *FERET*, *FVC2002 DB*$_2$, FRGC (Face Recognition Grand Challenge), *USF Human ID Gait Baseline*, *ENSIB*, et *GREYC-Keystroke*. Pour ce qui est des bases multimodales, les exemples suivants peuvent être cités : *XM2VTSDB*, *BANCA* et *BIOSECURE* [65].

3.4.2. Bases de données synthétiques

Ces bases contiennent des données synthétiques permettant de simuler des données biométriques réelles. Une base synthétique doit satisfaire deux propriétés. Premièrement, la performance issue d'une base synthétique doit être proche de celle obtenue avec une base de données réelle. Deuxièmement, une donnée dans la base synthétique ne doit pas représenter une donnée biométrique réelle d'un individu. La base *SFinGe* générée par le logiciel *SFinGe* [73],

développé par le laboratoire italien *BioLab*[1], est un exemple de base synthétique [65].

3.4.3. Bases de données virtuelles

Une base virtuelle est une base de données biométriques constituée d'individus virtuels [52]. Ces individus sont générés en associant une modalité biométrique d'une personne à une (ou plusieurs) autre (s) modalité (s) biométrique (s) d'une (ou plusieurs) autre (s) personne (s). Par exemple un individu virtuel peut être formé à partir du visage d'une personne et de l'empreinte digitale de quelqu'un d'autre. Ces bases virtuelles peuvent être utilisées pour deux raisons : soit parce qu'il n'existe pas de base réelle adaptée à un problème (taille, modalités, conditions d'acquisition, etc.) soit parce qu'on ne veut pas distribuer ou rendre disponible les données associées de personnes réelles pour des raisons de protection des données personnelles. Les bases virtuelles pourraient donc être utilisées afin de rendre public ou de diffuser des données biométriques multimodales tout en réduisant les problèmes de protection des données personnelles.

Par ailleurs, il convient d'évoquer le cas spécifique des exemples virtuels ainsi que leur nuance avec les personnes virtuelles. Les exemples virtuels sont en fait des exemples générés par de petites modifications (en général des déformations géométriques) des données initiales. Ces exemples virtuels sont parfois utilisés en test, lorsque l'ensemble de test est également trop petit. Les déformations appliquées aux données varient selon la nature des données. Par exemple pour les images comme le visage, les auteurs utilisent des translations de l'image d'un pixel dans une ou plusieurs directions afin de modifier l'image. Des travaux ont également utilisé le principe d'augmentation de la base d'apprentissage par des exemples virtuels mais en utilisant cette fois des personnes virtuelles. Dans la

1 . http ://biolab.csr.unibo.it/

référence [74], les auteurs utilisent des personnes virtuelles formées de personnes non incluses dans la base multimodale réelle, car elles ne possédaient pas les deux modalités (voix et signature), en les recombinant entre elles. Cela permet alors d'augmenter la taille de la base d'apprentissage pour apprendre les paramètres de la méthode de fusion.

3.5. Intervalle de confiance

En biométrie, les bases de données collectées sont utilisées pour évaluer la performance des systèmes biométriques. Cependant, ces bases ne sont pas représentatives de la population globale pour deux raisons principales. Premièrement, ces bases ne contiennent pas assez de personnes, et on a en général peu de données par personne. Deuxièmement, il y a souvent une différence entre le nombre de scores des utilisateurs légitimes et d'imposteurs, ce qui n'est également pas représentatif de la réalité. Enfin, les taux d'erreur (EER, WER, HTER et AUC) utilisés pour illustrer la performance globale du système dépendent du découpage *enrôlement* - *test*. Pour toutes ces raisons, il est nécessaire de calculer un intervalle de confiance à l'EER lors de la comparaison des systèmes biométriques. Cet intervalle de confiance est surtout indispensable lors de la comparaison des systèmes biométriques ayant des taux d'erreurs similaires. Cette section s'inspire des travaux effectués par les auteurs [52] et [65] qui ont fourni plus de détails sur la notion d'intervalle de confiance.

3.6. Les mesures de temps de traitements et occupation mémoire

Parlant du temps de traitement de l'information par le système, il s'agit d'un facteur très important pour l'évaluation de systèmes biométriques. Il est généralement mesuré en fonction de trois paramètres [65]. Premièrement, nous avons le *temps moyen d'enrôlement*. Il désigne le temps moyen mis par le système pour générer les modèles biométriques des individus, c'est-à-dire depuis la capture de la donnée biométrique jusqu'à son stockage dans la base de données. C'est ce modèle qui servira de référence lors de la comparaison. Nous avons

ensuite le *temps moyen de vérification*. Il désigne le temps moyen mis par le système pour l'acquisition des données biométriques requises et la comparaison de ces données avec le modèle correspondant. Ce temps ne dépend pas du nombre de personnes dans la base de données. Le troisième paramètre quant à lui est relatif au *temps moyen d'identification*. Il désigne le temps moyen pour l'acquisition des données biométriques requises et la comparaison de ces données avec l'ensemble des modèles existants dans la base. Cette information est considérablement impactée par le nombre d'utilisateurs du système.

L'espace mémoire requis par le système est aussi un important facteur qui entre en ligne de compte dans le cadre de l'évaluation de systèmes biométriques. Il est généralement mesuré en termes de *taille moyenne et maximale d'un modèle biométrique* et en *espace mémoire maximal alloué* pendant les phases d'enrôlement, de vérification et d'identification.

3.7. Evaluation de la sécurité d'un système biométrique

El-Abed [65] a effectué plusieurs travaux consacrés à l'évaluation biométrique. Au sujet de la sécurité en biométrie, il a évoqué la norme de l'Organisation Internationale de Normalisation ISO/IEC FCD 19792 [75] qui porte sur les aspects de l'évaluation de la sécurité des systèmes biométriques. Cette norme n'avait pas pour but de définir une méthodologie pour l'évaluation de ces systèmes, mais définit les principales exigences à prendre en considération lors de l'évaluation de ces systèmes. Outre les menaces et les vulnérabilités sur une application générique, la norme aborde des menaces liées à la performance du système et la qualité des données biométriques pendant la phase d'enrôlement. D'autres facteurs ont été également signalés tels que le respect des droits et des libertés fondamentales. La norme allègue que la gestion de stockage et la politique d'accès aux modèles biométriques est un facteur primordial lors du processus d'évaluation de tels systèmes.

Dimitriadis et Polemi [76] ont présenté une méthode d'évaluation quantitative pour la sécurité des systèmes biométriques. Ils présentent une liste de 12 vulnérabilités de ces systèmes et proposent quelques contre mesures associées. La méthode proposée calcule un facteur de risque associé à chaque vulnérabilité. Elle est facile à utiliser et efficace. Cependant, la méthode ne prend pas en considération les besoins de sécurité (confidentialité, intégrité, *etc.*), comme ils sont présents dans la méthode d'audit de sécurité EBIOS [77]. De plus, d'autres attaques et vulnérabilités doivent être prises en compte, surtout celles présentées dans la norme ISO/IEC FCD 19792 [75] qui ne sont pas prises en compte dans ces travaux, pour mieux évaluer et comparer les systèmes biométriques.

3.8. Les compétitions et plateformes

La comparaison des systèmes biométriques nécessite un protocole d'évaluation et une base de données bien définis. Le protocole d'évaluation assure que les systèmes biométriques sont testés sous les mêmes conditions. Nous présentons dans cette section un aperçu sur les compétitions les plus citées dans la littérature et qui sont structurées en trois catégories : les compétitions monomodales et multimodales puis les plateformes.

3.8.1. Compétitions monomodales

3.8.1.1. Concours de vérification des signatures ou Signature Verification Competition (SVC)

SVC [78] est une compétition de vérification par dynamique de signature qui a été organisée en collaboration avec la conférence internationale ICBA (*International Conference on Biometric Authentication*) en 2004. Le taux d'erreur EER a été utilisé comme indicateur de performance.

3.8.1.2. Concours de vérification des empreintes digitales ou Fingerprint Verification Competition (FVC)

Les compétitions de reconnaissance d'empreinte digitale ont été organisées par plusieurs universités[2] en 2000, 2002, 2004 et 2006. Les participants ont testé leurs algorithmes en fournissant leurs fichiers exécutables correspondants aux phases d'*enrôlement* et de *vérification*. Quatre bases de données, dont trois sont réelles et une synthétique générée par le logiciel *SFinGe*, ont été employées durant la compétition *FVC2006*. Les différents indicateurs de performance utilisés sont : la distribution de scores des utilisateurs légitimes et d'imposteurs, la taille des modèles biométriques, le temps moyen d'enrôlement et de vérification, taux d'échec à l'enrôlement et les courbes ROC.

3.8.1.3. Test des fournisseurs de reconnaissance de visage et évaluation du défi Iris ou Face Recognition Vendor Test (FRVT) et Iris Challenge Evaluation (ICE)

Les compétitions de reconnaissance faciale (FRVT) et d'iris (ICE) [79] sont des compétitions organisées par l'Institut National des Standards et de la Technologie (NIST). Ces compétitions utilisent les courbes ROC comme indicateur de performance.

3.8.1.4. Evaluation de la reconnaissance du locuteur ou Speaker Recognition Evaluation (SRE)

Les compétitions de reconnaissance vocale SRE ont été organisées par le NIST. Il y a eu un grand nombre des compétitions SRE[3] depuis 1997 et pour la 12ème fois en 2010.

2.http://bias.csr.unibo.it/fvc2006/
3.http://www.itl.nist.gov/iad/mig/tests/sre/

3.8.1.5. Evaluation de la technologie des fournisseurs d'empreintes digitales ou Fingerprint Vendor Technology Evaluation (FpVTE)

Cette competition a été organisée par le NIST en 2003. Quatre métriques ont été utilisées pour évaluer les systèmes biométriques. Il s'agit de ROC, DET, FAR et FRR.

3.8.1.6. Test indépendant de la technologie de reconnaissance de l'iris ou Independent Testing of Iris Recognition Technology (ITIRT2005)

ITIRT2005 a été organisée par International Biometric Group (IBG) en 2005. Comme indicateur de performance, le FNMR, FMR, Transactional-FNMR, Transactional-FMR, FTA, Transactional-FTA et FTE ont été utililisés.

3.8.1.7. Tests biométriques comparatifs ou Comparative Biometric Testing (CBT2006)

Le Comparative Biometric Testing ou CBT2006 a également été organisée par IBG en 2006. Les solutions biométriques ont été évaluées sur la base des métriques suivantes : FNMR, FMR, Transactional-FNMR, Transactional-FMR, FTA, Transactional-FTA et FTE.

3.8.2. Competitions multimodales

3.8.2.1. Campagne d'évaluation multimodale BioSecure ou BioSecure Multimodal Evaluation Campaign (BMEC)

BMEC[4] est une compétition organisée par le réseau d'excellence BioSecure[5] en 2007 [65]. Cette compétition comprenait des parties pour l'évaluation des systèmes biométriques monomodaux et une partie pour les systèmes multimodaux (fusion au niveau de scores). La base multimodale BioSecure a été utilisée lors de cette compétition.

4. http://biometrics.it-sudparis.eu/BMEC2007/
5. http://biosecure.it-sudparis.eu

3.8.2.2. Grand défi biométrique multiple ou Multiple Biometric Grand Challenge (MBGC)

MBGC est une compétition multimodale organisée par le NIST en 2009. L'objectif de cette compétition consiste à améliorer les systèmes de reconnaissance basés sur le visage et l'iris selon différentes conditions d'acquisition. Elle consiste également à évaluer les algorithmes de fusion, au niveau des images et au niveau de scores de ces deux modalités.

3.8.3. Plateformes

Outre les compétitions, il existe plusieurs plateformes permettant aux chercheurs et aux développeurs de tester leurs algorithmes par rapport à l'existant dans l'état de l'art. Nous présentons dans cette section un aperçu sur les plateformes existantes que sont BioSecure Reference and Evaluation Framework, le logiciel GREYC-Keystroke et FVC-onGoing [65].

3.8.3.1. Cadre de référence et d'évaluation de BioSecure ou BioSecure Reference and Evaluation Framework

En plus de la compétition multimodale BMEC, le réseau d'excellence BioSecure a proposé un Framework visant à comparer et évaluer les systèmes biométriques. Ce Framework comporte douze systèmes de référence relatifs aux modalités suivantes : visage, voix, iris, empreinte digitale, géométrie de la main et signature dynamique. Ces systèmes sont implémentés en quatre modules remplaçables (prétraitement, extraction de caractéristiques, génération du modèle biométrique et vérification) permettant aux développeurs et aux chercheurs de tester et d'évaluer une partie spécifique de leur système. L'indicateur de performance utilisé est la courbe DET avec son EER correspondant.

3.8.3.2. Dynamique de frappe au clavier GREYC ou GREYC-Keystroke

Le logiciel *GREYC-Keystroke*, développé dans le laboratoire de recherche GREYC, est un outil qui permet de comparer les systèmes biométriques basés sur

la dynamique de frappe au clavier. Il est également utilisé pour créer une base de données afin de comparer les différents algorithmes qui existent dans l'état de l'art sous les mêmes conditions d'acquisition. Les indicateurs de performance proposés sont : la distribution de scores des utilisateurs légitimes et d'imposteurs, les courbes ROC et le taux d'échec à l'acquisition (FTA) [65].

3.8.3.3. Concours de vérification des empreintes digitales en cours ou Fingerprint Verification Competition-onGoing (FVC-onGoing)

FVC-onGoing est une plateforme d'évaluation en ligne[6] dédiée aux algorithmes de reconnaissance des empreintes digitales. Elle est l'évolution des compétitions FVC, présentées dans la section précédente. La plateforme propose plusieurs bases de données regroupées en deux parties : la première (*Fingerprint Verification*) quantifie les deux modules d'enrôlement et de vérification, tandis que la seconde (*Fingerprint Matching (ISO)*) quantifie seulement le module de vérification sur des Templates ISO basés sur les minuties. Les indicateurs de performance proposés sont : le taux d'échec à l'enrôlement et l'acquisition (FTE et FTA), le taux de non-correspondance (FNMR) pour un taux de fausse correspondance fixé et vice versa, le temps moyen d'enrôlement et de vérification, la taille maximale requise pour stocker le modèle biométrique sur le support, la distribution de scores des utilisateurs légitimes et d'imposteurs, et la courbe ROC avec son EER correspondant.

Toutes ces compétitions et plateformes ont été synthétisées dans le tableau 3.1.

Tableau 3.1 : Compétitions biométriques internationales et plateformes [80]

Catégorie de compétition	Nom de la compétition	Année	Métrique de performance utilisée
Compétitions mono-modales	FVC	2000, 2002, 2004 et 2006	GMS et IMS, taille moyenne et maximale du modèle, temps moyen d'inscription et de vérification, FTE et courbes ROC
	FpVTE	2003	ROC, DET, FAR et FRR

6 . https://biolab.csr.unibo.it/FVCOnGoing

	SVC	2004	EER
	CBT2006	2006	FNMR, FMR, Transactional-FNMR, Transactional-FMR, FTA, Transactional-FTA et FTE
	ITIRT2005	2005	FNMR, FMR, T-FNMR, T-FMR, FTA, T-FTA et FTE
Compétitions multi-modales	BMEC	2007	Courbes ROC et leurs EERs correspondants
Platformes	BioSecure	2007	Courbes ROC et leurs EERs correspondants
	GREYC-Keystroke	2009	GMS et IMS, courbes ROC et taux FTA
	FVC-OnGoing	2009	FTE et FTA, FNMR pour FMR fixé et vice-versa, taille moyenne et maximale du modèle, temps moyen d'inscription et de vérification, GMS et IMS, courbes ROC et leurs EERs correspondants

Conclusion

Nous avons présenté dans ce chapitre les méthodes pour évaluer et comparer les performances d'un système biométrique monomodal et multimodal. Les principales propriétés de ces outils reposent respectivement sur les taux d'erreur, les bases de données, les compétitions et les plateformes. En dehors de ces trois aspects évoqués, il existe d'autres facteurs qui entrent en jeu dans la mise en œuvre de ces méthodes. Ces fateurs sont relatifs à la qualité des données acquises, l'usage en termes d'acceptabilité et de satisfaction et la sécurité en termes de robustesse contre la fraude. Dans ce chapitre, nous avons également vu qu'il existe un grand nombre de métriques statistiques d'évaluation de performances dont certaines sont complémentaires. Les outils d'évaluation développés dans la littérature ne mettent pas un accent sur les questions sécuritaires. Certains aspects de sécurité à prendre en compte ont été présentés. L'évaluation des systèmes biométriques en terme de sécurité est une phase importante à prendre en compte lors de la conception de ces systèmes. La partie suivante abordera les méthodes originales que nous avons développées.

Deuxième partie : Méthodes développées

Chapitre 4 : Architecture de fusion séquentielle adaptée de scores

Sommaire

Introduction

La biométrie unimodale fait face à une variété de problèmes tels que la non-universalité et les attaques par usurpation. La biométrie multimodale réduit les contraintes de la biométrie unimodale en combinant plusieurs modalités biométriques. La fusion des scores résultant du traitement de chaque modalité se fait selon les architectures en parallèle (fusion globale), en série (avec un seul seuil de décision) ou séquentiellement (avec deux seuils de décision) [81]. Allano et al. [53] ont montré que l'architecture séquentielle présente plus d'avantages que l'architecture en parallèle et celle en série. Sur un autre plan, Jain et al. [82] ont mené une étude sur la technique d'intégration de la biométrie douce dans les systèmes multibiométriques. Cette technique permet d'améliorer la précision de reconnaissance. Dans ce chapitre, nous proposons une architecture de fusion de scores qui combine ces deux principes. L'objectif principal est de fournir une nouvelle architecture capable d'améliorer les performances des systèmes multibiométriques. La section 2 présente les avantages de la biométrie douce dans les systèmes multibiométriques. Ensuite, nous mettons un accent sur la stratégie séquentielle de fusion des scores. Sur la base des deux principes mentionnés précédemment, un cadre de l'architecture de fusion séquentielle adaptée de scores est proposé. Enfin, les instructions détaillées de l'algorithme dérivé sont construites.

4.1. Méthode d’intégration des métadonnées dans les systèmes multibiométriques

En 2004, Jain et al. [82] ont présenté une catégorie de système multibiométrique combinant des modalités biométriques primaires (tels que le visage et les empreintes digitales) et des attributs de biométrie douce (tels que le genre, la taille, le poids, la couleur des yeux, etc.). Les modalités de biométrie douce ne peuvent pas être utilisées pour distinguer les individus de manière fiable car le même attribut est susceptible d'être partagé par plusieurs personnes différentes dans la population cible. Cependant, lorsqu'elles sont utilisées conjointement avec des modalités de biométrie primaire, les performances du système d'authentification peuvent être considérablement améliorées. Les attributs de biométrie douce facilitent également le filtrage (ou l'indexation) des bases de données biométriques volumineuses en limitant le nombre d'entrées à rechercher dans la base de données. Dans ce travail, les auteurs ont proposé un cadre pour l'intégration de l'information de biométrie douce à un système de biométrie primaire. Ce cadre comporte deux sous-systèmes pour la reconnaissance biométrique. Le premier sous-système basé sur les identifiants biométriques traditionnels (empreintes digitales, visage et géométrie de la main) est appelé système biométrique primaire. Il peut être unimodal ou multimodal. Le deuxième sous-système est appelé système biométrique secondaire. Il est basé sur des caractéristiques de biométrie douce (âge, genre, couleur des yeux, couleur de la peau et taille).

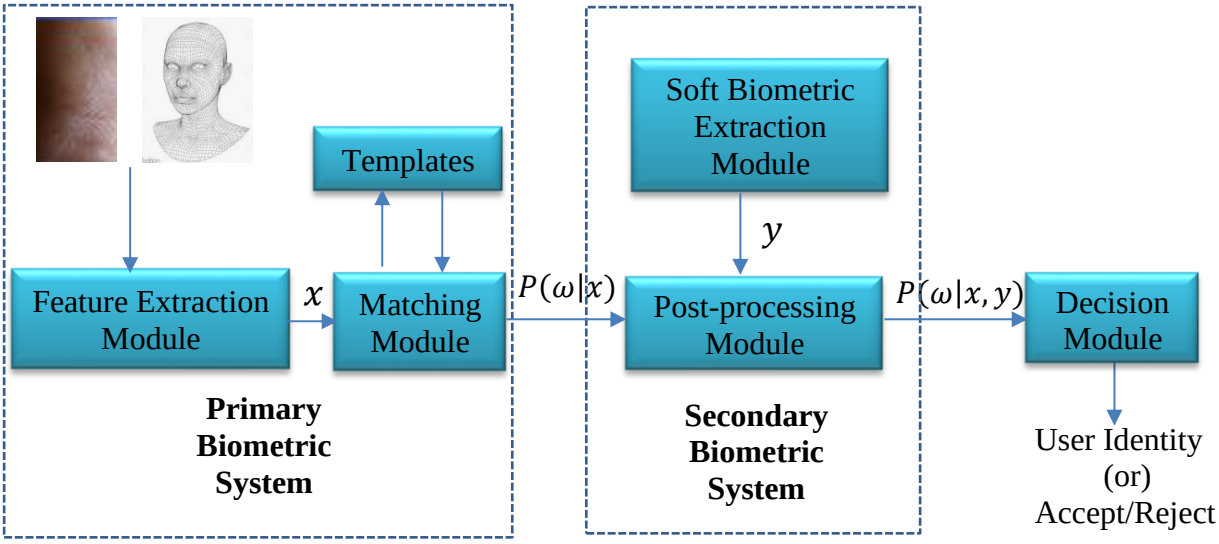

Figure 4.1 : Intégration des caractéristiques de biométrie douce à un système de biométrie primaire [82]

Les résultats expérimentaux ont montré qu'en ajoutant un trait de biométrie douce à une modalité de biométrie pure, on pouvait améliorer les performances de reconnaissance de 5%. En utilisant la règle de Bayes, la probabilité de reconnaissance d'un utilisateur à partir de ses modalités de biométrie pure d'une part et de ses caractéristiques de biométrie douce d'autre part est calculée à l'aide de la formule suivante:

$$P(\omega_i|x,y) = \frac{p(y|\omega_i)\, P(\omega_i|x)}{\sum_{i=1}^{n} p(y|\omega_i) P(\omega_i|x)} \qquad (3.1)$$

x est l'identifiant de biométrie pure ;

$y = [y_1, y_2, \dots, y_k, y_{k+1}, y_{k+2}, \dots, y_m]$ est le vecteur des attributs de biométrie douce, où y_1 à y_k sont des variables continues et y_{k+1} à y_m sont des variables discrètes ;

$\omega_1, \omega_2, \dots, \omega_n$ représentent les n utilisateurs enregistrés dans la base de données;

$P(\omega_i \mid x), i = 1,2, \dots, n$ est la probabilité que l'utilisateur test soit ω_i étant donné l'identifiant x.

En supposant que les variables de biométrie douce soient indépendantes [82], l'équation (3.1) devient :

$$P(\omega_i|x,y) = \frac{p(y_1|\omega_i)\dots P(y_k|\omega_i)\, P(y_{k+1}|\omega_i)\dots P(y_m|\omega_i)\, P(\omega_i|x)}{\sum_{i=1}^{n} p(y_1|\omega_i)\dots P(y_k|\omega_i)\, P(y_{k+1}|\omega_i)\dots P(y_m|\omega_i)\, P(\omega_i|x)} \quad (3.2)$$

Dans (3.2), $p(y_j|\omega_i), j = 1, 2, \dots, k$ est évaluée à partir de la densité conditionnelle de la variable y_j pour l'utilisateur w_i. De l'autre côté, la probabilité discrète $p(y_j|\omega_i), j = k+1, k+2, \dots, m$ représente la probabilité que l'utilisateur w_i soit affecté à la classe y_j. Afin de simplifier le problème, nous supposons que la précision du module de classification est indépendante de l'utilisateur, sur la base de l'indicateur biométrique y_j.

Soit $p(y) = \sum_{i=1}^{n} p(y_1|\omega_i) \dots P(y_k|\omega_i)\, P(y_{k+1}|\omega_i) \dots P(y_m|\omega_i)\, P(\omega_i|x)$ (3.3)

Le logarithme de $P(\omega_i|x,y)$ en (3.2) peut être exprimé comme suit :

$$\log P(\omega_i|x,y) = \log p(y_1|\omega_i) + \dots + \log p(y_k|\omega_i) + \log P(y_{k+1}|\omega_i) + \dots + \log P(y_m|\omega_i) + \log P(\omega_i|x) - \log p(y) \quad (3.4)$$

Compte tenu de l'importance relative des différentes modalités concernées, des pondérations leur sont attribuées. On obtient la fonction de discrimination suivante :

$$g_i(x,y) = a_0 \log P(\omega_i|x) + a_1 \log p(y_1|\omega_i) + \dots + a_k \log p(y_k|\omega_i) + a_{k+1} \log P(y_{k+1}|\omega_i) + \dots + a_m \log P(y_m|\omega_i) \quad (3.5)$$

avec $\sum_{i=0}^{m} a_i = 1 \; et \; a_0 >> a_i, \; i = 1, 2, \dots, m$. Notez que les coefficients $a_i, i = 1, 2, \dots, m$ représentent les poids attribués aux caractéristiques de la biométrie douce et que a_0 est le poids attribué à l'identifiant de la biométrie pure. $a_0 >> a_i$ signifie que a_0 est largement supérieure à a_i.

4.2. La stratégie de fusion séquentielle de scores

En 2009, Allano [52] a présenté la stratégie de fusion séquentielle des scores (voir architecture à la figure 2.8). Cette méthode de fusion repose sur le principe du test séquentiel du rapport de vraissemblance (Sequential Probability Ratio Test ou SPRT) [6]. Les auteurs considèrent l'hypothèse H_0 (l'utilisateur est le client) et l'hypothèse alternative H_1 (l'utilisateur est un imposteur). Chacune des deux

hypothèses est associée respectivement à une erreur dite de première espèce correspondant au rejet de H_0 alors que l'hypothèse est vraie et une erreur de deuxième espèce correspondant à l'acceptation de H_0 alors que l'hypothèse est fausse. La probabilité de l'erreur de première espèce notée α est égale au Taux de Faux Rejets (FRR) et la probabilité de l'erreur de deuxième espèce notée β est égale au Taux de Fausses Acceptations (FAR). α et β sont représentés par les équations suivantes :

$$\alpha = FRR \tag{3.6}$$

et

$$\beta = FAR \tag{3.7}$$

Le rapport de vraissemblance RV est défini comme suit :

$$RV = \frac{P(X_n|H_0)}{P(X_n|H_1)} \tag{3.8}$$

Le test séquentiel du rapport de vraissemblance (SPRT) définit k_0 et k_1 comme les critères d'arrêt autour de la frontière k. L'option de rejet représente la zone intermédiaire (zone d'incertitude) qui mène à l'étape suivante $(k+1)$ avec l'ajout de données supplémentaires. Ainsi, à l'étape k, H_0 est accepté si :

$$\frac{P(x_1,\ x_2, \dots, x_k|H_0)}{P(x_1,\ x_2, \dots, x_k|H_1)} \geq k_1 \tag{3.9}$$

H_0 est rejetée (et H_1 acceptée) si :

$$\frac{P(x_1,\ x_2, \dots, x_k|H_0)}{P(x_1,\ x_2, \dots, x_k|H_1)} \leq k_0 \tag{3.10}$$

Nous sommes dans la zone d'incertitude si :

$$k_0 < \frac{P(x_1,\ x_2, \dots, x_k|H_0)}{P(x_1,\ x_2, \dots, x_k|H_1)} < k_1 \tag{3.11}$$

Considérons un échantillon de taille $n, X_n = (x_1, x_2, \dots, x_n)$.

Sur la base du principe de double seuillage et de l'option de rejet, le SPRT définit deux seuils k_0 (seuil bas) et k_1 (seuil haut) autour de la frontière k (voir théorème *de Neyman-Pearson*) tels que :

$$k_0 = \frac{\alpha}{1-\beta} \tag{3.12}$$

et

$$k_1 = \frac{1-\alpha}{\beta} \tag{3.13}$$

En supposant l'hypothèse d'indépendance sur les échantillons testés [52], le rapport de vraisemblance est calculé comme suit :

$$\frac{P(x_1,\ x_2, \dots, x_k|H_0)}{P(x_1,\ x_2, \dots, x_k|H_1)} = \frac{P(x_1,\ x_2, \dots, x_{k-1}|H_0)}{P(x_1,\ x_2, \dots, x_{k-1}|H_1)} \frac{P(x_k|H_0)}{P(x_k|H_1)} = \prod_{i=1}^{k} \frac{P(x_i|H_0)}{P(x_i|H_1)} \tag{3.14}$$

Le test séquentiel peut être résumé par l'équation de la zone d'incertitude à l'étape k comme suit :

$$\log\left(\frac{\alpha}{1-\beta}\right) < \sum_{i=1}^{k} \log\left(\frac{P(x_i|H_0)}{P(x_i|H_1)}\right) < \log\left(\frac{1-\alpha}{\beta}\right) \tag{3.15}$$

Nous pouvons faire une representation graphique du SPRT à travers la figure 4.2.

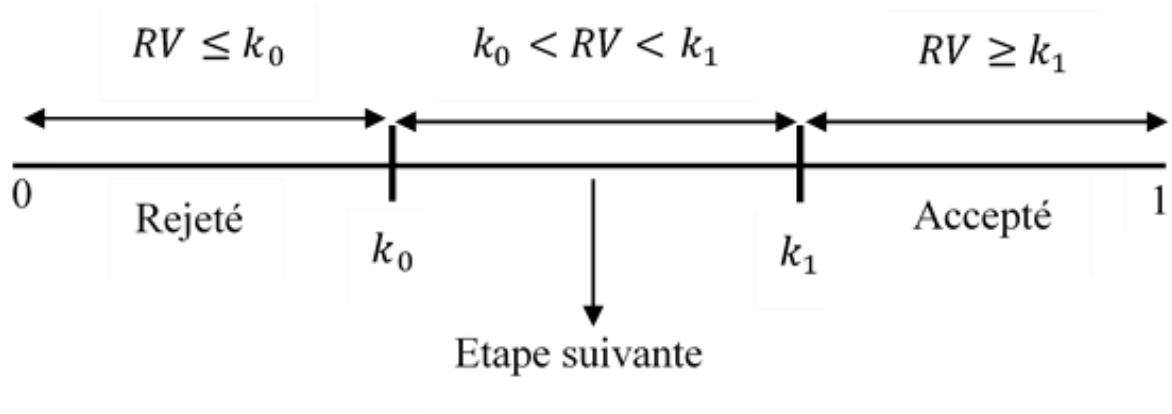

Figure 4.2 : Représentation graphique du test séquentiel du rapport de vraisemblance

La méthode de fusion séquentielle des scores met en œuvre plusieurs modalités de biométrie pure. Cela donne de meilleurs résultats par rapport à la méthode traditionnelle de fusion en série. Ces meilleurs résultats concernent d'une part l'amélioration des performances de reconnaissance et d'autre part la réduction du temps et la difficulté d'utilisation de plusieurs systèmes biométriques [52].

4.3. Architecture pour la fusion séquentielle adaptée de scores

Afin de capitaliser sur les deux principes décrits dans les sections 4.1 et 4.2 à travers les équations (3.1) à (3.15), nous proposons un nouveau modèle pour la fusion de scores (figure 4.3).

Soit $X_m = (x_1, x_2, \dots, x_m), m > 1$ un vecteur de caractéristiques biométriques pures ;

Soit $Y_{mn} = \begin{pmatrix} y_{11} & \cdots & y_{1k} & \cdots & y_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{l1} & \cdots & y_{lk} & \cdots & y_{ln} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{m1} & \cdots & y_{mk} & \cdots & y_{mn} \end{pmatrix}, n \geq 1$ la matrice des caractéristiques des métadonnées (ou de la biométrie douce); pour $j = (1, 2, \dots, n)$, soit $y_{ij} = (y_{i1}, y_{i2}, \dots, y_{ik}, y_{ik+1}, y_{ik+2}, \dots, y_{in})$ l'ensemble des attributs de métadonnées, où y_{i1} à y_{ik} sont des variables continues (taille, poids, couleur de peau, etc.) et y_{ik+1} à y_{in} sont des variables discrètes (race, marques, maquillage, etc.) ;

soit $\omega_r = (\omega_1, \omega_2, \dots, \omega_q)$ les utilisateurs inscrits dans la base de données; sur la base de l'attribut x_i un utilisateur ω_r est reconnue par la probabilité à priori exprimée comme suit : $P(\omega_r|x_i)$;

A partir des équations (3.1) à (3.14), le rapport de vraissemblance est calculé comme suit :

$$RV = \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{P(x_i, y_{ij}|H_o)}{P(x_i, y_{ij}|H_1)} \tag{3.16}$$

En considérant l'équation (3.15), le test séquentiel peut être résumé par l'équation de la zone d'incertitude à l'étape k comme suit :

$$\log\left(\frac{\alpha}{1-\beta}\right) < \sum_{i=1}^{k} \sum_{j=1}^{n} \log\left(\frac{P(x_i, y_{ij}|H_0)}{P(x_i, y_{ij}|H_1)}\right) < \log\left(\frac{1-\alpha}{\beta}\right) \tag{3.17}$$

En référence au schéma de pondération utilisé dans l'équation (3.5), la fonction de discrimination est définie par l'expression suivante :

$$g_{ij}(x_i, y_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n}[a_0 \log P(x_i|H_0) - a_0 \log P(x_i|H_1) + a_j \log P(y_{ij}|H_0) - a_j \log P(y_{ij}|H_1)] \quad (3.18)$$

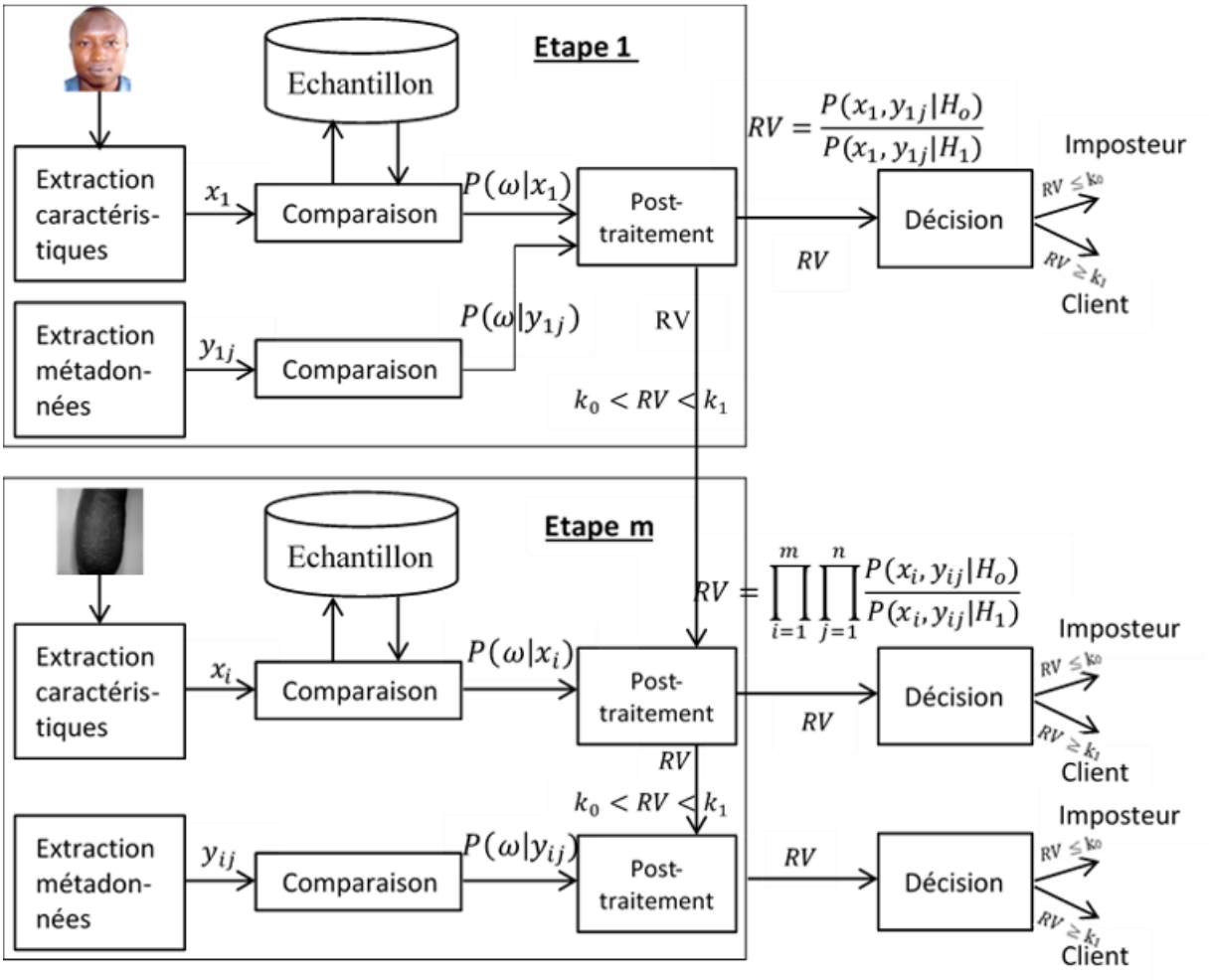

Figure 4.3 : Cadre proposé pour la fusion séquentielle adaptée

Sur la base du framework ci-dessus, nous proposons une nouvelle architecture de fusion de scores (voir figure 4.4), appelée *stratégie de fusion séquentielle adaptée*. Sur la figure 4.4, MB signifie Modalité Biométrique; MD signifie Métadonnée, RV signifie Rapport de Vraissemblance et ⊕ représente le symbole de fusion.

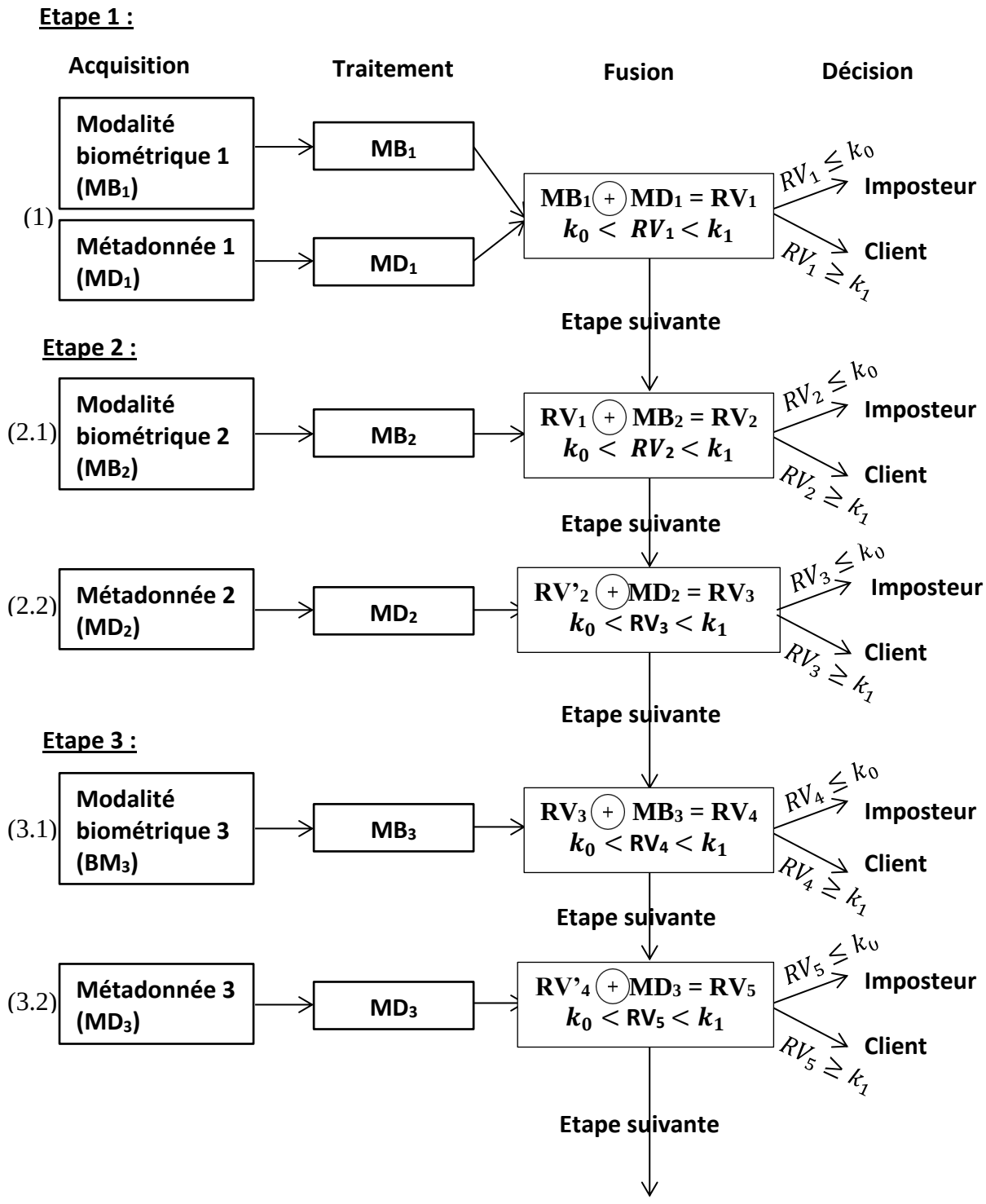

Figure 4.4 : Architecture de fusion séquentielle adaptée

Remarque :

Au niveau de la figure 4.4, il est important de préciser qu'à partir de l'étape 2 ($i = 2$), chaque première sous-étape (i.1) correspond uniquement au traitement de la modalité de biométrie pure ; ainsi, le poids $a_0 = 1$ et le score obtenu (rapport de

vraisemblance) est noté RV_i. Le passage aux deuxièmes sous-étapes notées i.2 implique la prise en compte des métadonnées et oblige à recalculer le rapport de vraisemblance de l'étape i.1 noté RV'_i avec $a_0 < 1$.

4.4. Algorithme de fusion séquentielle adaptée de scores

Pour une meilleure description de l'architecture de fusion séquentielle adaptée, un cadre a été conçu (figure 4.3). L'algorithme dérivé du cadre proposé permet de définir les variables d'entrée et de sortie avant la présentation des instructions sur 59 lignes (voir figure 4.5) [81]. Dans le but de déterminer la performance de notre algorithme de fusion séquentielle adaptée, nous nous sommes intéressés à sa complexité. Nous calculons à cet effet la complexité temporelle qui permet de mesurer sa vitesse d'exécution. En négligeant les affections et les instructions hors boucles, le pire des cas dans l'algorithme serait de parcourir toute la boucle "while ($i < m$)" et pour chaque itération de cette boucle, de parcourir totalement les deux boucles "for ($j = 1$ à n)" pour avoir un nombre d'opérations de $2 \times m \times n$. En ajoutant la première boucle "for ($j = 1$ à n)" on aura un nombre total d'opérations de : $n + 2 \times m \times n$. En négligeant le premier terme, on maintient le nombre d'opérations à $2 \times m \times n$. Ce qui nous permet d'obtenir une complexité temporelle de $O(n^2)$. Il s'agit d'une complexité de type quadratique.

Algorithm : Adapted sequential fusion

Input:

- Let PR (Probability Ratio) be a variable.
- Let x_i be the set of primary biometric features, with $i \leftarrow (1, 2, ..., m)$, $m > 1$.
- Let $y_{ij} \leftarrow (y_{i1}, y_{i2}, ..., y_{ik}, y_{ik+1}, y_{ik+2}, ..., y_{in})$ be the set of metadata features, with $j \leftarrow (1, 2, ..., n)$, $n \geq 1$.
- Let a_j, $j \leftarrow (0, 1, 2, ..., n)$ be the set of weights attributed to the biometric modalities with $\sum_{j=0}^{n} a_j \leftarrow 1$; $a_0 \gg a_j$, $j \leftarrow 1, 2, ..., n$; a_0 is attributed to the primary biometric modality ; a_j , $j \leftarrow 1, 2, ..., n$ are attributed to the soft biometrics modalities.
- Let $\alpha \leftarrow FRR$; $\beta \leftarrow FAR$; $k_0 \leftarrow \frac{\alpha}{1-\beta}$; $k_1 \leftarrow \frac{1-\beta}{\alpha}$.

Output:

- Let "Result" be a Boolean variable that takes the values "Genuine" or "Impostor"

LOOP Process

$i \leftarrow 1$
// In the next line, $a_0 < 1$
$PR \leftarrow a_0 \log P(x_i|H_0) - a_0 \log P(x_i|H_1)$
for $j = 1$ to n **do**
  $PR \leftarrow PR + a_j \log P(y_{ij}|H_0) - a_j \log P(y_{ij}|H_1)$
  $j \leftarrow j + 1$
end for
if $PR \leq k_0$ **then**
  Result $\leftarrow$ "Impostor"
  Break
else
  if $PR \geq k_1$ **then**
    Result $\leftarrow$ "Genuine"
    Break
  else
    $i \leftarrow i + 1$
    while $i \leq m$ **do**
      // In the next line, $a_0 \leftarrow 1$
      $PR \leftarrow PR + a_0 \log P(x_i|H_0) - a_0 \log P(x_i|H_1)$
      if $PR < k_0$ **then**
        Result $\leftarrow$ "Impostor"
        Break
      else
        if $PR \geq k_1$ **then**
          Result $\leftarrow$ "Genuine"
          Break
        else
          // In the next line, $a_0 < 1$
          $PR \leftarrow a_0 \log P(x_{i-1}|H_0) - a_0 \log P(x_{i-1}|H_1)$
          for $j = 1$ to n **do**
            $PR \leftarrow PR + a_j \log P(y_{i-1j}|H_0) - a_j \log P(y_{i-1j}|H_1)$
            $j \leftarrow j + 1$
          end for
          $PR \leftarrow PR + a_0 \log P(x_i|H_0) - a_0 \log P(x_i|H_1)$
          for $j = 1$ to n **do**
            $PR \leftarrow PR + a_j \log P(y_{ij}|H_0) - a_j \log P(y_{ij}|H_1)$
            $j \leftarrow j + 1$
          end for
          if $PR \leq k_0$ **then**
            Result $\leftarrow$ "Impostor"
            Break
          else
            if $PR \geq k_1$ **then**
              Result $\leftarrow$ "Genuine"
              Break
            else
              $i \leftarrow i + 1$
              if $i > m$ **then**
                Result $\leftarrow$ "Genuine"
                // At the last step we choose Genuine for personnal applications and Impostor for critical applications
              end if
            end if
          end if
        end if
      end if
    end while
  end if
end if
return Result

Algorithme 1 : Algorithme de fusion séquentielle adaptée

Conclusion

Les systèmes multibiométriques visent à améliorer la précision de reconnaissance d'un système d'authentification des personnes en combinant les preuves présentées par plusieurs sources d'information. La précision de reconnaissance peut être améliorée en ajoutant des informations auxiliaires au système. Ce chapitre a présenté un nouveau cadre de fusion séquentielle adaptée de scores afin d'améliorer la précision de reconnaissance des systèmes multibiométriques. Un algorithme est proposé pour cette architecture de fusion. L'architecture séquentielle réduit le temps moyen de reconnaissance de 45% [52] et l'intégration de la biométrie douce induit une amélioration de la précision de reconnaissance de 5% [82]. La nouvelle architecture proposée combine ces deux principes et est donc conçue pour apporter des améliorations à la fois en termes de temps de reconnaissance et de précision de reconnaissance. Le framework fera l’objet d’une implémentation dans la troisième partie. Au niveau du chapitre suivant, nous allons effectuer une étude sur la typologie des métadonnées en lien avec les vulnérabilités des systèmes biométriques.

Chapitre 5 : Typologie des métadonnées et vulnérabilités biométriques

Sommaire

Introduction

Bien qu'étant plus fiables et plus sécurisées que les méthodes traditionnelles d'authentification, les techniques biométriques sont l'objet de vulnérabilités qui constituent des défis à relever. Face à la multiplication des cas d'usurpation d'identité et de fraudes, la biométrie est de plus en plus utilisée pour assurer la protection des biens et des personnes dans plusieurs domaines tels que les applications commerciales, médico-légales et gouvernementales. Dans ce chapitre, nous présentons une analyse des causes profondes des vulnérabilités biométriques et fournissons une typologie détaillée des métadonnées dans l'adaptation biométrique. Dans un premier temps, une analyse des métadonnées a été effectuée. Cette analyse a permis de ranger les métadonnées en deux groupes, selon qu'elles sont liées à l'utilisateur ou au système d'acquisition. En ce qui concerne le système d'acquisition, deux cas sont à considérer. Il s'agit du niveau "capteur" et du niveau "environnement d'acquisition". Une fois l'univers des métadonnées clarifié, un focus a été mis sur leur rôle dans la lutte contre les vulnérabilités biométriques. Ainsi, les vulnérabilités étudiées ont été classifiées en deux grandes catégories que sont les limites intrinsèques et les attaques adverses. De façon détaillée, les attaques adverses sont subdivisées en trois sous-groupes à savoir les problèmes d'administration, la non sécurisation de l'infrastructure et les failles biométriques. A tous ces différents niveaux, le rôle possible des métadonnées a été mis en exergue avec des scénarios qui peuvent

être envisagés. Dans les chapitres 6 et 7, l'un des scénarios envisagés sera mis en œuvre, notamment le cas de la combinaison de la couleur de peau à la reconnaissance faciale.

5.1. Mode opératoire de la sécurité informatique

La transformation digitale en plein essor est un processus qui consiste à intégrer complètement les technologies digitales dans l'ensemble des activités d'une organisation. « Dans ce monde numérique, l'étape d'authentification est souvent considérée comme le maillon faible de la sécurité informatique, à cause du nombre important des cas d'usurpation d'identité » a affirmé Belguechi dans son manuscrit de thèse soutenu en 2015. L'authentification des utilisateurs représente alors l'un des défis les plus importants dans la sécurisation des biens. En référence à la sécurité informatique, la biométrie concerne l'utilisation des caractéristiques morphologiques, comportementales, biochimiques et bioélectriques pour déterminer ou vérifier l'identité d'un utilisateur.

La sécurité consiste à assurer la "protection" des "biens" contre toutes formes de "menaces" [83]. La menace représente le type d'action susceptible de nuire dans l'absolu, tandis que la vulnérabilité (appelée parfois faille ou brèche) représente le niveau d'exposition face à la menace dans un contexte particulier. Enfin la contre-mesure est l'ensemble des actions mises en œuvre en prévention de la menace. Les contre-mesures à mettre en œuvre ne sont pas uniquement des solutions techniques mais également des mesures de formation et de sensibilisation à l'intention des utilisateurs, ainsi qu'un ensemble de règles clairement définies. A partir des notions de menace, de vulnérabilité et de contre-mesure, nous pouvons définir la notion de risque comme étant un danger susceptible de subvenir à tout moment. Le risque en terme de sécurité est généralement caractérisé par l'équation suivante :

$$Risque = \frac{Menace \times Vulnérabilité}{Contre\_mesure}$$

Figure 5.1 : Equation du risque [83]

Le mode opératoire de la sécurité peut être représenté par la figure 5.2.

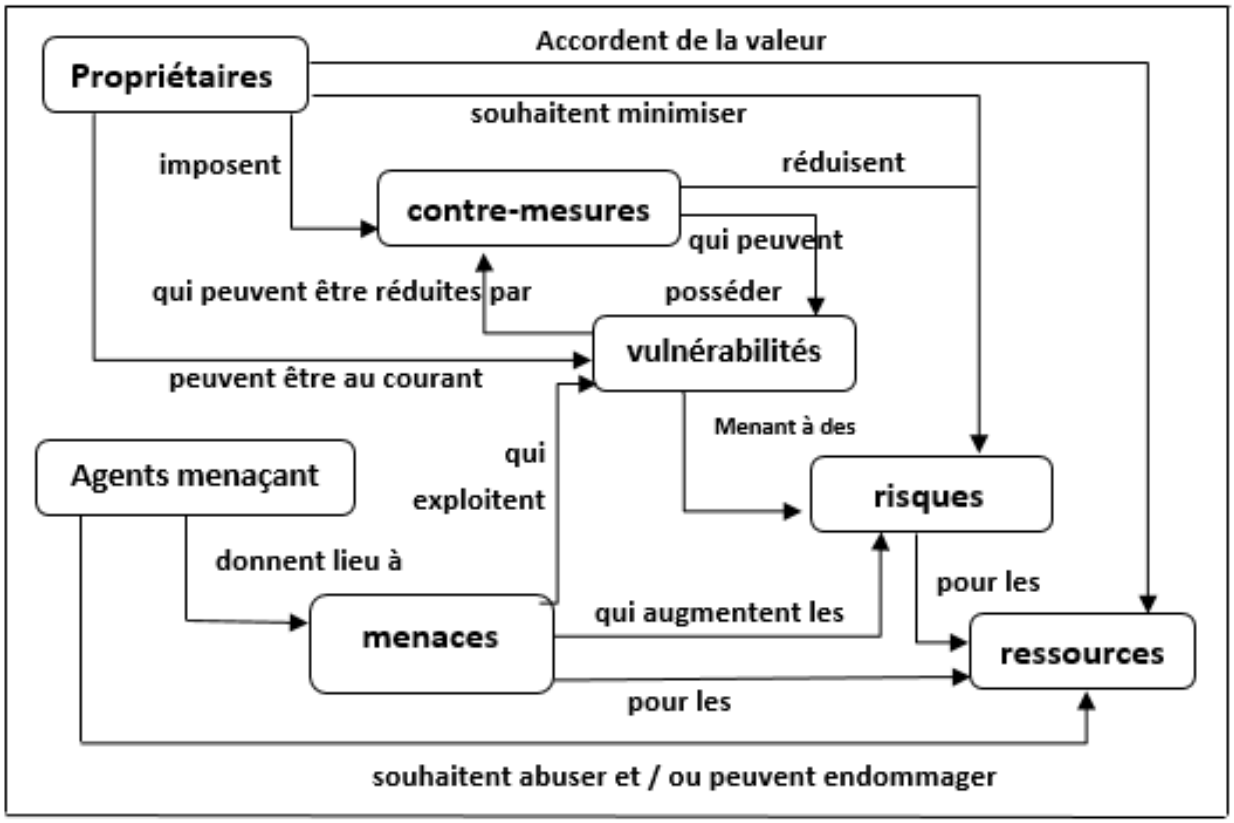

Figure 5.2 : Mode opératoire de la sécurité (adapté de [83])

Comme on peut le constater, les vulnérabilités se situent au cœur du mode opératoire de la sécurité. Les systèmes d'authentification biométrique se positionnent comme les moyens de contre-mesure mis en place pour assurer la protection des biens. Mais ces systèmes font face à un certain nombre de défis et de vulnérabilités. Parmi les solutions à ces défis et vulnérabilité, l'utilisation des métadonnées figure en bonne place. Nous présentons dans les sections suivantes un aperçu sur les métadonnées en biométrie avant d'aborder leur rôle pour faire face aux vulnérabilités identifiées.

5.2. Analyse des métadonnées et typologie de l'adaptation en biométrie

Les métadonnées ou informations auxiliaires sont des données non biométriques qui sont soit combinées avec les données de biométrie pure (empreinte digitale, visage, iris, signature, etc.) ou exploitées par le système pour faire de l'adaptation en biométrie. L'objectif de l'adaptation est de proposer des moyens pour faire face aux défis des systèmes biométriques.

En général, il existe deux grandes méthodes pour adapter les systèmes biométriques, la première liée à l'utilisateur (données de biométrie douce) et la seconde liée au système d'acquisition. Dans le premier cas concernant les données de biométrie douce, il existe une multitude de traits provenant de l'utilisateur qui peuvent être extraits et combinés avec les modalités de biométrie pure. Jain et al. [82] furent parmi les premiers à explorer ce domaine. Ils ont fait des expérimentations sur les traits suivants : Taille, genre et ethnie. Plus tard, Dantcheva et al. [84] ont exploré les effets d'autres traits de biométrie douce tels que: Age, couleur de peau, couleur des cheveux, couleur des yeux, marques, etc.

Il faut noter qu'en fonction des auteurs, certains traits sont classés tantôt parmi les modalités de biométrie pure ou parmi celles de biométrie douce. C'est le cas de la démarche qui est tantôt considérée comme une modalité comportementale [2] et tantôt comme un trait de biométrie douce [84].

Pour ce qui est du système d'acquisition, l'adaptation peut se faire d'une part au niveau du capteur (volume, flash, etc.) et d'autre part en prenant en compte le contexte d'opérationnalisation (luminosité, bruit, etc.). Ici, les métadonnées ne sont pas directement fusionnées avec les données de biométrie pure. La prise en compte se fait selon deux approches. Soit l'administrateur du système procède à des réglages des différents paramètres tels que le volume, la luminosité ambiante, le flash, etc. Soit par contre, le dispositif d'acquisition est pré-dimensionné pour assurer de façon automatique le contrôle de ces différents paramètres. Un tel

dispositif sera suffisamment intelligent pour ne pas déclencher par exemple une prise d'image du visage lorsque la luminosité n'est pas à un niveau minimal requis.

Par ailleurs, certains systèmes biométriques intègrent la qualité des données dans le processus de fusion. L'objectif est (a) d'attribuer automatiquement des poids aux modalités participantes, atténuant ainsi les erreurs introduites par des données d'entrée de mauvaise qualité [85], ou (b) d'invoquer de manière appropriée les modalités en cascade, maximisant ainsi la précision de la reconnaissance [86]. Les données de biométrie douce et les indices de qualité sont appelés informations auxiliaires dans le contexte de la fusion biométrique.

Selon Allano [52], les systèmes biométriques monomodaux ou multimodaux sont généralement ajustés en fonction de l'application. Les ajustements se font à plusieurs niveaux qui correspondent aux différents modules de la structure d'un système biométrique (voir figure 2.9). Une fois que l'application a été définie et les modalités choisies, les modules à définir sont l'extraction des caractéristiques, l'appariement, la fusion, le classement et la décision. Le paramétrage de ces modules dépend donc de l'application et de la population cible.

Selon Jain et Ross [1], l'adaptation en biométrie multimodale peut être réalisée à deux niveaux : (a) définition d'un seuil de décision spécifique à l'utilisateur et (b) attribution d'un coefficient de pondération pour chaque modalité biométrique. Les résultats expérimentaux prouvent que les seuils spécifiques à l'utilisateur améliorent les performances du système d'environ 2%, tandis que les pondérations spécifiques à l'utilisateur améliorent les performances d'environ 3%. Une autre façon d'adapter les systèmes consiste à prendre en compte des informations contextuelles (température, luminosité, bruit, etc.) ou des informations personnelles (couleur des yeux, âge, taille, etc.), appelées métadonnées [13,47,52,87].

Plusieurs travaux ont porté sur le rôle des métadonnées dans l'amélioration des performances biométriques. Dans la référence [82], les auteurs montrent les avantages d'utiliser les informations de l'utilisateur sur le genre, l'origine ethnique et la taille, en plus des empreintes digitales. L'utilisation de ces données de biométrie douce entraîne une amélioration d'environ 5% par rapport au système biométrique principal (basé sur l'empreinte digitale seule). Le tableau ci-dessous présente une typologie des métadonnées dans l'adaptation biométrique.

Tableau 5.1 : Typologie des métadonnées dans l'adaptation biométrique [6]

Niveau d'adaptation	**Utilisateur**			**Système d'acquisition**	
Type d'adaptation	Seuil de décision	Système de pondération (basé sur qualité de donnée)	Biométrie douce	Capteur	Contexte
Exemple	Voix: 0,7 Iris: 0,6	Empreinte digitale: 0,6 Face: 0,4	Genre; Taille; Couleur de peau, des yeux et des cheveux	Volume (microphone); Flash (caméra)	Luminosité, bruit, température
Donnée manipulée	Donnée biometrique (image ou signal)		Metadonnée		
Nécessité d'expérimentation	Après expérimentation		Sans expérimentation (immuable)		Après expérimentation

Dans la section suivante, un accent particulier sera mis sur l'effet des métadonnées dans la lutte contre les vulnérabilités biométriques.

5.3. Rôle des métadonnées face aux défis et vulnérabilités biométriques

Les défis des systèmes biométriques se situent essentiellement au niveau des limites en termes de performance et d'acceptabilité [6]. Par ailleurs,

l'authentification biométrique est confrontée à des vulnérabilités qui ont été identifiées à huit niveaux sur une architecture biométrique générique [88]. Jain et al. [89] ont ensuite listé au total quatre différents niveaux de vulnérabilités des systèmes biométriques. Il s'agit des limites intrinsèques, des problèmes d'administration, la non sécurisation de l'infrastructure et les failles biométriques. Les trois derniers niveaux ont été regroupés dans les attaques adverses, provenant de l'extérieur du système. Ces quatre niveaux de vulnérabilité ont été synthétisées à travers un modèle en arrêtes de poisson (voir figure 5.3).

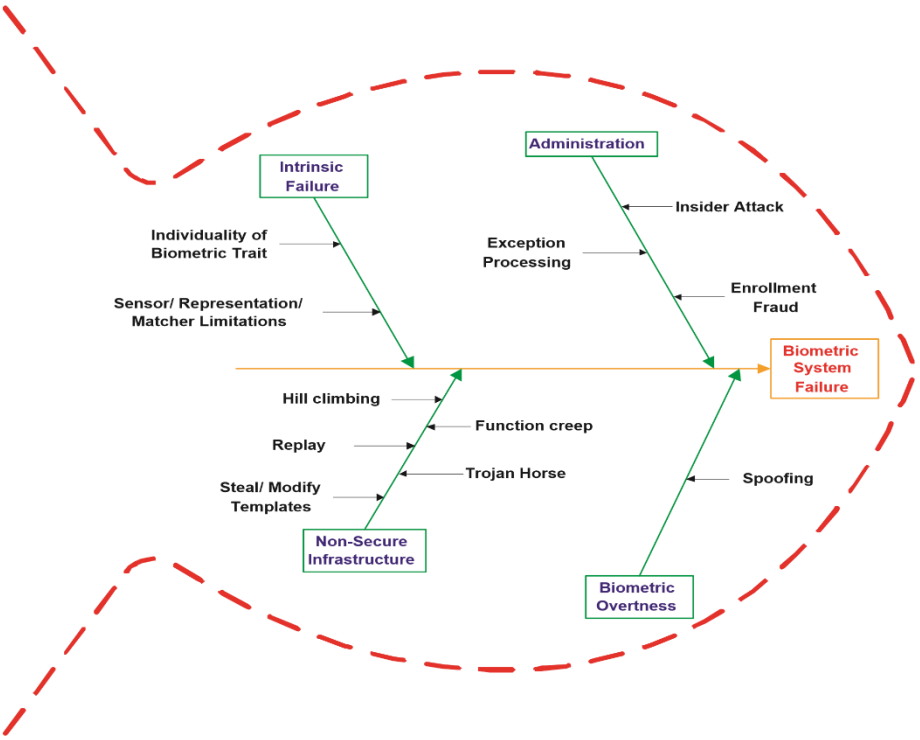

Figure 5.3 : Modèle en arrête de poisson pour catégoriser les vulnérabilités des systèmes biométriques [89]

5.3.1. Les défaillances intrinsèques des systèmes biométriques

Une défaillance intrinsèque est une défaillance de la sécurité du système biométrique qui engendre une décision incorrecte prise par ce dernier de façon native. Elle peut survenir même en l'absence d'un effort explicite de la part d'un attaquant pour contourner le système. Ce type d'échec est également connu sous le nom de "attaque-zéro-effort". Un système de vérification biométrique peut commettre deux types d'erreurs dans la prise de décision. Il s'agit des fausses

acceptations et des faux rejets. Dans le premier cas, les fausses acceptations sont généralement causées par un manque d'unicité dans le trait biométrique, ce qui peut entraîner une grande similitude entre les caractéristiques d'utilisateurs différents. Les variations intra-utilisateur et la similarité entre utilisateurs peuvent également être causées par l'utilisation de caractéristiques non saillantes et d'appariants non robustes. A ce niveau, l'ajout de métadonnées peut s'avérer efficace pour lutter contre ce problème. Parmi les cas les plus probables, il y a ceux de deux vrais jumeaux ou des frères qui présentent une parfaite similitude du visage par exemple. La prise en compte des données de biométrie douce telles que les marques et le maquillage permettra de faire la distinction entre deux individus même s'ils ont exactement le même visage car les algorithmes de reconnaissance de visage ne tiennent pas compte de ces données accessoires.

Un utilisateur légitime peut être faussement rejeté par le système biométrique en raison des grandes différences entre le modèle stocké de l'utilisateur et les ensembles de caractéristiques biométriques d'entrée. Ces variations intra-utilisateur peuvent être dues à une interaction incorrecte de l'utilisateur avec le système biométrique. C'est le cas par exemple de l'effet des torsions sur une image d'empreinte digitale ou l'effet des changements de pauses sur une image de visage. Les métadonnées n'apportent pas de solution spécifique pour ces cas de variations intra-utilisateur. Mais de plus en plus, les algorithmes de vérification biométrique prennent en compte ces problématiques. Djara et al. [19] ont proposé un algorithme de reconnaissance par l'empreinte digitale sans contact qui adresse cette question. Sur un autre plan, il est à noter que les faux rejets peuvent être dus au bruit introduit au niveau du capteur comme par exemple, les traces résiduelles laissées sur un capteur d'empreinte digitale. Les métadonnées n'apportent pas une solution à cette limite. La biométrie multimodale permet d'offrir une alternative aux cas de faux rejets. En effet, lorsque l'une des modalités est affectée de bruit, le système est capable de l'ignorer et de passer à la modalité suivante.

Parfois, un capteur peut ne pas acquérir le trait biométrique d'un utilisateur en raison des limites de la technologie de détection ou des conditions environnementales défavorables. Par exemple, lorsqu'un utilisateur se présente devant un capteur d'empreinte digitale avec des doigts secs ou mouillés, le système ne sera pas capable d'acquérir des images de bonne qualité. De même, un système d'authentification par la voix sera perturbé par l'émission de bruit sonore provenant de l'environnement d'opération. Cela entraîne des erreurs de non-inscription (FTE) ou de non-acquisition (FTA). Pour les limites liées à la technologie de détection, les métadonnées exploitées peuvent être le volume (cas de la modalité voix) ou le flash (cas du visage, de l'empreinte digitale sans contact, etc.). Pour les limites du contexte d'opération, les paramètres tels que la luminosité, le bruit et la température sont utilisés comme métadonnées pour faire face aux erreurs intrinsèques.

5.3.2. Les attaques adverses

Une personne externe mal intentionnée s'organise pour lancer une attaque contre le système biométrique en exploitant une multitude d'opportunités. Les trois formes d'attaques possibles sont : attaques de l'administration, attaques liées à l'infrastructure non sécurisée et attaques liées aux failles biométriques.

5.3.2.1. Attaques liées à l'administration du système

C'est une attaque encore appelée attaque d'initié. Elle provient en général de personnes familières malveillantes qui exploitent des vulnérabilités créées suite à une mauvaise administration du système biométrique. Il s'agit entre autres de l'intégrité du processus d'inscription (par exemple, la validité des justificatifs d'identité présentés lors de l'inscription), la collusion (ou la contrainte) entre l'adversaire et l'administrateur système ou un utilisateur légitime et l'abus des procédures de traitement des exceptions.

Les initiés sont souvent familiers avec les données de l'organisation et les méthodes qui sont en place pour les protéger. Cela rend plus facile pour l'initié de contourner les contrôles de sécurité dont ils ont connaissance. Les menaces internes sont plus difficiles à défendre contre les attaques que de l'extérieur, car l'initié a déjà accès légitime à l'information et les biens de l'organisation. Dans ces conditions, le recours aux métadonnées n'apportera pas de parade à ce type d'attaque dans l'absolue. Par contre, l'utilisation additionnelle des traits de biométrie douce peut rendre le système plus complexe et requérir plus d'effort de la part de l'attaquant.

5.3.2.2. L'infrastructure non sécurisée

L'infrastructure d'un système biométrique comprend le matériel, les logiciels et les canaux de communication entre les différents modules. Un adversaire peut manipuler l'infrastructure biométrique de plusieurs manières, ce qui peut conduire à des violations de la sécurité. Ratha et al. [88] ont regroupé les points d'attaques sur un système biométrique générique en 8 niveaux. La figure 4.4 illustre les emplacements possibles de ces attaques dans un système biométrique générique :

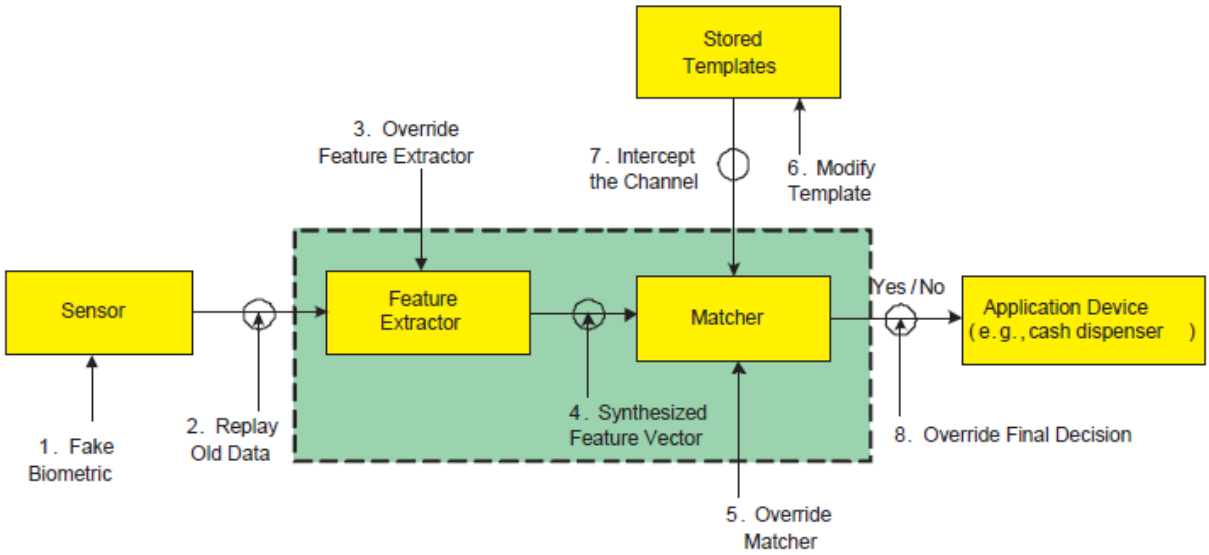

Figure 5.4 : Emplacements possibles d'attaques dans un système biométrique générique [89]

Premier niveau :

Données biométriques falsifiées : une reproduction de la donnée biométrique utilisée sera présentée au capteur biométrique. Dans le cas de l'authentification par l'empreinte digitale, l'attaquant peut présenter un faux doigt face à un capteur avec contact ou présenter juste l'image d'un doigt face à un capteur sans contact. Il faut déjà noter que cette forme d'attaque est plus courante avec l'empreinte digitale qui est une modalité très utilisée avec un pouvoir discriminatoire élevé. Par contre avec d'autres modalités telles que le visage, la rétine, la signature, l'ADN, il est plus difficile de présenter une donnée falsifiée. Par ailleurs, si le système biométrique intègre la prise en compte d'un trait de biométrie douce à combiner avec la biométrie pure, cette forme d'attaque devient plus compliquée à mettre en œuvre. Prenons par exemple un cas d'authentification par l'empreinte digitale et la taille de l'utilisateur. Même si l'empreinte digitale est falsifiée, l'imposteur sera bloqué au moment d'enregistrer la taille. En considérant un autre cas d'authentification par le visage combinée avec la couleur de peau extraite de façon automatique à partir de l'image du visage, un imposteur aura du mal à falsifier un visage tout en prenant en compte les spécificités et les exigences liées à la détection de la couleur de peau.

Deuxième niveau :

Transmission de données biométriques interceptées : Ici, l'attaquant rejoue une ancienne donnée biométrique enregistrée dans le système sans passer par le capteur biométrique. C'est le cas avec la présentation d'une ancienne copie de l'image de l'empreinte. Etant donné que l'attaquant contourne le capteur biométrique en fournissant au système une ancienne donnée enregistrée, les métadonnées n'auront aucun effet contre cette forme d'attaque.

Troisième niveau :

Attaque sur le module d'extraction de paramètres : ce module pourrait être remplacé par un cheval de Troie de manière à produire des informations choisies

par l'attaquant. L'utilisateur légitime ne se rend pas compte que ce module est a été corrompu et fourni des informations conformes aux directives du pirate. Le module d'extraction de paramètres étant compromis par le pirate, les métadonnées ne seront pas efficaces face à ce genre d'attaque.

Quatrième niveau :

Altération de paramètres extraits : après l'obtention de données par le module d'extraction de paramètres, celles-ci sont altérées voire remplacées par d'autres données définies par l'attaquant. Pour les attaques liées à l'infrastructure non sécurisée, à partir du deuxième niveau d'attaque jusqu'au huitième niveau, nous sommes dans des situations où le système biométrique est corrompu et ne fournira que des réponses conformément au dessein du pirate. Les métadonnées ne seront pas d'une efficacité dans ces contextes.

Cinquième niveau :

Le module de calcul de similarité est remplacé par un module malveillant : ce module pourrait être remplacé par un cheval de Troie afin de produire artificiellement de hauts ou bas scores.

Sixième niveau :

Altération de la base de données : la base de modèles biométriques est disponible localement, à distance ou distribuée sur plusieurs serveurs. Dans ce type d'attaque, l'attaquant modifie un ou plusieurs modèles afin d'autoriser un imposteur voire d'empêcher un utilisateur légitime d'y accéder.

Septième niveau :

Attaque sur le canal entre la base de données et le module de calcul de similarité: dans ce type d'attaque, les modèles sont altérés sur le lien de transmission reliant la base de modèles et le module de calcul de similarité.

Huitième niveau :

Altération des décisions (accepté ou rejeté) : ce type d'attaque altère la décision booléenne (oui ou non) pris par le module de calcul de similarité. La dangerosité de cette attaque est élevée puisque même si le système est robuste en terme de performance, il a été rendu inutile par ce type d'attaque.

5.3.2.3. Ouvertures/failles biométriques

Une personne malveillante peut acquérir discrètement les caractéristiques biométriques d'un utilisateur légitime et de les utiliser pour créer une production artificielle du trait biométrique (empreintes digitales prélevées à la surface d'un capteur avec contact). Dès lors, si le système biométrique n'est pas en mesure de faire la distinction entre une présentation biométrique en direct et une parodie artificielle, un adversaire peut contourner le système en présentant des traits falsifiés. Ce cas rejoint parfaitement le premier niveau d'attaque sur l'infrastructure non sécurisée. L'analyse effectuée à ce niveau est également valable ici. Il convient de préciser que la biométrie douce représente une parade à cette forme d'attaque pour deux raisons. En effet, l'utilisateur est obligé de se présenter en personne devant le dispositif d'acquisition de données ou bien cette donnée est automatiquement extraite à partir d'une modalité de biométrie pure ayant un certain nombre de caractéristiques difficilement respectables par les productions artificielles.

Le tableau ci-dessous présente de façon synthétique pour chaque catégorie de vulnérabilité, l'existence d'effet ainsi que les actions possibles des métadonnées. Pour chaque type de vulnérabilité, des exemples de métadonnées utilisables ont été énumérés.

Tableau 5.2 : Typologie de l'action des métadonnées sur les vulnérabilités biométriques

Catégorie de vulnérabilité	**Echecs intrinsèques**	**Attaques adverses**		
		Administration	**Infrastructure non sécurisée**	**Failles biométriques**
Existence d'effet des métadonnées	Oui (excepté le cas de faux rejets)	Partiel	Oui (1 cas sur 8)	Oui
Actions possibles	1- Réduisent les fausses acceptations 2- Limitent les non-inscriptions et les non-acquisitions	Rendent le système plus difficile à attaquer	Empêchent l'attaque par présentation de données biométriques falsifiées	Empêchent l'attaque par présentation de données biométriques falsifiées
Exemples de métadonnées	1- Marques (tatouages, scarifications, etc.) ; maquillage 2- Volume, flash ; Luminosité, bruit, température	Poids, les formes du corps, etc.	Taille, couleur des yeux, etc.	Démarche, genre, couleur de peau, etc.

Conclusion

Dans le but de répondre aux besoins liés à la fiabilité dans l'authentification et l'identification des personnes, la biométrie se présente comme une technologie efficace. Elle est mise en œuvre dans de nombreuses applications commerciales, gouvernementales et médico-légales par l'utilisation des caractéristiques

physiologiques et comportementales spécifiques à chaque personne. Mais au-delà de leur efficacité, les techniques biométriques font face à des vulnérabilités qui nécessitent d'être étudiées afin d'apporter des solutions appropriées. Ce chapitre s'est consacré à la présentation d'une typologie complète des métadonnées sur les vulnérabilités des systèmes biométriques. En outre, nous avons discuté de la catégorisation des métadonnées biométriques, à savoir le niveau utilisateur (c.-à-d. biométrie douce) et le niveau du système d'acquisition (c.-à-d. le capteur et le contexte d'acquisition). Dans la dernière partie à venir, nous allons mettre en application les méthodes précédemment développées.

Troisième partie : Application et résultats expérimentaux

Chapitre 6 : Classification de la couleur de peau du visage humain

Sommaire

Introduction

La multi biométrie adaptée consiste à introduire des données auxiliaires dans le processus d'authentification biométrique permettant d'améliorer les performances de reconnaissance. Ces données auxiliaires peuvent provenir de l'utilisateur, du capteur ou de l'environnement d'acquisition. Dans le cas où les données auxiliaires proviennent de l'utilisateur, on parle de biométrie douce. [6] ont réalisé une typologie de l'adaptation biométrique. Dans la gamme variée de données d'adaptation biométrique, les données de biométrie douce font de plus en plus l'objet de recherches. La biométrie douce comprend, sans toutefois s'y limiter, la taille, le poids, le genre, la couleur de la peau, la couleur des cheveux, la couleur des yeux [90].

La couleur de la peau est l'un des moyens les plus remarquables de différenciation humaine et a été largement utilisée pour définir les races humaines [91]. Parmi ces modalités de biométrie douces, la couleur de peau apparaît comme pertinente en termes de pouvoir de discrimination. La couleur de peau étant une modalité de

biométrie douce, elle ne peut à elle seule permettre d'authentifier un individu de façon fiable mais elle peut être fusionnée avec d'autres modalités pures pour la réalisation d'un système multi-modal robuste. Nous proposons à travers ce manuscrit une nouvelle approche d'authentification par classification de la couleur de peau.

6.1. Les travaux antérieurs

La détection et la classification de la couleur de peau intéressent les scientifiques à cause de leurs nombreuses applications possibles. La détection peut être utilisée comme étape préliminaire dans certaines applications de vision par ordinateur comme la reconnaissance de nudité sur les sites web, la détection des visages, la détection des maladies de la peau, etc. Ainsi, dans le domaine de la détection de la couleur de peau plusieurs travaux ont été menées à travers différentes approches parmi lesquelles nous pouvons citer ceux de Singh et al. [92] qui ont utilisé la détection de la couleur de peau pour la localisation des visages sur des images. Pour cela, ils ont combiné les trois espaces colorimétriques RGB, YCbCr et HSI afin d'obtenir un nouvel algorithme de détection de visage basé sur la couleur de la peau et dont les résultats expérimentaux montrent une précision de 95,18%. Plus tard, Abd El-Hafeez [93] a proposé un système pour détecter les régions de couleur de peau dans les images extraites de documents PDF. Ce système est basé sur les espaces colorimétriques log opponent et HSV qui ont été modifiés afin d'améliorer les performances de détection. Les résultats obtenus indiquent que la technique du log opponent permet d'avoir une meilleure détection de la couleur de peau par rapport à la technique basée sur l'espace HSV. En 2015, [94] ont mis au point un algorithme de détection et de reconnaissance de la couleur de la peau utilisant une technique de cartographie de la couleur de la peau. Cet algorithme donne un résultat prometteur pour la valeur de seuil de Cb = [100, 127] et Cr = [130, 175] dans l'espace colorimétrique YCbCr. Une approche basée sur des seuils de détection dynamique a été proposée plus récemment par Patil et al. [95].

Les résultats expérimentaux obtenus grâce à leur méthode donnent une précision de 0,9857. Sur la base d'une combinaison des espaces colorimétriques RGB, HSV et YCbCr, Kolkur et al. [96] ont développé un système pour la détection de la couleur de peau. L'algorithme développé est capable de traiter des images sous diverses conditions d'éclairage telle que la luminosité, etc. La technique appelée RGB-H-CMYK, qui utilise trois espaces colorimétriques à savoir RGB, H (teinte du HSV) et CMYK a été développée par les auteurs [97] qui ont appliqué des règles basées sur des seuils dans leur méthode. A travers différentes combinaisons telles que RC (RGB et CMYK), RH (RGB et H) et RHC (RGB et H et CMYK), l'image d'entrée dans ces trois schémas de couleurs hybrides est explorée puis chaque pixel est qualifié de pixel de peau lorsqu'au moins deux règles votent en sa faveur. Cette technique de détection de couleur de peau a donné une précision de 89 % dans les résultats expérimentaux.

Quant à la classification de la couleur de peau, elle peut être utilisée comme métadonnées dans le domaine de la biométrie pour vérifier l'identité des personnes. Dans ce domaine, Yoon et al. [98] ont fait la description d'une technique automatique de classification de la couleur de la peau. Cette technique permet une modélisation précise des non-uniformités du teint, tout en évitant la contamination par les cheveux, les yeux, l'arrière-plan et les ombres. Quatre ans plus tard, Bhoyar et Kakde [99] ont présenté une approche de classification de la couleur de peau basée sur les pixels, pour détecter les pixels de peau et les pixels non liés à la peau dans les images couleur, en utilisant un classificateur symétrique à réseau neuronal. Leur classificateur a donné un taux de détection de plus de 90% avec 7% de faux positifs en moyenne. Même en présence de plusieurs groupes ethnique sur une même image, leur classificateur est capable de classer les pixels de peau. Il convient de mentionner également les travaux de Bianco et al. en 2015 [100] qui ont conçu une méthode de classification de la couleur de peau utilisant la détection automatique de visages et de corps dans une image, pour initialiser

de manière adaptative des classificateurs de peau ad-hoc individuels. Les expérimentations résultantes de leurs travaux montrent que le classificateur est moins dépendant des changements de couleur de peau en raison des niveaux de bronzage, des races, des genres et des conditions d'éclairage.

Toutes ces différentes méthodes bien qu'ayant leurs avantages et inconvénients ont permis d'avoir des résultats prometteurs dans le domaine de la détection et la classification de la couleur de peau.

6.2. Base de données utilisée

Etant donné que nous sommes dans un contexte de détection du visage par la couleur de peau, une base de données contenant des images en niveau de gris n'est pas appropriée. En outre, d'autres conditions nécessitent d'être respectées pour garantir une bonne représentativité des images de visage à utiliser. Il s'agit entre autres de la pose (différentes positions du visage), de l'expression (joie, colère, tristesse, etc.), de l'arrière-plan (présence ou non d'objets) ou de l'éclairage (conditions de luminosité). Cette contrainte a contribué au choix de la base de données de visages du Caltech [101]. Recueillie à l'institut de technologie de Californie, cette base de données comporte 450 images de visages de taille 896x592 au format JPEG. Il s'agit de captures de 27 personnes uniques sous différentes conditions d'éclairage et d'arrière-plans avec une variation d'expressions. Après analyse des données de cette base, nous pouvons présenter le tableau 6.1 :

Tableau 6.1 : Informations sur la base de données

Nombre de personnes	Nombre d'images
16	20 images au moins
19	17 images au moins
5	5 images exactement

2	1 image exactement
1	7 images exactement

Nous constatons que la base de données n'est pas uniforme en termes de nombre d'image par individu comme le montre le tableau 6.1. Afin d'uniformiser la base de données pour la suite des travaux, nous avons pris les 19 personnes ayant au moins 17 images et nous les avons tous ramenées à 17 images exactement. Nous utilisons 12 images pour la formation du modèle (apprentissage) et 5 pour les tests, soit 70,6 % et 29,4 % respectivement. Ainsi nous avons au total 323 images dont 228 pour la formation du modèle et 95 pour le test du modèle. Cette base de données utilisée normalement pour la reconnaissance faciale sera aussi utilisée dans le processus de classification de la couleur de peau. Voici ci-dessous un échantillon d'images de cette base de données montrant par exemple la diversité en termes d'expression et de conditions de capture.

Figure 6.1 : Echantillon d'images (source [101])

6.3. La détection du visage

La première étape dans le processus d'authentification par la couleur de peau est la détection du visage sur une image. La détection et la reconnaissance de visage à partir d'une image ou d'une vidéo est un sujet populaire dans la recherche biométrique [102]. Dans ce manuscrit, la détection de visage est réalisée en utilisant le framework python avec le package OpenCV (Open Source Computer Vision). Ce système contient trois modules que sont la détection, la formation et la reconnaissance. Fondamentalement, le module de détection détecte le visage qui est sur une image puis nous sélectionnons uniquement la zone de visage détectée (zone d'intérêt) en faisant un cadrage sur cette zone. La détection de visage utilise des classificateurs, qui sont des algorithmes qui détectent ce qu'est un visage ou non sur une image. Les classificateurs ont été formés pour détecter les visages à l'aide de milliers, voire de millions d'images afin d'obtenir une plus grande précision [103].

La figure 6.2 montre des exemples de détections de visages effectuées grâce au classificateur de visage choisi. Les figures 6.2.1 et 6.2.2 proviennent de la base de données décrite ci-dessus alors que les figures 6.2.3 et 6.2.4 proviennent de nos propres captures.

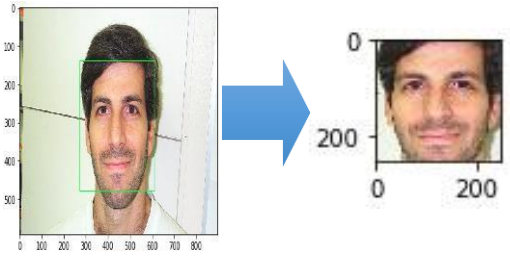

Figure 6.2.1 : individu 1

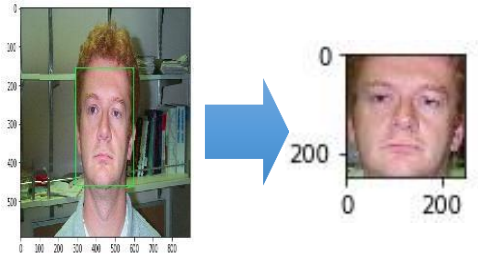

Figure 6.2.2 : individu 2

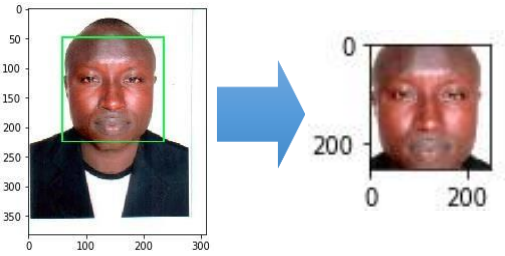

Figure 6.2.3 : individu 3

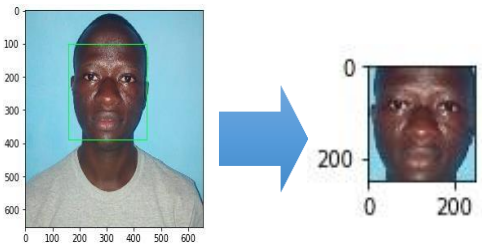

Figure 6.2.4 : individu 4

Figure 6.2 : Visages détectés et encadrés

6.4. La détection de la peau

La suite du processus consiste à faire une détection de la peau sur l'image afin d'éliminer toutes les parties inutiles et de garder uniquement les pixels de peau. La détection de peau est le processus de recherche de pixels et de régions de couleur de peau sur une image ou dans une vidéo [104]. Plusieurs travaux ont été effectués dans le domaine de la détection de la peau humaine sur des images ou des vidéos. Dans notre système, après extraction du visage sur une image on obtient une image contenant dans une large majorité des pixels de peau comme le montre la figure 6.2. Ainsi l'extraction du visage élimine une grande partie de pixels inutiles. En raison des variations des conditions d'éclairage, des paramètres matériels du capteur et de la plage de coloration de la peau chez les êtres humains, un modèle de couleur de peau prédéfini ne peut pas capturer avec précision la large distribution des couleurs de peau sur des images individuelles [105]. Il apparaît donc nécessaire de procéder au choix de l'espace colorimétrique approprié.

6.4.1. Choix de l'espace colorimétrique

Le choix de l'espace colorimétrique est très important dans le processus de détection de la peau. Lorsque l'espace colorimétrique RGB standard est utilisé, la détection de la peau peut être très complexe, difficile dans des conditions d'éclairage et de contraste variables. Par conséquent, l'image doit être convertie dans un autre espace colorimétrique invariant ou au moins insensible aux changements d'éclairage, tels que HSV [104]. Le modèle de couleur HSV est une représentation cylindrique du modèle RGB standard [106]. HSV signifie Teinte, Saturation et Valeur. La teinte est mesurée en degrés et varie de 0 à 360. Elle forme la couleur de base. La saturation et la valeur (luminosité) déterminent la proximité des blancs et des noirs respectivement. Dans le modèle de base, elles varient de 0 à 100, mais dans la bibliothèque OpenCV utilisée pour l'étape de détection de visage, elles varient de 0 à 255. Pour convertir l'image du modèle RGB en HSV, chaque pixel de l'image est soumis à la transformation suivante [104]:

-Il convient de trouver les valeurs maximales et minimales de R, G, B, C_{max} et C_{min} et de calculer leur différence M ;

-La Teinte H est calculée par la formule suivante :

$$H = \begin{cases} 0, & C_{max} = 0 \\ 60 \times \dfrac{G-B}{M}, & C_{max} = R \\ 60 \times \dfrac{B-R}{M} + 120, & C_{max} = G \\ 60 \times \dfrac{R-G}{M} + 240, & C_{max} = B \end{cases} \tag{5.1}$$

-La saturation S est calculée par la formule suivante :

$$S = \begin{cases} \dfrac{M}{C_{max}}, & C_{max} \neq 0 \\ 0, & C_{max} = 0 \end{cases} \tag{5.2}$$

-La Valeur V est calculée par la formule suivante :

$$V = C_{min} \tag{5.3}$$

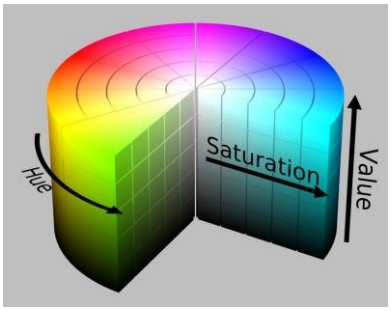

Figure 6.3 : Modèle de couleur HSV [107]

6.4.2. Segmentation de la couleur de peau par seuillage

La segmentation d'image est une opération de traitement d'image qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis. Les pixels sont ainsi regroupés en régions, qui constituent un pavage ou une partition de l'image. Segmenter une image en couleurs peut être extrêmement coûteux. Afin de simplifier les traitements, de nombreuses recherches se sont orientées vers la binarisation de l'image. Le seuillage global est l'une des méthodes de binarisation. Les techniques de seuillage global tentent de trouver une valeur de seuil unique appropriée (S) à partir de l'image globale. Les pixels sont séparés en deux classes en utilisant l'équation suivante [108] :

$$I_b(x,y) = \begin{cases} 0, & si\ \ I_f(x,y) \leq S \\ 1, & si\ I_f(x,y) > S \end{cases} \tag{5.4}$$

Avec x et y les coordonnées des pixels de l'image, $I_f(x,y)$ le pixel de l'image d'entrée et $I_b(x,y)$ le pixel de l'image binaire.

Dans notre système un intervalle de valeurs représente la couleur de peau et constituent une première classe. Les autres valeurs hors de cet intervalle constituent la deuxième classe. Suivant la formule de l'équation 5.4, les pixels se retrouvant dans cette plage (première classe) sont alors conservés et les autres sont éliminés. Ainsi la segmentation de la peau est effectuée à l'aide du seuillage empirique dans l'espace colorimétrique HSV grâce à une plage de couleur de peau prédéfinie. Pour les composantes H, S et V les valeurs des bornes inférieures et supérieures sont respectivement de : [0, 48, 80] et [20, 255, 255]. La figure 6.4 montre quelques résultats obtenus grâce à cette méthode. Sur la première ligne, nous avons les images de visages détectés (zone d'intérêt) et en dessous de chaque image d'entrée, nous avons l'image obtenue après segmentation.

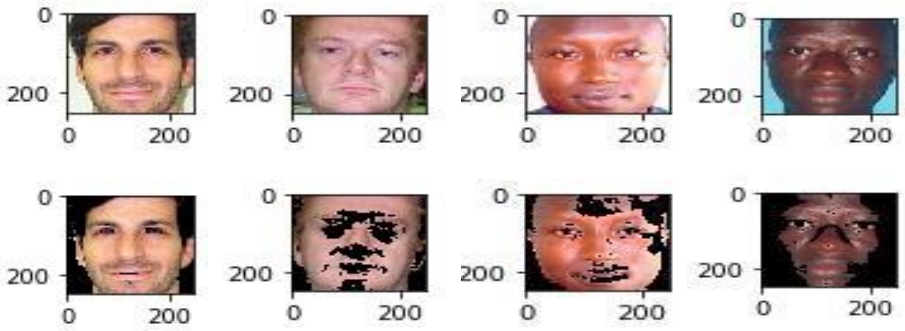

Figure 6.4 : Segmentation par seuillage de la peau

6.5. Extraction des couleurs dominantes sur la peau

Sur une même peau on peut observer des variations de couleurs d'une zone de la peau à une autre. Après détection de la peau du visage, on passe alors à la détection des couleurs dominantes sur la peau. Pour atteindre cet objectif nous utilisons l'algorithme des k-moyennes de la méthode du clustering.

6.5.1. Le clustering

Le clustering est une des méthodes d'analyse des données. Il vise à diviser un ensemble de données en différents « paquets » homogènes, en ce sens que les données de chaque sous-ensemble partagent des caractéristiques communes, qui correspondent le plus souvent à des critères de proximité (similarité informatique)

que l'on définit en introduisant des mesures et classes de distance entre objets. Ainsi le clustering signifie la création de groupes d'objets basés sur leurs caractéristiques de telle sorte que les objets appartenant aux mêmes groupes sont similaires et ceux appartenant à des groupes différents sont dissemblables [109].

6.5.2. L'algorithme des k-moyennes

Il s'agit de l'algorithme de clustering de partition le plus utilisé [110]. L'algorithme des k-moyennes commence par choisir k points représentatifs comme centroïdes initiaux. Chaque point est ensuite attribué au plus proche centroïde basé sur une mesure de proximité particulière choisie. Une fois les grappes formées, les centroïdes pour chaque cluster (groupe ou classe) sont mis à jour. L'algorithme répète ensuite ces deux étapes de manière itérative jusqu'à ce que les centroïdes ne changent pas et qu'aucun autre critère de convergence n'est rencontré.

Étant donné n points de données $x_1 \dots, x_n$, dans R^p et une k-partition $C = (C_1, \dots, C_k)$ de l'ensemble $O = \{1, \dots, n\}$ des "objets" sous-jacents avec des classes non vides $C_i \subset O$, le critère de variance ou d'inertie est donné par [111] :

$$g_n(C) := \sum_{i=1}^{k} \sum_{i \in c_i} \| x_l - x'_{c_l} \|^2 \rightarrow \min_{C} \tag{5.5}$$

Où x'_{c_l} désigne le centroïde des points de données x "appartenant" à la classe C_i (c'est-à-dire avec $l \in C_i$). Nous recherchons une partition k de O avec la valeur minimale du critère $g_n(C)$. On parle de minimisation de la distance intra-classe. La figure 5.5 ci-dessous illustre le processus de formation des groupes lors du déroulement de l'algorithme des k-moyennes.

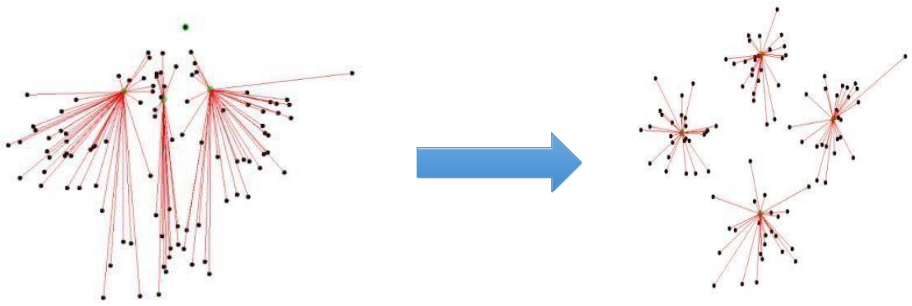

Figure 6.5 : Graphe du déroulement de l'algorithme des k-moyennes

Nous utilisons cette méthode de clustering afin de déterminer les groupes de pixels de peau ayant des couleurs semblables. Afin d'avoir une bonne séparation des groupes, il convient de bien choisir le nombre de clusters. En effet quand le nombre de cluster est faible, certains groupes normalement différents sont obligés de se mettre en ensemble pour former un nouveau groupe et cela ne permet donc pas une bonne séparation des groupes. Sur la figure 6.6, nous avons utilisé pour un même individu, différents clusters et généré les barres de couleurs. On constate alors que les couleurs obtenues varient et sont mieux séparées quand le nombre de clusters augmente.

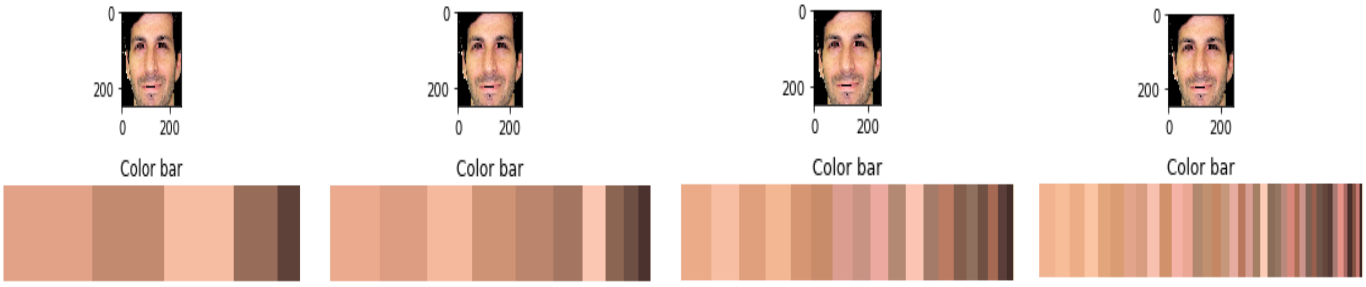

Figure 6.6 : Variation des couleurs en fonction du nombre de clusters

Cependant l'augmentation du nombre de clusters augmente également le temps d'exécution du programme. Le tableau 6.2 montre que le temps d'exécution est approximativement doublé lorsqu'on double le nombre de clusters.

Tableau 6.2 : Temps d'exécution de l'algorithme en fonction du nombre de clusters.

Nombre de clusters	Temps d'exécution en seconde
5	1.410046
10	3.229928
20	7.212621
40	14.290107

6.6. Choix du nombre optimal de clusters

Pour la suite de l'expérimentation, il convient de choisir le nombre de clusters à utiliser pour chaque image. Nous avons opté pour la détermination graphique de ce nombre. Les valeurs caractéristiques du taux de faux rejets (FRR) et du taux de fasses acceptations (FAR) ont été mesurées respectivement pour 10, 20, 30, 40, 50 et 60 clusters. La représentation graphique des courbes de ces deux taux est représentée sur la figure 6.7.

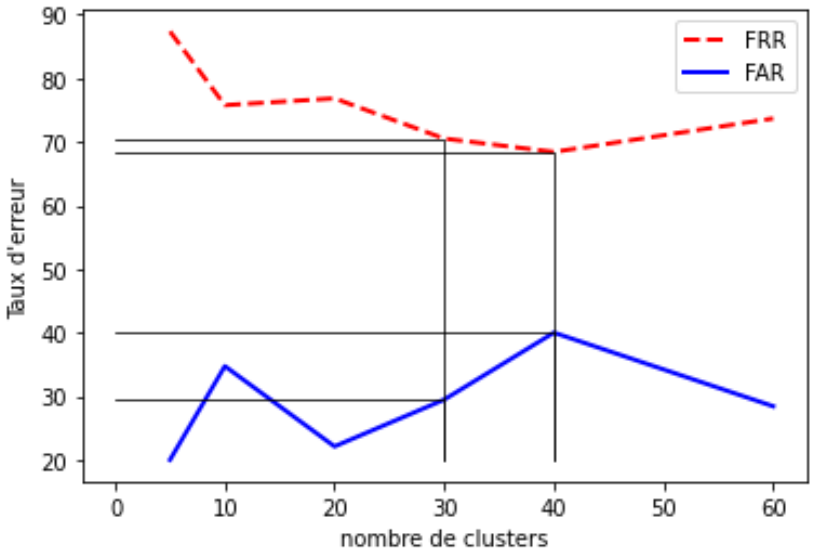

Figure 6.7 : Evolution des taux d'erreurs en fonction du nombre de clusters

Sur la figure 6.7, nous pouvons observer un rapprochement des deux courbes au niveau où le nombre de clusters est égal à 40. Cependant, le taux FAR étant élevé au niveau de 40 clusters, nous avons préféré utiliser 30 clusters. En effet, la variation du taux FRR est faible entre le niveau des 30 clusters et celui des 40 clusters. Par contre, le taux FAR est plus faible au niveau des 30 clusters. C'est ce qui justifie notre choix pour les 30 clusters. En prenant en compte les contraintes de temps d'exécution au niveau du tableau 6.2 et l'analyse de la figure 6.7, nous avons alors utilisé 30 clusters pour la suite de l'implémentation.

Ensuite les groupes (clusters) sont classés par ordre croissant de dominance sur la peau. Ainsi on arrive à avoir les couleurs de peau dominantes. La figure 6.8 illustre quelques résultats.

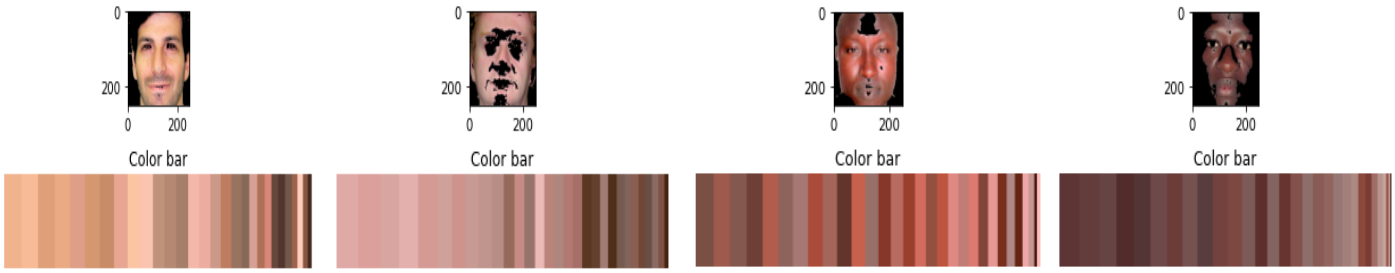

Figure 6.8 : Extraction de couleurs dominantes sur la peau

6.7. La prédiction de l'identité

Afin de pouvoir prédire l'identité d'un individu, nous formons d'abord un modèle qui enregistre les informations utiles relatives à l'identité de chaque personne dans une base de données. Pour un individu, la couleur de peau varie d'une région à une autre du visage encore que cette couleur est influencée par les conditions d'éclairage comme le montre la figure 6.9. Ainsi pour une même personne les composantes R, G et B de la couleur dominante sur la peau peuvent varier d'une image à une autre. On ne peut donc pas enregistrer une valeur de couleur fixe pour une personne. La variation de la couleur implique une variation des trois composantes R, G et B. Nous utilisons alors les variations suivant les composantes R, G et B pour prédire l'identité d'un individu. Pour la création du

modèle, nous utilisons les douze (12) images disponibles par individu dans la base de données et procédons comme suit pour chaque individu :

1. Pour chaque image, on prend les deux premières couleurs dominantes suivant chacune des trois composantes de l'espace colorimétrique RGB. On enregistre ces valeurs dans trois différentes tables pour les douze images. Ainsi chacune des trois tables contient 24 valeurs par composante, à raison de deux couleurs dominantes multipliées par les douze images.
2. Ensuite, on applique l'algorithme des k-moyennes sur chacune des trois tables afin d'avoir en sortie deux classes de valeurs par table. La première classe est celle des valeurs faibles et la seconde représente les valeurs élevées. On sélectionne ensuite le centroïde de chacune des classes obtenues. Cela permet d'avoir par classe une valeur représentative caractérisant la classe. Ainsi chacune des tables donne deux valeurs qui correspondent aux informations utiles sur chacune des composantes de l'espace colorimétrique RGB. Ces deux valeurs permettront d'avoir un intervalle de confiance suivant chacune des composantes R, G et B.
3. Les deux valeurs obtenues par composante sont enregistrées dans la base de données pour l'individu. On a alors au total six (6) valeurs enregistrées dans le modèle par individu.

Le tableau 6.3 présente les valeurs enregistrées dans le modèle pour les 19 utilisateurs de la base de données. Dans ce tableau, sur chaque ligne correspondant à un individu, les deux premières valeurs représentent l'intervalle de confiance pour la composante en R, les deux valeurs suivantes représentent l'intervalle de confiance pour la composante en G et les deux dernières valeurs représentent l'intervalle de confiance de la composante B. Après la création du modèle, vient alors la phase de prédiction de l'identité des individus. Lors d'une prédiction d'identité on détermine la couleur dominante sur l'image de visage en entrée et on vérifie si au moins deux des valeurs des composantes R, G et B obtenues

appartiennent aux intervalles de confiance enregistrés dans le modèle suivant chacune de ces composantes. Après expérimentation, le temps de traitement d'une image pour l'extraction de la couleur est de 7.21s.

Tableau 6.3 : Données enregistrées dans le modèle.

[197, 218, 161, 195, 153, 180]

[219, 115, 84 , 162, 169, 68]

[162, 233, 195, 127, 197, 135]

[222, 69 , 127, 183, 178, 78]

[231, 35 , 200, 138, 186, 33]

[240, 220, 202, 173, 171, 202]

[221, 152, 209, 149, 188, 125]

[220, 70 , 180, 78 , 57 , 180]

[216, 85 , 179, 163, 175, 159]

[212, 237, 230, 206, 199, 178]

[241, 117, 205, 101, 213, 129]

[220, 74 , 142, 38 , 157, 87]

[230, 186, 178, 148, 169, 194]

[214, 235, 145, 182, 167, 197]

[241, 10 , 181, 12 , 7 , 202]

[222, 172, 134, 177, 187, 149]

[250, 242, 190, 218, 223, 203]

[231, 86 , 64 , 212, 72 , 212]

[216, 14 , 175, 4 , 180, 6]

6.8. Effet des variations intra-classe

Les variations de couleur intra-classe peuvent être provoquées par des facteurs tels que l'effet de l'âge, le bronzage, la dépigmentation, les conditions d'éclairage, le maquillage, etc. La figure 6.9 illustre ces variations pour trois individus. Pour chaque individu, deux images différentes sont utilisées pour extraire les 30 couleurs dominantes ainsi que leur importance. Les barres de couleurs obtenues

montrent la différence entre les deux images bien qu'elles proviennent d'un même individu.

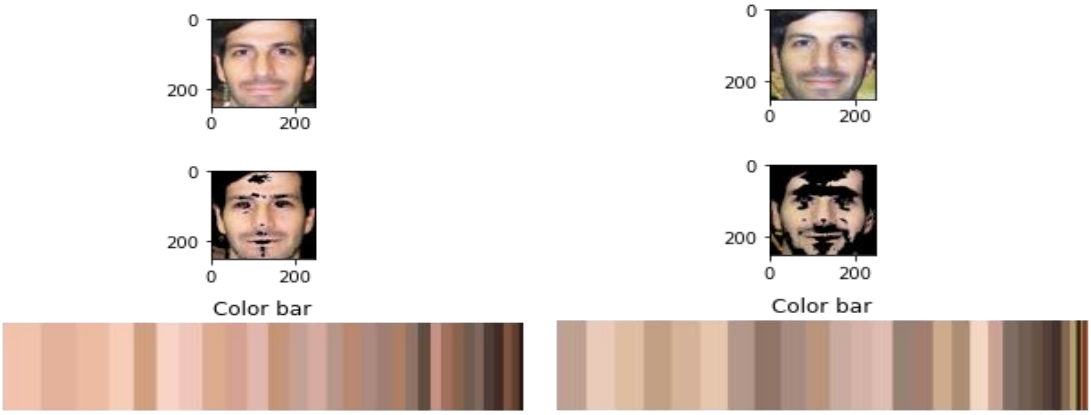

Figure 6.9.1 : individu 1

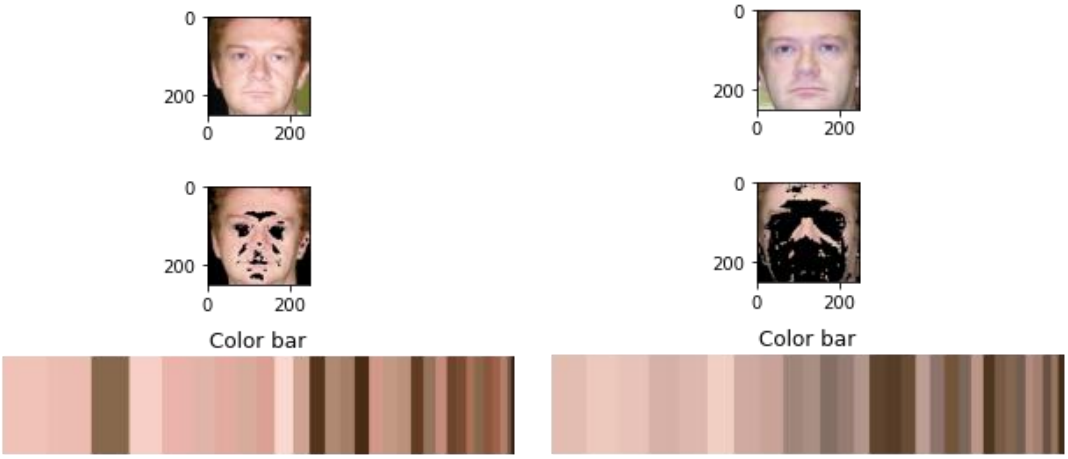

Figure 6.9.2 : Individu 2

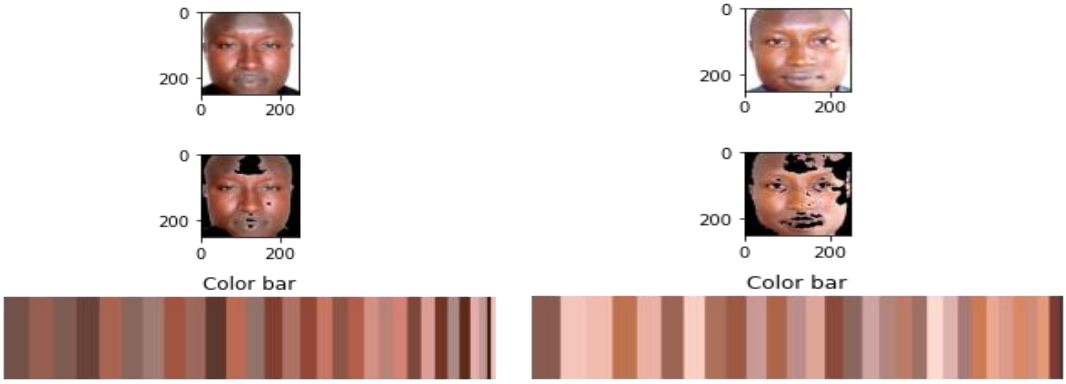

Figure 6.9.3 : Individu 3

Figure 6.9 : Variations intra-classe de différents individus

6.9. Résultats expérimentaux

Pour évaluer les performances du système de classification de la couleur de peau, nous avons opté pour les taux d'erreurs caractéristiques FAR et FRR annoncés à la section 6.7. Pour les tests du système nous avons utilisé cinq (5) images par individu. Nous mesurons deux paramètres liés aux taux d'erreurs des systèmes d'authentification biométrique [65]:

-Taux de fausses acceptations (false acceptance rate, FAR) : Proportion des transactions des imposteurs acceptées par erreur. Pour ce taux nous utilisons 5 images de personnes différentes que nous comparons avec le modèle d'un individu. L'opération est répétée pour les 19 individus présents dans la base de données. Nous avons obtenu un taux de 29.47% pour ce paramètre.

-Taux de faux rejets (false rejection rate, FRR) : Proportion des transactions des utilisateurs légitimes rejetées par erreur. Ces transactions sont rejetées, par l'algorithme de correspondance, en raison de non-correspondance à tort ainsi que ceux rejetées en raison d'un échec à l'acquisition. Pour ce taux nous utilisons 5 images de la même personne et nous les comparons au modèle de cette personne. L'opération est répétée également pour les 19 individus présents dans la base de données. Nous avons obtenu un taux de 70.53% pour ce deuxième paramètre.

Tableau 6.4 : Taux d'erreurs du système d'authentification

Taux	Valeur
Taux de fausses acceptations, FAR	29.47%
Taux de faux rejets, FRR	70.53%

6.10. Discussion des résultats

La détection de la peau a été faite par segmentation. La méthode de seuillage globale a été utilisée. Elle a l'avantage d'être rapide à mettre en œuvre. Il existe plusieurs méthodes de segmentation sur la détection de la peau humaine qui donnent des résultats intéressants. Entre autres, nous pouvons citer la méthode de fusion proposée par Tan et al. [112] pour la détection de la peau. Cette méthode pourrait permettre d'avoir de meilleurs résultats. Par ailleurs, dans notre système nous utilisons les composantes de l'espace colorimétrique RGB pour authentifier un individu. Nous pourrions éventuellement avoir de meilleurs résultats en utilisant une combinaison des espaces colorimétriques RGB, YCbCr et HSI tel que proposé par Singh et al. [92]. Sur un autre plan, il convient de souligner l'effet non négligeable des facteurs provoquant les variations intra-classe. Ces variations sont inhérentes à l'être humain ; cela justifie la mise en place d'un système de pondération tel que proposé par Jain et al. [82] en 2004.

Conclusion

D'après les résultats obtenus, on remarque que le taux de faux rejets est très élevé (70.53%). Cela signifie que le système rejette beaucoup des personnes qui normalement devraient être reconnues et authentifiées. Par contre le taux de fausses acceptations est relativement bas (29.47%), comparativement au taux de faux rejets. Ces résultats sont caractéristiques des systèmes d'authentification de haute sécurité. Cependant, il convient de garder à l'esprit que le système n'est pas conçu pour être utilisé dans une approche d'authentification monomodale. Il est conçu pour être associé à des modalités de biométrie pure (visage, empreinte digitale, etc.) dans une approche de multibiométrie. L'une des perspectives de notre étude est l'implémentation d'un système d'authentification par la biométrie douce basée sur la couleur de peau en utilisant une combinaison des espaces colorimétriques RGB, YCbCr et HSI. Ainsi il serait possible de faire une étude comparative entre l'utilisation d'un seul espace colorimétrique d'une part et la

combinaison de plusieurs espaces colorimétriques d'autre part. Un autre champ d'investigation sera la conception d'un système capable de contourner les limites liées à l'influence des facteurs externes comme les conditions d'éclairage afin d'améliorer la détection de la peau. Le dernier chapitre à venir (chapitre 7) sera consacré à la mise en œuvre de l'architecture de fusion séquentielle adaptée.

Chapitre 7 : Mise en œuvre de la fusion séquentielle adaptée

Sommaire

Introduction

La multimodalité présente de nombreux avantages dont entre autres la robustesse aux impostures, une solution au problème d'universalité mais aussi et surtout elle permet une amélioration importante des performances de vérification. Par ailleurs, les résultats issus de l'intégration des métadonnées dans l'authetification biométrique ont montré un autre niveau d'amélioration des performances. Au chapitre 3 nous avons mis au point une nouvelle architecture de fusion de scores à savoir l'architecture de fusion séquentielle adaptée. Le présent chapitre vise à évaluer les performances de cette nouvelle architecture. A cet effet, le système d'authentification par la couleur de peau sera combiné dans une première étape avec l'authentification du visage. Ensuite, l'empreinte digitale sans contact sera utilisée comme deuxième modalité de biométrie pure avant de procéder à la combinaison des trois modalités en utilisant l'algorithme de fusion séquentielle adaptée. Il sera ainsi possible d'apprécier le niveau de contribution de la couleur de peau dans l'amélioration de la performance du système d'authentification par le visage et l'empreinte digitale sans contact.

7.1. Environnement d'implémentation

7.1.1. Matériel de travail

L'implémentation des différentes méthodes présentées dans les chapitres précédents a été faite en utilisant un ordinateur dont les spécifications techniques sont décrites ci-dessous :

- Type : Portatif ;
- Processeur : Intel Core i5 ;
- Fréquence processeur : 2,3 GHz ;
- Mémoire Ram : 4 Go ;
- Carte graphique : Intel HD Graphics 520 ;
- Système d'exploitation : Linux ;
- Distribution système : Ubuntu 18.10.

7.1.2. Bases de données utilisées

En référence aux deux modalités de biométrie pure à utiliser pour l'authentification, deux bases de données seront utilisées. La première base de données est constituée d'images de visages de différentes personnes. Elle est exploitée pour la reconnaissance faciale et la classification de la couleur de peau du visage. La deuxième base de données est constituée d'images d'empreintes digitales. Cette seconde base sera utilisée pour la reconnaissance par empreinte digitale. La base de données d'images du visage ainsi que les protocoles d'implémentation ont été décrites dans le chapitre 6. Quant à la base de données d'empreintes digitales, elle a été conçue par Djara et al. [113]. Il s'agit d'images d'empreintes digitales acquissent sans contact. La base de données comporte 420 images d'empreintes digitales comprenant 28 séries de doigts différents avec 15 acquisitions pour chaque doigt. La figure 7.1 présente un échantillon d'images de cette base de données. Pour les besoins de tests du système multimodal (visage et empreinte digitale), nous avons dû composer une base de données virtuelle de 19 personnes (le maximum au niveau du visage) ayant chacune 15 images de chaque

modalité (le maximum au niveau de l'empreinte digitale). Pour la formation du modèle de reconnaissance d'empreintes digitales basé sur l'apprentissage par transfert, nous avons utilisé 11 images. Les 4 images restantes ont été utilisées pour le test du modèle de reconnaissance.

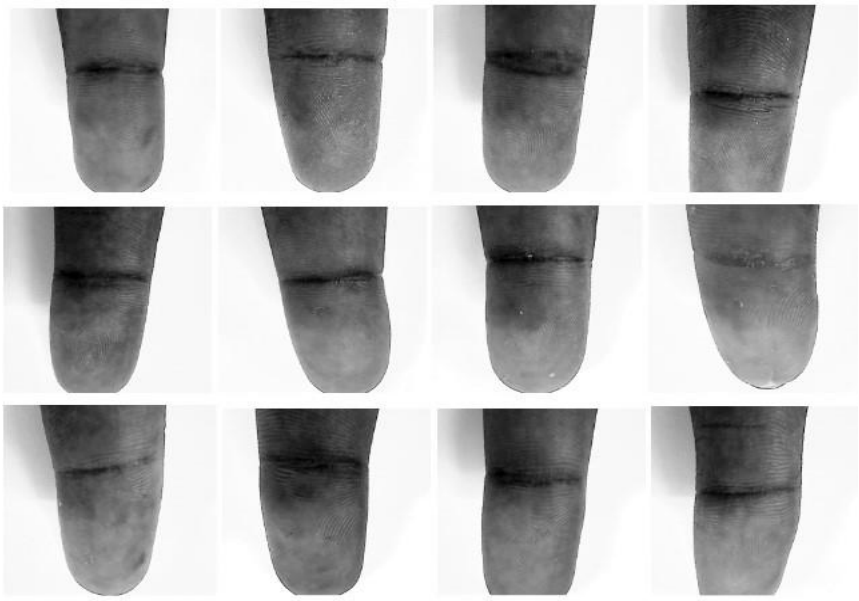

Figure 7.1 : Echantillon d'images de la base de données d'empreintes digitales [113]

7.1.3. Le cadre logiciel d'implémentation

Ozgur et al. [114] ont mené une étude comparative entre Matlab, Python et R. L'étude a abouti à la préférence de l'environnement Python pour l'implémentation, comparativement à Matlab et R. Python est un langage de programmation dont l'implémentation la plus courante est celle en C (aussi connu sous le nom de CPython). C'est un langage de programmation interprété, multi-paradigme et multiplateformes. Il favorise la programmation fonctionnelle et orientée objet. Il est doté d'un typage dynamique fort, d'une gestion automatique de la mémoire par ramasse-miettes et d'un système de gestion d'exceptions ; il est ainsi similaire à Perl, Ruby, Scheme, Smalltalk et Tcl. Le langage Python est placé sous une licence libre proche de la licence BSD et fonctionne sur la plupart des plates-formes informatiques, des smartphones aux ordinateurs centraux, de

Windows à Unix avec notamment GNU/Linux en passant par macOS, ou encore Android, iOS, et peut aussi être traduit en Java ou .NET. Il est conçu pour optimiser la productivité des programmeurs en offrant des outils de haut niveau et une syntaxe simple à utiliser. Non seulement Python est un langage de programmation, mais il consiste en une grande bibliothèque standard. Cette bibliothèque est structurée pour se concentrer sur la programmation générale et contient des modules pour les OS spécifiques, les threads, la mise en réseau et les bases de données.

Par ailleurs, une étude a été menée par Piatetsky [115] sur les outils utilisés dans les sciences de données entre 2016 et 2019. Cette étude révèle que depuis 2017, Python est l'outil le plus utilisé pour le traitement de données au regard des différentes possibilités qu'il offre. Pour ce faire, nous avons porté notre choix sur ce langage.

7.2. Authentification par le visage

7.2.1. Les méthodes de reconnaissance faciale

Les généralités sur la reconnaissance faciale ayant été abordées dans la section 2.1.1.1, nous allons aborder dans ce chapitre les méthodes de reconnaissance. Les algorithmes de reconnaissance de visage sont principalement classés en deux catégories : les approches locales et les approches globales. A ces deux approches s’ajoutent celles hybrides.

7.2.1.1. Les méthodes locales

Elles se fondent sur une analyse bas niveau de l'image numérique à travers des caractéristiques telles que les yeux, le nez, la bouche, la couleur de peau. Les statistiques locales et les positions de ces caractéristiques géométriques interviennent souvent dans l'analyse de l'image. Le point fort de ce type de méthode est leur faculté à apporter une vision/information multi-niveaux sur le contenu de l'image. Ces méthodes sont robustes aux occlusions partielles mais vulnérables au bruit, aux changements d'expressions et de poses de tête, ce qui

entraîne des imprécisions au niveau des caractéristiques [116]. Parmi ces approches locales, nous avons par exemple les opérateurs LBP [117], l'apprentissage adaptatif d'Adaboost [118], les descripteurs à base de SIFT [119].

7.2.1.2. Les méthodes globales

Au niveau de ces méthodes, on utilise le visage dans sa globalité. Une phase préalable d'analyse par des méthodes statistiques est généralement réalisée afin d'obtenir une représentation et/ou une classification des visages. Ensuite, vient une phase de classification en utilisant notamment, des méthodes basées sur l'Intelligence Artificielle telles que les Modèles de Markov Cachés (HMM), les réseaux de neurones (NN), les Machines à Vecteurs de Support (SVM). Les principaux avantages des méthodes globales statistiques sont leur représentation compacte, leur mise en œuvre rapide et la complexité moyenne des calculs de base nécessaires dans la phase de matching. Par contre, elles sont sensibles aux variations d'éclairage, de pose et d'expression faciale. En effet, la moindre variation des conditions de l'environnement ambiant entraîne des changements inéluctables dans les valeurs des pixels qui sont traités directement.

7.2.1.3. Les méthodes hybrides

Ces méthodes permettent d'associer les avantages des méthodes globales et locales en combinant la détection de caractéristiques géométriques (ou structurales) avec l'extraction de caractéristiques d'apparence locales. Elles augmentent la stabilité de la performance de reconnaissance lors de changements de pose, d'éclairage et d'expressions faciales. Parmi ces méthodes nous avons les techniques modulaires telles que la PCA modulaire, et la LDA modulaire ; les techniques d'appariement de graphes telles que le Gabor-EBGM (Gabor Elastic Bunch Graph Matching) et le HOG-EBGM.

7.2.2. L'authentification du visage sous OpenCV

Dans cette section nous allons expliciter la méthode utilisée pour la reconnaissance faciale. Il existe trois algorithmes de reconnaissance faciale inclus dans OpenCV : LDA ou Fisherfaces [120], PCA ou Eigenfaces [121] et LBPH [122]. L'algorithme basé sur le LBPH étant moins affecté par les conditions d'éclairage que les algorithmes basés sur Eigenfaces et Fisherfaces ; nous avons alors opté pour la méthode LBPH. L'algorithme du LBPH utilise l'opérateur LPB et crée en suite des histogrammes représentants les caractéristiques importantes extraites du visage. Cet algorithme sera mieux décrit dans la suite. Comme annoncé dans la section 2.1.1.1, la reconnaissance faciale comporte trois principales étapes : Détection de visage, extraction des caractéristiques et reconnaissance de visage.

7.2.2.1. Détection du visage

La détection de visage a été ébauchée dans la section 5.4 ; nous apportons ici des détails concernat les classificateurs utilisés. Il est à noter que OpenCV intègre deux types de classificateurs, LBP (Local Binary Patterns) et Haar Cascades que nous utilisons dans le cadre de ce travail. Le principe de fonctionnement de ce classificateur part d'une ondelette de Haar qui est une fonction mathématique produisant des ondes carrées avec un début et une fin et qui sert à créer des motifs en forme de boîte pour reconnaître les signaux avec des transformations soudaines [122]. Un exemple est illustré à la figure 7.2.

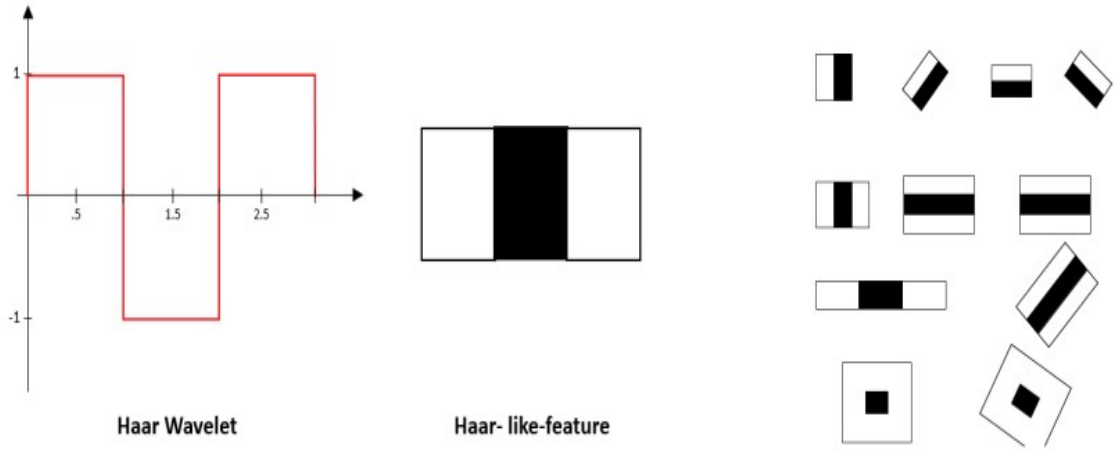

Figure 7.2 : Une ondelette de Haar et les caractéristiques résultantes de type Haar [122]

En combinant plusieurs ondelettes, on peut créer une cascade qui peut identifier des bords, des lignes et des cercles avec des intensités de couleurs différentes. Pour analyser une image à l'aide des cascades Haar, une échelle plus petite que l'image cible est sélectionnée. Elle est ensuite placée sur l'image, et la moyenne des valeurs des pixels de chaque section est prise. Si la différence entre deux valeurs dépasse un seuil donné, elle est considérée comme une correspondance. La détection des visages s'effectue en associant une combinaison de différentes caractéristiques de type Haar. Par exemple, front, sourcils et le contraste des yeux ainsi que le nez avec les yeux. Haar Cascade utilise l'algorithme d'apprentissage Adaboost, qui sélectionne un petit nombre de caractéristiques importantes dans un vaste ensemble pour obtenir un résultat efficace de classification [118]. La figure 7.3 montre les caractéristiques que recherche l'algorithme de détection de visage dans une image ou une vidéo.

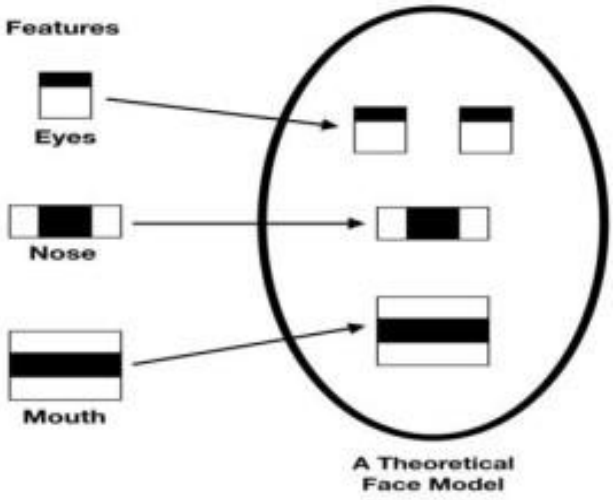

Figure 7.3 : Caractéristiques extraites par Adaboost

Dans la section 6.3, figure 6.2, nous avons donné une illustration de quelques visages détectés par l'algorithme LBPH.

7.2.2.2. Extraction des caractéristiques

Pour extraire les informations pertinentes sur une image de visage, nous allons utiliser l'opérateur LBP. L'opérateur LBP est l'un des descripteurs de texture les plus performants et a été largement utilisé dans diverses applications. L'idée d'utiliser le LBP pour la description des visages est motivée par le fait que les visages peuvent être vus comme une composition de micro-motifs bien décrits par un tel opérateur [123]. Il s'est avéré extrêmement discriminant et ses principaux avantages, à savoir son invariance aux changements monotoniques de niveaux de gris et son efficacité de calcul, le rendent convenable aux tâches d'analyse d'image exigeantes. Les étapes du déroulement de l'algorithme sont les suivantes :

1. On crée une fenêtre 3X3 et on la déplace sur toute l'image. À chaque déplacement (chaque partie locale de l'image), on compare le pixel au centre avec ses pixels environnants. On indique les voisins avec une valeur d'intensité inférieure ou égale au pixel central par 1 et le reste par 0.
2. Une fois que nous avons lu ces valeurs 0/1 dans la fenêtre 3X3 dans le sens des aiguilles d'une montre, on obtient un modèle binaire tel que 11100011,

qui est local à une zone particulière de l'image. On obtient au terme une liste de modèles binaires locaux.

3. Ensuite, après avoir obtenu une liste de modèles binaires locaux, on convertit chacun d'eux en un nombre décimal à l'aide d'une conversion binaire en décimal.
4. Enfin on crée l'histogramme de toutes ces valeurs décimales.

Une description plus formelle de l'opérateur LBP peut être donnée comme suit [124] :

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p S(g_c - g_p) \tag{6.1}$$

Avec (x_c, y_c) les coordonnées du pixel central c de valeur en niveau de gris g_c ; et g_p étant la valeur en niveau de gris du pixel voisin p. S est la fonction de signe définie comme :

$$S(x) = \begin{cases} 1, & \text{si } x \geq 0 \\ 0, & \text{sinon} \end{cases} \tag{6.2}$$

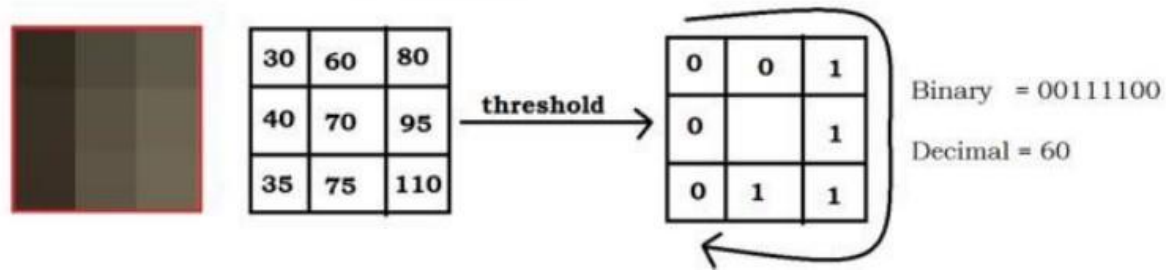

Figure 7.4 : Opérateur LBP

7.2.2.3. Authentification du visage

Il s'agit de la dernière étape du processus de reconnaissance faciale. Lorsqu'une nouvelle image est transmise au programme, l'histogramme de cette nouvelle image est créé et comparé avec les histogrammes préalablement enregistrés dans le modèle de référence. Le programme retournera enfin l'étiquette et le nom associé à la meilleure correspondance. La figure 7.5 en donne une illustration pour un visage provenant de la base de données utilisée.

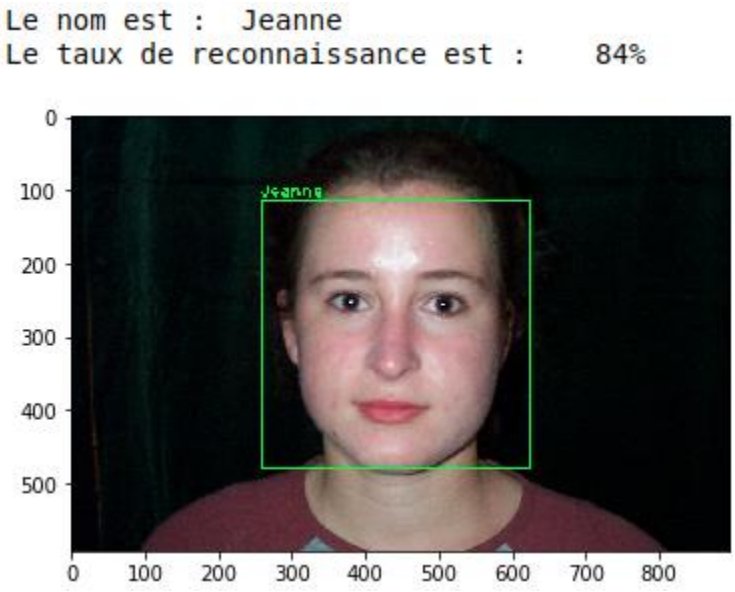

Figure 7.5 : Reconnaissance faciale d'un individu sous OpenCV

Dans le but de déterminer le taux d'égale erreur (EER), nous avons fait varier les taux de faux rejet (FRR) et de fausses acceptations (FAR) en fonction du seuil de décision. Cela nous a permis d'obtenir le graphique illustré à la figure 7.6

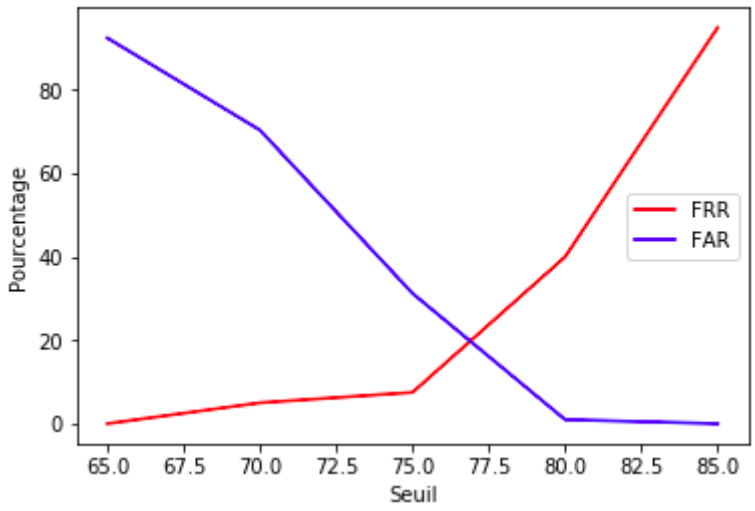

Figure 7.6 : Courbe EER de la reconnaissance faciale

7.3. Authentification par l'empreinte digitale sans contact

7.3.1. Extraction des caractéristiques en utilisant les réseaux de neurones convolutionnels

Les généralités sur l'empreinte digitale ont été présentées dans la section 2.1.1.2. La plupart des algorithmes proposés pour la reconnaissance d'empreintes digitales sont basés sur la correspondance des minuties. Mais ces dernières années, le développement de l'intelligence artificielle, précisément de l'apprentissage automatique (machine learning) permet de résoudre de nombreux problèmes dans divers domaines. Le deep learning est utilisé aujourd'hui pour résoudre divers problèmes tels que la classification, la segmentation, le sous-titrage d'images, l'analyse des émotions, la reconnaissance des visages et la détection d'objets, et a considérablement amélioré les performances par rapport aux approches traditionnelles [125]. Dans cette direction, les réseaux de neurones convolutionnels (CNN) ont connu un grand succès dans plusieurs domaines de la vision par ordinateur, y compris la reconnaissance des empreintes digitales. Pour la reconnaissance des empreintes digitales sans contact, nous avons opté pour une méthode d'apprentissage par transfert appliquée aux CNN. En effet, les connaissances acquises par un réseau de neurones pour une tâche peuvent être

transférées à un autre réseau de neurones pour résoudre un problème similaire : on parle alors de *tranfer learning*.

Dans l'apprentissage par transfert, un modèle formé à une tâche est réadapté à une autre tâche connexe, habituellement par une certaine adaptation vers la nouvelle tâche. Pour notre cas nous avons utilisé l'architecture des modèles MobileNets. MobileNets est une famille de modèles de vision par ordinateur pour TensorFlow, conçue pour maximiser efficacement la précision tout en tenant compte des ressources limitées pour une application sur appareil ou embarquée. Les MobileNets sont des modèles de petite taille, à faible latence et à faible consommation d'énergie, paramétrés pour répondre aux contraintes de ressources de divers cas d'utilisation. Ils peuvent être utilisés pour la classification, la détection, l'intégration et la segmentation [126]. Pour tenir compte de notre contexte expérimental, nous avons adapté l'architecture MobileNets à nos besoins pour la classification des images d'empreinte digitale sans contact. La figure 7.7 montre la courbe d'apprentissage du modèle adapté.

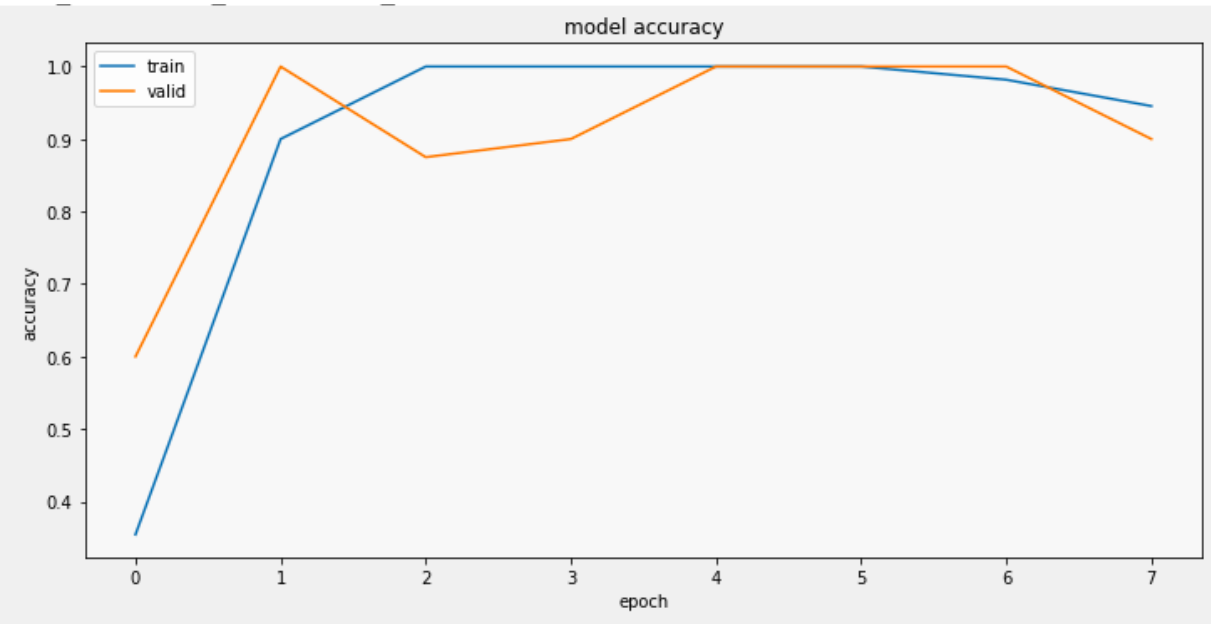

Figure 7.7 : Courbe d'apprentissage du modèle de reconnaissance des empreintes digitales

7.3.2. La comparaison (matching)

Cette étape permet de vérifier si l'empreinte d'un individu lui correspond. Une fois le système de reconnaissance entraîné avec les images d'empreinte, on lui présente d'autres images inconnues pour voir si le système arrive à les faire correspondre au bon individu. Notre système de reconnaissance fonctionne et permet de reconnaître avec une précision absolue l'identité d'un individu quelconque. Ce constat est etayé dans la rubrique 7.4.2.

7.4. Les résultats de la fusion séquentielle adaptée des trois modalités

7.4.1. Les résultats de l'authentification faciale combinée avec la couleur de peau

Pour la reconnaissance faciale nous avons utilisé la méthode basée sur l'opérateur LBP inclus dans OpenCV. Ainsi nous avons formé un modèle de reconnaissance pour les images de notre base de données. Ensuite nous avons fusionné ce système avec celui de la couleur de peau. Cette fusion constitue la première étape de l'algorithme de fusion séquentielle adaptée.

Les figures 7.8, 7.9 et 7.10 présentent la courbe ROC (receiver operating characteristic) respectivement (i) du système de reconnaissance faciale avec l'algorithme LBPH, (ii) du système de reconnaissance faciale avec la méthode séquentielle et (iii) du système de reconnaissance faciale combiné avec la couleur de peau. Cette courbe (ROC) constitue l'une des méthodes les plus couramment utilisées afin d'évaluer la performance globale d'un système d'authentification biométrique. L'avantage de cette méthode est qu'on obtient une représentation compacte de la performance d'un système biométrique pour ses différents paramétrages au travers d'une seule courbe, qui permet de comparer objectivement différents systèmes biométriques [65]. Il s'agit en fait de la représentation de la relation entre le taux de fausses acceptations (FAR) et le taux de faux rejets (FRR) pour les différentes valeurs du seuil de décision, respectivement en abscisses et en ordonnées. L'AUC (Area Under the Curve) de

la courbe ROC est une mesure qui permet de voir facilement l'efficacité du système. Cette valeur mesure l'intégralité de l'aire à deux dimensions située sous l'ensemble de la courbe ROC (par calculs d'intégrales) de (0,0) à (1,1). Les valeurs d'AUC sont comprises dans une plage de 0 à 1. Un modèle dont 100% des prédictions sont erronées a un AUC de 0,0. Si toutes ses prédictions sont correctes, son AUC est de 1,0. Ainsi, un système d'authentification biométrique est efficace lorsque son AUC tend vers 1.

Les résultats expérimentaux nous ont donné les valeurs respectives de (i) 0,826 pour le système de reconnaissance faciale avec l'algorithme LBPH (voir figure 7.8), (ii) 0,916 pour le système de reconnaissance faciale avec la méthode séquentielle (voir figure 7.9) et (iii) 0,944 pour le système de reconnaissance faciale combiné avec la couleur de peau (voir figure 7.10). Ces résultats montrent une amélioration progressive des performances de reconnaissance d'une méthode à une autre. De façon particulière, l'augmentation de l'AUC entre la figure 7.9 et la figure 7.10 montre que la couleur de peau a eu un impact positif sur la reconnaissance faciale et donc accroît les performances de la méthode séquentielle de 2,8%.

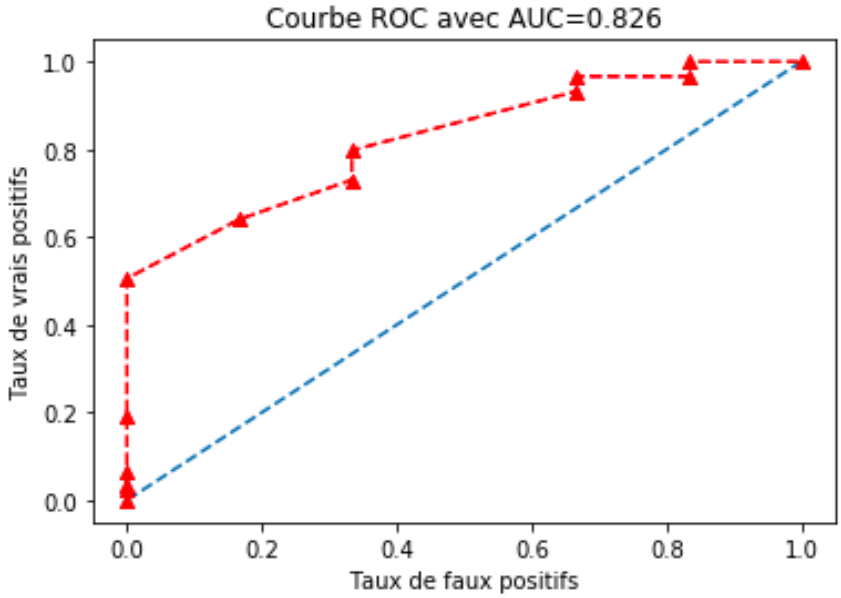

Figure 7.8 : Courbe ROC du visage (méthode LBPH)

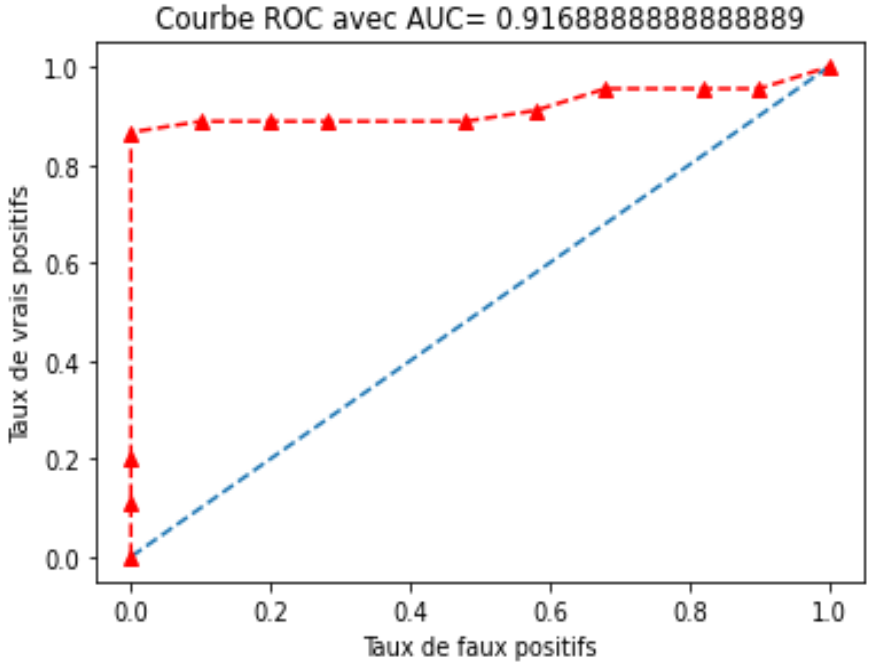

Figure 7.9 : Courbe ROC du visage combiné à la couleur de peau (fusion séquentielle)

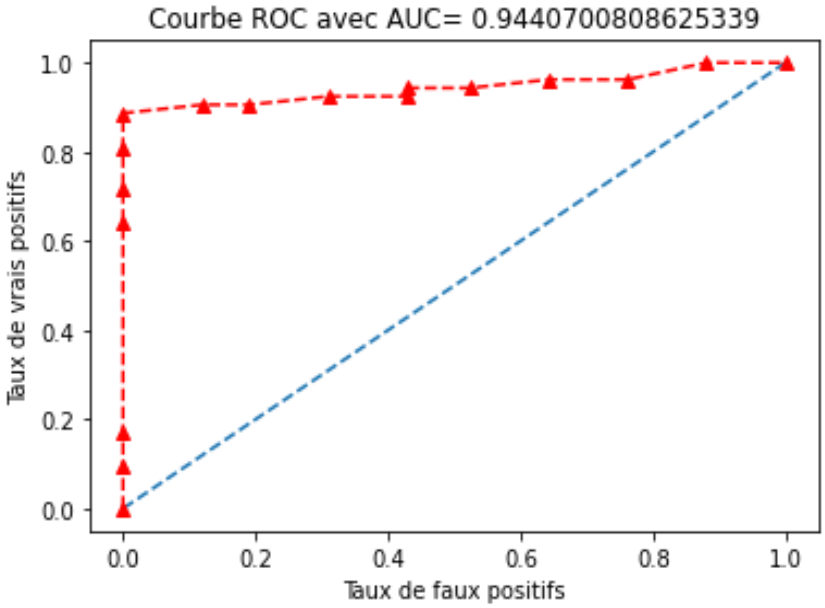

Figure 7.10 : Courbe ROC du visage combiné à la couleur de peau (fusion séquentielle adaptée)

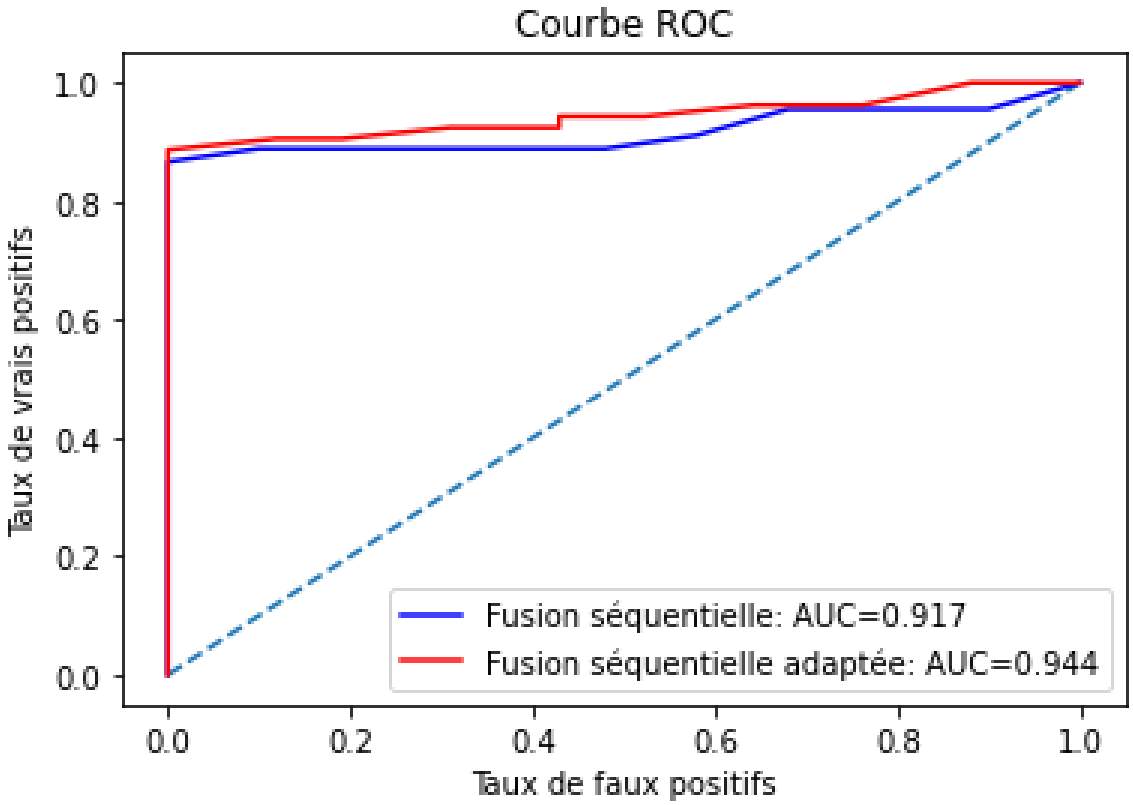

Figure 7.11 : Courbes ROC du visage combiné à la couleur de peau (fusion séquentielle et fusion séquentielle adaptée)

7.4.2. Les résultats de l'authentification par l'empreinte digitale sans contact

Au niveau de la reconnaissance par l'empreite digitale, une méthode basée sur l'intelligence artificielle a été utilisée. Il s'agit de la méthode d'apprentissage par pransfert. Les résultats de l'authentification des utilisateurs de la base de données sont consignés dans la matrice de confusion présentée à la figure 7.12.

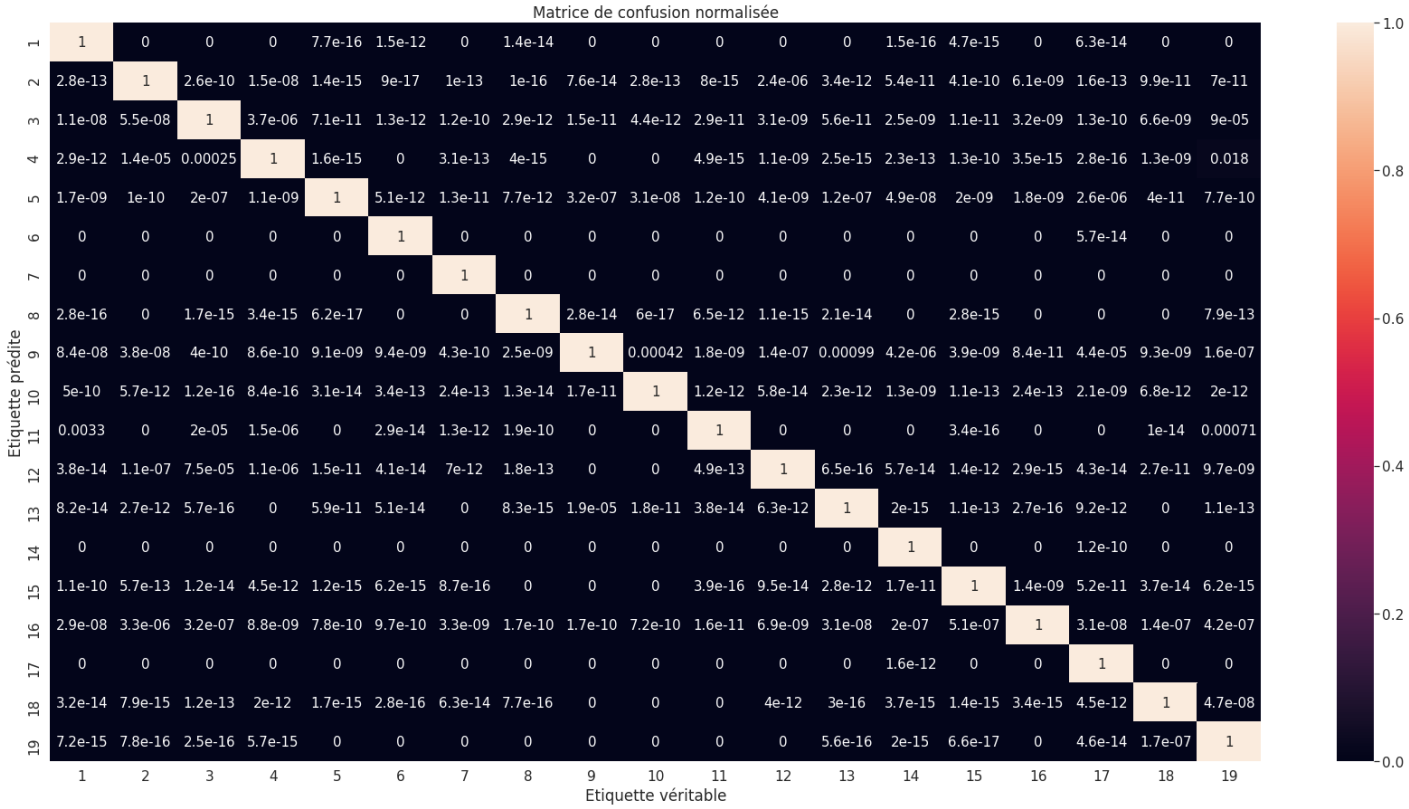

Figure 7.12 : Matrice de confusion de la reconnaissance par empreinte digitale

Ces résultats montrent que la méthode utilisée parvient à effectuer une reconnaissance sans faille des individus présents dans la base de données, sans aucune confusion. Cette situation est due à la faible taille de notre base de données. Par ailleurs, le fait que la base de données ait été acquise en condition supervisée contribue aussi à ce niveau de perfection des résultats.

7.4.3. Les résultats combinés des trois modalités

La différence majeure entre notre architecture avec celle de fusion séquentielle se situe au niveau de l'introduction de métadonnées dans le processus de fusion. La

première étape de notre processus inclus la combinaison d'une donnée de biométrie pure à savoir le visage puis une donnée de biométrie douce qu'est la couleur de peau. L'augmentation de l'AUC au niveau de la figure 6.10 (combinaison du visage et de la couleur de peau) par rapport à celui de la figure 6.9 (visage seul) montre que l'introduction d'une métadonnée dans le processus de reconnaissance permet d'améliorer les performances du système. Cette amélioration peut être constatée à travers l'augmentation de la proportion d'individus reconnus dès la première étape. L'augmentation de la proportion des individus reconnus dès la première étape permet de gagner en temps et en ressources matérielles. Le tableau 6.1 présente une étude comparée du temps d'exécution des deux algorithmes (fusion séquentielle et fusion séquentielle adaptée). Deux paramètres ont été pris en compte à savoir le temps moyen d'exécution et son écart type.

Tableau 7.1 : Comparaison du temps de traitement des algorithmes "Fusion séquentielle" et "Fusion séquentielle adaptée"

Fusion séquentielle : Etape1 (Visage)	Fusion séquentielle : Etape1+Etape2 (Visage et empreinte)	Fusion séquentielle adaptée : Etape1 (Visage et couleur de peau)	Fusion séquentielle adaptée : Etape1+Etape2 (visage et couleur de peau plus empreinte)
Temps moyen : 1,25s Ecart-type : 0,05	Temps moyen : 13s Ecart-type : 1,29	Temps moyen : 1,60s Ecart-type : 0,25	Temps moyen : 14s Ecart-type : 1,63

Les écart-types de faibles valeurs présents dans le tableau 6.1 témoignent de la validité des temps moyens que nous utilisons pour nos analyses. Le temps moyen t de chaque méthode de fusion dépend de la proportion d'utilisateurs p étant reconnus dès la première étape. Ce temps est donné par les formules suivantes, par méthode :

1. Méthode séquentielle

$$t = 1{,}25p + 13(1 - p) \quad (6.3)$$

2. Méthode séquentielle adaptée

$$t = 1{,}6p + 14(1 - p) \quad (6.4)$$

Les équations 6.3 et 6.4 nous montrent que le temps t est linéairement dépendant et inversement proportionnel à la proportion d'utilisateurs étant reconnus à la première étape de ces deux méthodes de fusion de scores. Les résultats expérimentaux nous ont donné une proportion p de 0,41 (39/95) pour la fusion séquentielle et 0,49 (47/95) pour la fusion séquentielle adaptée. Sur cette base, nous avons calculé le temps moyen de fusion de scores pour chacune des deux méthodes. Les résultats se présentent ainsi qu'il suit :

1. Méthode séquentielle : $p = 41\%$

Temps moyen fusion séquentielle (en secondes) : $t = 1{,}25p + 13(1 - p)$ soit $t = 13 - 11{,}75p = 13 - 4{,}8175 = 8{,}1825$

2. Méthode séquentielle adaptée : $p = 49\%$

Temps moyen fusion séquentielle adaptée (en secondes) : $t = 1{,}6p + 14(1 - p)$ soit $t = 14 - 12{,}4p = 14 - 6{,}076 = 7{,}924$

L'algorithme de fusion séquentielle a une complexité de type linéaire ($O(n)$) tandis que notre algorithme de fusion séquentielle adaptée a une complexité de type quadratique ($O(n^2)$), donc bien plus complexe. Mais après l'implémentation, il ressort que l'algorithme de fusion séquentielle adaptée a une vitesse d'exécution plus élevée. Cela est dû à l'apport de la couleur de peau en termes de performance d'authentification.

Conclusion

Dans ce chapitre nous avons procédé à la mise en œuvre des différentes méthodes développées au niveau des chapitres 4 et 5 tout en exploitant les résultats obtenus

au chapitre 6. L'architecture séquentielle adaptée mise au point et qui se décline à travers un algorithme de 59 lignes d'instructions a fait l'objet d'une implémentation sous le framework Python avec le package OpenCV. Deux modalités de biométrie pure (visage et empreinte digitale sans contact) et une modalité de biométrie douce (couleur de peau du visage) ont été utilisées pour l'implémentation.

La méthode basée sur l'opérateur LBP a été utilisée pour la reconnaissance faciale. En ce qui concerne la couleur de peau, l'algorithme des k-moyenne a permis de ressortir les couleurs dominates sur le visage. Pour l'empreinte digitale, nous avons fait recours à une méthode d'intelligence artificielle (Machine Learning), notamment la méthode d'apprentissage par transfert ou trnasfer learning. La matrice de confusion obtenue montre que l'algorithme ne fait aucune confusion entre les utilisateurs enregistrés dans la base de données. Les résultats issus de la fusion globale ont montré une amélioration des performances avec une augmentation de la proportion des utilisateurs reconnus à la première étape. Cette proportion est de 49% pour notre architecture de fusion séquentielle adaptée contre 41% pour l'architecture de fusion séquentielle.

Conclusion générale et perspectives

Les travaux effectués dans le cadre de cette thèse ont trait aux systèmes multibiométriques et plus particulièrement la prise en compte des métadonnées avec pour finalité d'améliorer les performances de reconnaissance. Nous avons commencé par effectuer un état de l'art détaillé sur les systems biométriques développés à ce jour. Nous avons ainsi contribué à enrichir le vocabulaire biométrique avec de nouveaux concepts tels que la *"biométrie multi-origine"* et la *"fusion séquentielle adaptée"*. Cette biométrie multi-origine constitue la sixième catégorie des systèmes multibiométriques. Nous avons également proposé une nouvelle catégorisation des modalités biométriques en quatre classes à savoir : les modalités morphologiques, comportementales, biochimiques et bioélectriques. En ce qui concerne l'évaluation des performances des systèmes biométriques, nous avons proposé une nouvelle méthode d'évaluation qui peut être appliquée à n'importe quel type de modalité biométrique.

En substance, nous avons présenté un aperçu des différentes méthodes biométriques, des systèmes multibiométriques avant de traiter des schémas de fusion. Nous nous sommes intéressés ensuite à la fusion de scores dans les systèmes multibiométriques avant de proposer une nouvelle architecture de fusion séquentielle adaptée. Pour la mise en œuvre de ce nouveau cadre de fusion, nous avons développé une méthode originale pour l'authentification des personnes par la couleur de peau. Le nouveau cadre de fusion a fait l'objet d'une implémentation sous Python avec le package OpenCV. Pour ce faire, le visage et l'empreinte digitale ont été utilisés comme modalités de biométrie pure tandis que la couleur de peau (extraite du visage) a été utilisée comme modalité de biométrie douce (métadonnée). Les résultats obtenus montrent une amélioration des performances en termes de taux de reconnaissance et de temps d'exécution. Ces résultats ouvrent la voie à un certain nombre de perspectives.

Le système que nous avons mis en place est un système multibiométrique sans contact, utilisant un seul capteur (une caméra) mais avec deux captures. Non seulement qu'il est non intrusif, ce système sera efficace dans des situations d'épidémie par exemple. Par contre, l'une des limites des systèmes multibiométriques (précisément les systèmes multi-capteurs, multi-modaux et multi-origines) est la multiplication des capteurs à utiliser en raison du nombre de modalités surtout. Pour les systèmes multi-instances, multi-échantillons et multi-algorithmes, le même capteur est utilisé pour les différentes captures. Une autre limite de ces systèmes est le nombre de captures à faire pour disposer des données à traiter. Une solution à cette problématique (capteurs/captures) passe par le développement de systèmes multibiométriques capables non seulement d'utiliser un seul capteur, mais surtout de pouvoir extraire plusieurs données de différentes modalités. L'un des scénarios possibles serait par exemple la capture d'une image de visage dont l'individu soulève la main à la hauteur du visage. Une telle image comporterait aussi bien le visage que la main de l'individu. Un traitement de cette image peut permettre d'une part de détecter le visage et d'autre part de détecter la main. L'image de la main peut founir des données sur la paume de main ainsi que les empreintes digitales sans contact. Dans ce scénario, il est possible de disposer à partir d'un seul capteur et d'une seule capture, d'au moins trois modalités de biométrie pure et une modalité de biométrie douce. Les trois modalités de biométrie pure étant le visage, la paume de main et les empreintes digitales sans contact. La modalité de biométrie douce sera celle de la couleur de peau. Une telle approche contribuera à faciliter la généralisation des systèmes multibiométriques et particulièrement les systèmes adaptés à l'utilisateur.

Le système d'authentification par la couleur de peau que nous avons mis en place est considéré comme un système de biométrie douce à l'étape actuelle des recherches. Sur la base des résultats obtenus, nous envisageons de mener d'autres travaux visant à effectuer une comparaison entre la couleur de peau et d'autres

modalités de biométie pure. L'objectif étant de voir la possibilité de classer la couleur de peau comme modalité de biométrie pure. En outre, nous avons mis en lumière l'importance des variations intra-classes dans l'authentification par la couleur de peau. La couleur de peau sur une image est sujette à plusieurs impacts tels que les conditions d'éclairage. Un autre champ d'investigation sera donc la conception d'un système capable de contourner les limites liées à l'influence de ces facteurs externes afin d'améliorer la détection de la peau. Un accent particulier doit être mis sur les algorithmes de détection de la peau et d'extraction des couleurs dominantes. L'algorithme de fusion séquentielle adaptée que nous avons développé présente une complexité de type quadratique (polynomiale). Il serait intéressant d'améliorer cette complexité pour l'amener à une forme linéaire.

En perspectives, nous comptons étudier le lien entre la biométrie et l'émotion. En effet, il a été prouvé que les émotions (exprimées ou ressenties) peuvent être mesurées et servir de base à une prise de décision dans le cadre d'un interrogatoire de police par exemple. Ainsi, l'une des perpectives de notre travail est d'explorer l'impact de l'émotion sur la performance des systèmes biométriques. Les différentes valeurs mesurées de l'émotion peuvent être utilisées comme métadonnée à combiner avec d'autres données de biométrie pure. L'un des scénarios envisagés sera d'extraire l'émotion exprimée par les individus à travers le visage.

Références

[1] Anil K. Jain and Arun Ross. Learning user-specific parameters in a multibiometric system. Appeared in Proc. International Conference on Image Processing (ICIP), Rochester, New York, September, pp. 22-25, DOI: 10.1109/ICIP.2002.1037958, 2002.

[2] Harbi AlMahafzah and Ma'en Zaid AlRwashdeh. A survey of multibiometric systems. International Journal of Computer Application, volume 43 No 15 April, pp. 36-43, DOI : 10.5120/6182-8612, 2012.

[3] M. Nageshkumar, P. K. Mahesh and M. N. Shanmukha Swamy. An efficient secure multimodal biometric fusion using palmprint and face image. International Journal of Computer Science Issues, Vol. 2, pp. 49-53, arXiv:0909.2373v1, 2009.

[4] Anil Jain, Karthik Nandakumar and Arun Ross. Score normalization in multimodal biometric systems. Elsevier, Pattern Recognition 38, pp. 2270-2285, DOI:10.1016/j.patcog.2005.01.012, 2005.

[5] Norman Poh and Samy Bengio. Can Chimeric Persons Be Used in Multimodal Biometric Authentication Experiments? In: Renals S., Bengio S. (eds) Machine Learning for Multimodal Interaction, MLMI 2005, Lecture Notes in Computer Science, vol 3869, Springer, Berlin, Heidelberg, 2005.

[6] Tahirou Djara, Abdou-Aziz Sobabe and Antoine Vianou. Incorporating Metadata in Multibiometric Score-Level Fusion: an Optimized Architecture. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November 2019, pp. 5290–5305, DOI : 10.35940/ijitee.A4118.119119, 2019.

[7] Mohamad El-Abed and Christophe Charrier. Evaluation of Biometric Systems. New Trends and Developments in Biometrics, pp. 149 - 169, <10.5772/52084>, <hal-00990617>, 2012.

[8] Pierre Dagnelie. Les Mots "Biomètre", "Biométrie" Et "Biométrique" Au Dix-Neuvième Siècle. Faculté des Sciences agronomiques B-5030 Gembloux (Belgique), Biometric Bulletin 5 (3), pp. 3-4, 1988.

[9] Ramadan Gad, AYMAN EL-SAYED, Nawal El-Fishawy, and M. Zorkany. Multi-biometric systems: A state of the art survey and research directions. International Journal of Advanced Computer Science and Applications, Vol. 6, No. 6, pp. 128-138, DOI: 10.14569/IJACSA.2015.060618, 2015.

[10] Mohamad El-Abed, Romain Giot, Baptiste Hemery and Christophe Rosenberger. Evaluation of Biometric Systems: A Study of Users' Acceptance and Satisfaction. Inderscience International Journal of Biometrics (IJBM), pp.1-27. <10.1504/IJBM.2012.047644>, <hal-00984024>, 2012.

[11] Abdou-Aziz SOBABE, Tahirou DJARA and Antoine VIANOU. Biometric System Vulnerabilities: A Typology of Metadata. Advances in Science, Technology and Engineering Systems Journal, Special Issue on Multidisciplinary Sciences and Engineering, Vol. 5, No. 1, pp. 191-200, January 2020, https://dx.doi.org/10.25046/aj050125, 2020.

[12] N. Gopal and Dr. R. K. Selvakumar. Multimodal biometric identification system - An overview. International Journal of Engineering Trends and Technology (IJETT) – Volume 33 Number 7-March (2016), pp. 351-355, DOI: 10.14445/22315381/IJETT-V33P267, 2016.

[13] Anil K. Jain, Patrick Flynn and Arun A. Ross. Handbook of Biometrics. © 2008 Springer Science+Business Media, LLC, 2008.

[14] Davide Maltoni, Dario Maio, Anil K. Jain and Salil Prabhakar. Handbook of Fingerprint Recognition. © Springer-Verlag London Limited, 2009.

[15] Souhila Guerfi. Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D. Traitement du signal et de l'image, Ph.D thesis, Université d'Evry-Val d'Essonne, <tel-00623243>, 2008.

[16] Anil K. Jain and Ajay Kumar. Biometric recognition: An overview. Chapter 3, © Springer Science + Business Media B.V., pp 49-79, DOI : 10.1007/978-94-007-3892-8_3, 2012.

[17] Gandhimathi Amirthalingam and Radhamani. G. A Multimodal Approach for Face and Ear Biometric System. International Journal of Computer Science Issues (IJCSI), Vol. 10, Issue 5, No 2, September 2013, 2013.

[18] Muhtahir O. Oloyede and Gerhard P. Hancke. Unimodal and multimodal biometric sensing systems: A review. IEEE Access, Volume 4, pp. 7532-7555, Digital Object Identifier: 10.1109/ACCESS.2016.2614720, 2016.

[19] Tahirou Djara, Marc Kokou Assogba and Antoine Vianou. A contactless fingerprint verification method using a minutiae matching technique. International Journal of Computer Vision and Image Processing, Volume 6, Issue 1, January-June 2016, pp. 12-27, DOI: 10.4018/IJCVIP.2016010102, 2016.

[20] Florence Francis-Lothai and David B.L. Bong. A fingerprint matching algorithm using bit-plane extraction method with phase-only correlation. Int. J. Biometrics, Vol. 9, No. 1, pp. 44–66, DOI: 10.1504/IJBM.2017.084135, 2017.

[21] Ajay Kumar and Cyril Kwong. Towards contactless, low-cost and accurate 3D fingerprint identification. IEEE Transactions on pattern analysis and machine intelligence, Vol. 37, No. 3, pp. 681-696, DOI: 10.1109/TPAMI.2014.2339818, 2015.

[22] Tahirou Djara. Caractérisation et reconnaissance des empreintes digitales : application en biométrie sans contact. Ph.D thesis. Université d'Abomey-Calavi, Bénin, 2013.

[23] Michèle Gagnon. Tout savoir sur la biométrie. http://biometrics.over-blog.com/pages/La_geometrie_de_la_main-2019729.html, visité le 20 novembre 2018.

[24] Arun A. Ross, Karthik Nandakumar and Anil K. Jain. Handbook of Multibiometrics. Publisher by Springer, Boston, MA, © Springer Science+Business Media, LLC 2006, 2006.

[25] Gary Heiting Od. The Retina: Where Vision Begins. October (2017), https://www.allaboutvision.com/resources/retina.htm. Visité le 20 octobre 2018.

[26] C. A. Oyeleye, T. M. Fagbola, R. S. Babatunde, and A. A. Adigun. An exploratory study of odor biometrics modality for human recognition. International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 9, November 2012, pp. 1-10, ID: IJERTV1IS9205, 2012.

[27] P. Inbavalli and G. Nandhini. Body Odor as a Biometric Authentication. International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5 (5), pp. 6270-6274, 2014.

[28] Sherif N. Abbas, Mohammed Abo–Zahhad, Sabah M. Ahmed and Mohammed Farrag. Heart-ID: human identity recognition using heart sounds based on modifying mel- frequency cepstral features. IET Biom, 2016, Vol. 5 Iss. 4, © The Institution of Engineering and Technology, pp. 284-296, 2016.

[29] Hafs Toufik. Reconnaissance Biométrique Multimodale basée sur la fusion en score de deux modalités biométriques : l'empreinte digitale et la signature manuscrite cursive en ligne. Ph.D Thesis, Badji Mokhtar University-Annaba, 2016.

[30] Q. Zhang. Wavelets networks: The radial structure and an efficient initialization procedure. Technical Report of Linköping University, LITH-ISY-I-1423,1992.

[31] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.Y. Glorennec, H. Hjalmarson, and A. Juditsky. Nonlinear black-box modeling in system identification: Unified overview. Automatica, Vol.31, no.12, pp.1691-1724, 1995.

[32] A. Juditsky. Wavelet estimators: Adapting to unknown smoothness. Technical Report IRISA, June 1994.

[33] Q. Zhang. Using Wavelet Network in Nonparametric Estimation. Publication Interne 833, IRISA, June 1994.

[34] Q. H. Zhang. Using Wavelet Network in Nonparametric Estimation. IEEE Trans. Neural Networks, Vol.8, pp. 227-236, 1997.

[35] Samer Chantaf. Biométrie par signaux physiologiques. Ph.D Thesis, Université Paris-Est Créteil, France, 2011.

[36] J. Zhang, G. G. Walter, Y. Miao, and W. N. W. Lee. Wavelet Neural Networks for function learning. IEEE Trans. Signal process, Vol.43, No.6, pp.1485-1497, June 1995.

[37] S. Mallat. A theory for multiresolution signal decomposition: The wavelet transform. IEEE Trans. Pattern Anal. Machine Intell., Vol.11, pp. 674-693, July 1989.

[38] N. Draper and H. Smith. Applied regression analysis, Series in Probability and Mathematical Statistics, Wiley, 1981, Second edition.

[39] A. Nait-Ali and R. Fournier. Signal and Image Processing for Biometrics. John Wiley & Sons, ISBN 978-1-84821-385-2, 2012.

[40] K. Aloui. Biométrie du cerveau humain. Ph.D Thesis, Université Paris-Est Créteil, France, 2012.

[41] K. Aloui, Amine Nait-Ali and S. Nacer. A novel approach based Brain Biometrics: some preliminary results for Individual identification. IEEE Workshop on Computational Intelligence in Biometrics and Identity Management, Paris, France, April 2011.

[42] M. Manoj Kumar and N. B. Puhan. Off-line signature verification: upper and lower envelope shape analysis using chord moments. In Biometrics, IET, vol.3, no.4, pp. 347-354, 2014.

[43] Anil K. Jain, Arun Ross and Salil Prabhakar. An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology, vol. 14, no. 1, January 2004, pp. 4-20, DOI: 10.1109/TCSVT.2003.818349, 2004.

[44] N. Poh. Multi-system biometric authentication. Ph.D thesis, EPFL, Lausanne, 2006.

[45] Anis Chaari. Nouvelle approche d'identification dans les bases de données biométriques basée sur une classification non supervisée. Ph.D thesis. Université d'Evry-Val d'Essonne, France, 2009.

[46] Haryati Jaafar and Dzati Athiar Ramli. A Review of Multibiometric System with Fusion Strategies and Weighting Factor. International Journal of Computer Science Engineering (IJCSE), Vol. 2 No.04, July 2013, ISSN: 2319-7323, 2013.

[47] A. Ross. An introduction to multibiometrics. Appeared in Proc. of the 15th European Signal Processing Conference (EUSIPCO), (Poznan, Poland), September 2007, DOI: 10.1007/978-0-387-71041-9_14, 2007.

[48] C. Sanderson and Kuldip K. Paliwal. Identity Verification Using Speech and Face Information. Published in Digital Signal Processing, Vol. 14, No. 5, pp. 449-480, 2004.

[49] Tin Kam Ho, Jonathan J. Hull and Sargur N. Srihar. Decision combination in multiple classifier systems. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(1), January 1994, pp. 66–75, DOI: 10.1109/34.273716, 1994.

[50] Kunal Kumar and Mohammed Farik. A Review Of Multimodal Biometric Authentication Systems. International Journal Of Scientific & Technology Research, Volume 5, Issue 12, December 2016, 2016.

[51] Y. Faridah, Haidawati Nasir, A. K. Kushsairy and Sairul I. Safie. Multimodal biometric algorithm: A survey. Biotechnology 15 (5), pp. 119.124, DOI: 10.3923/biotech.2016.119.124, 2016.

[52] Lorene Allano. La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles. Ph.D thesis,

Institut National des Télécommunications dans le cadre de l'école doctorale SITEVRY en co-accréditation avec l'Université d'Evry-Val d'Essonne, 2009.

[53] Lorene Allano, Sonia Garcia-Salicetti and Bernadette Dorizzi. A low cost incremental biometric fusion strategy for a handheld device. S+SSPR 2008 : Joint IAPR International Workshops on Structural, Syntactic and Statistical Pattern Recognition, Dec 2008, Orlando, United States. pp. 842 -851, 10.1007/978-3-540-89689-0_88., hal-01375832, 2008.

[54] Vijay M. Mane and Dattatray V. Jadhav. Review of Multimodal Biometrics: Applications, challenges and Research Areas. International Journal of Biometrics and Bioinformatics (IJBB), Volume 3, Issue 5, pp. 90-95, 2009.

[55] Christina-Angeliki Toli and Bart Preneel. A Survey on Multimodal Biometrics and the Protection of their Templates. IFIP Advances in Information and Communication Technology, September 2014, DOI : 10.1007/978-3-319-18621-4_12, 2014.

[56] Ahmed Banafa. What is affective computing? https://www.bbvaopenmind.com/en/what-is-affective-computing/, Visited on Saturday 27 october (2018).

[57] Erik Cambria. Affective computing and sentiment analysis. IEEE Computer Society, pp. 102-107, DOI: 10.1109/MIS.2016.31, 2016.

[58] Mohammad Soleymani, David Garcia, Brendan Jou, Björn Schuller, Shih-Fu Chang and Maja Pantic. A survey of multimodal sentiment analysis. Image and Vision Computing, DOI : http://dx.doi.org/10.1016/ j.imavis.2017.08.003, 2017.

[59] K. Ahmad. Affective Computing and Sentiment Analysis: Metaphor, Ontology, Affect and Terminology, 174 pages, 2011.

[60] Imen Tayari, Nhan Le Thanh, Chokri Ben Amar. Modélisation des états émotionnels par un espace vectoriel multidimensionnel. Rapport de recherche, ISRN I3S/RR–2009-19–FR, Décembre 2009, 2009.

[61] Tahirou Djara, Abdoul Matine Ousmane, Antoine Vianou. Mood and personality influence on emotion. © ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2019, Springer Nature Switzerland AG 2019, AFRICATEK 2018, LNICST 260, pp. 166–174, https://doi.org/10.1007/978-3-030-05198-3_15, 2019.

[62] Dr. Clive Chandler and Rachel Cornes. Biometric Measurement of Human Emotions. 5th IEEE International Conference on Advanced Computing & Communication Technologies [ICACCT-2011] ISBN 81-87885-03-3, 2011.

[63] Precise Biometrics. Understanding Biometric Performance Evaluation. © Precise Biometrics AB-SPA 133 1000 4160/wpbpe RA, 2014.

[64] P. Jonathon Phillips, Alvin Martin, C.L. Wilson and Mark Przybocki. An Introduction to Evaluating Biometric Systems. National Institute of Standards and Technology, © 2000 IEEE, 2000.

[65] Mohamad El-Abed. Évaluation de système biométrique. Cryptographie et sécurité [cs.CR]. Ph.D thesis, Université de Caen, France, 2011.

[66] ISO/IEC 19795-1. Information technology – biometric performance testing and reporting – part 1 : Principles and framework, 2006.

[67] James P. Egan. Signal detection theory and ROC-analysis. By Academic Press, New York, 1975.

[68] F. Cherifi, B. Hemery, R. Giot, M. Pasquet, and C. Rosenberger. Performance evaluation of behavioral biometric systems. In Behavioral Biometrics for Human Identification : Intelligent Applications, pp. 57–74, 2009.

[69] J. Bhatnagar and A. Kumar. On estimating performance indices for biometric identification. Pattern Recognition, 42, pp. 1803–1815, 2009.

[70] D. Faraggi and B. Reiser. Estimation of the area under the ROC curve. Statistics in medicine, 21, pp. 3093–3106, 2002.

[71] R. Tronci, G. Giacinto, and F. Roli. Designing multiple biometric systems : Measures of ensemble effectiveness. Engineering Applications of Artificial Intelligence, 22 :pp. 66–78, 2009.

[72] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics, 1947.

[73] R. Cappelli, D. Maio, and D. Maltoni. Synthetic fingerprint-database generation. In International Conference on Pattern Recognition (ICPR), pp. 744–747, 2002.

[74] Franziska Wolf, Tobias Scheidat, and Claus Vielhauer. Study of applicability of virtual users in evaluating multimodal biometrics. In Multimedia Content Representation, Classification and Security, volume 4105, pages 554_561. Springer Science & Business, New York, VG-Wort : 13 Normseiten, 2006.

[75] ISO/IEC FCD 19792. Information technology – security techniques – security evaluation of biometrics, 2008.

[76] C. Dimitriadis and D. Polemi. Application of multi-criteria analysis for the creation of a risk assessment knowledgebase for biometric systems. In international conference on biometric authentication (ICB), volume 3072, pp. 724–730, 2004.

[77] EBIOS. Expression des besoins et identification des objectifs de sécurité (EBIOS). Technical report, L'Agence nationale de la sécurité des systèmes d'information (ANSSI), 2004.

[78] D.Y. Yeung, H. Chang, Y.M. Xiong, S. George, R. Kashi, T. Matsumoto, and G. Rigoll. SVC2004 : First International Signature Verification Competition. In International Conference on Biometric Authentication (ICBA'04), pp. 16 – 22, 2004.

[79] J. P. Phillips, T. W. Scruggs, A. J. O'toole, P. J. Flynn, K.W. Bowyer, C. L. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale results. Technical report, National Institute of Standards and Technology, 2007.

[80] Tahirou Djara, Abdou-Aziz Sobabe, Macaire Bienvenu Agbomahena and Antoine Vianou. Practical method for evaluating the performance of a biometric algorithm. © ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2019, Springer Nature Switzerland AG 2019, AFRICATEK 2018, LNICST 260, pp. 125–132, https://doi.org/10.1007/978-3-030-05198-3_11, 2019.

[81] Abdou-Aziz SOBABE, Tahirou DJARA and Antoine VIANOU. A framework for combination of sequential architecture and soft biometrics in multibiometric scores fusion. BioSMART 2019 Proceedings, 3rd International Conference on Bio-engineering for Smart Technologies, Paris, 24th-26th April 2019, ©2019 IEEE, pp. 164-167, DOI : 10.1109/BIOSMART.2019.8734247, 2019.

[82] Anil K. Jain, Karthik Nandakumar, Xiaoguang Lu, and Unsang Park. Integrating faces, fingerprints, and soft biometric traits for user recognition.

Proceedings of Biometric Authentication Workshop, LNCS 3087, Prague, May 2004, pp. 259-269, DOI:10.1007/978-3-540-25976-3_24, 2004.

[83] Common Criteria. Common Criteria for Information Technology Security Evaluation. Version 3.1, Revision 4, September 2012, 2012.

[84] Antitza Dantcheva, Carmelo Velardo, Angela D'Angelo and Jean–Luc Dugelay. Bag of soft biometrics for person identification: New trends and challenges. Multimedia Tools and Applications 51(2), January 2010), DOI : 10.1007/s11042-010-0635-7, 2010.

[85] Karthik Nandakumar, Yi Chen, Anil K. Jain and Sarat C. Dass. Quality-based Score Level Fusion in Multibiometric Systems. 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, pp. 473-476, DOI: 10.1109/ICPR.2006.951, 2006.

[86] Engin Erzin, Yücel Yemez and A. Murat Tekalp. Multimodal Speaker Identification Using an Adaptive Classifier Cascade Based on Modality Reliability. IEEE transactions on multimedia, vol. 7, n° 5, 2005.

[87] Mejda Chihaoui, Akram Elkefi, Wajdi Bellil and Chokri Ben Amar. A survey of 2D face recognition techniques. Computers 2016, 5, 21, September, doi:10.3390/computers5040021, www.mdpi.com/journal/computers, 2016, Visited on Saturday 28 september (2019).

[88] N. K. Ratha, J. H. Connell, and R. M. Bolle. An analysis of minutiae matching strength. In Audio- and Video-Based Biometric Person Authentication, pp. 223–228, 2001.

[89] Anil K. Jain, Karthik Nandakumar and Abhishek Nagar. Biometric template security. Published in EURASIP Journal on Advances in Signal Processing, Special Issue on Biometrics, January 2008, pp. 1-20, DOI:10.1155/2008/579416, 2008.

[90] Antitza Dantcheva, Jean-Luc Dugelay and Petros Elia. Person recognition using a bag of facial soft biometrics (BoFSB). 2010 IEEE International Workshop on Multimedia Signal Processing, Saint Malo, pp. 511-516, DOI: 10.1109/MMSP.2010.5662074, 2010.

[91] N. G. Jablonski, G. Chaplin. The evolution of human skin coloration. Journal of Human Evolution, 39, 57-106. DOI: 10.1006/jhev.2000.0403, 2000.

[92] Sanjay Kr. Singh, D. S. Chauhan, Mayank Vatsa and Richa Singh. A Robust Skin Color Based Face Detection Algorithm. Tamkang Journal of Science and Engineering, Vol. 6, No. 4, pp. 227-234, 2003.

[93] Tarek Abd El-Hafeez. A New System for Extracting and Detecting Skin Color Regions from PDF Documents. International Journal on Computer Science and Engineering (IJCSE), Vol. 02, No 09, pp. 2838-2846, 2010.

[94] Amit Kumar and Shivani Malhotra. Real-time Human Skin Color Detection Algorithm using Skin Color Map. Proceedings of the 9th INDIACom, INDIACom-2015, 2nd International Conference on Computing for Sustainable Global Development, 2015.

[95] Pallavi R. Patil, A. H. Karode and S. R. Human skin detection using image fusion. International Journal of Electronics and Communication Engineering and Technology (IJECET), Communication Engineering and Technology, Volume 8, Issue 4, July-August 2017, pp. 13–21, 2017.

[96] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat and J. Jatakia. Human Skin Detection Using RGB, HSV and YCbCr Color Models. ICCASP/ICMMD-2016, Advances in Intelligent Systems Research, Vol. 137, pp. 324-332, 2017.

[97] Ashish Kumar and P. Shanmugavadivu. Skin Detection Using Hybrid Colour Space of RGB-H-CMYK. Progress in Advanced Computing and Intelligent Engineering, Advances in Intelligent Systems and Computing, vol 713, © Springer 2019, Singapore, https://doi.org/10.1007/978-981-13-1708-8_7, 2019.

[98] Sangho Yoon, Michael Harville, Harlyn Baker and Nina Bhatii. Automatic Skin Pixel Selection And Skin Color Classification. International Conference on Image Processing (ICIP), November 2006, © 2005 IEEE, DOI: 10.1109/ICIP.2006.312630, 2006.

[99] Kishor Bhoyar and Omprakash Kakde. Skin Color Detection Model Using Neural Networks and its Performance Evaluation. Journal of Computer Science, December 2010, DOI: 10.3844/jcssp.2010.963.968, 2010.

[100] Simone Bianco, Francesca Gasparini and Raimondo Schettini. Adaptive Skin Classification Using Face and Body Detection. IEEE Transactions on Image Processing, August 2015, DOI: 10.1109/TIP.2015.2467209, 2015.

[101] Markus Weber, Frontal Face Database, California Institue of Technology, 1999, http://www.vision.caltech.edu/html-files/archive.html/, Visited on Saturday 28 september (2019).

[102] M. M. Madhuram, B. P. Kumar, L. Sridhar, N. Prem, V. Prasad. Face Detection and Recognition Using OpenCV. International Research Journal of Engineering and Technology (IRJET), Volume-05 Issue-10, 2018.

[103] Vince Tabora. Face Detection Using OpenCV With Haar Cascade Classifiers, https://becominghuman.ai/face-detection-using-opencvwith-haar-cascade-classifiers-941dbb25177, 2019 (accessed 13 February 2020).

[104] K. Nikolskaia, N. Ezhova, A. Sinkov and M. Medvedev. Skin Detection Technique Based on HSV Color Model and SLIC Segmentation Method. Ural Workshop on Parallel, Distributed, and Cloud Computing for Young Scientists (Ural-PDC), 2018.

[105] Q. Zhu, K.-T. Cheng, C.-T. Wu, Y.-L. Wu. Adaptive Learning of an Accurate Skin-Color Model. Sixth IEEE International Conference on Automatic Face and Gesture Recognition. pp. 37-42. DOI: 10.1109/AFGR.2004.1301506, 2004.

[106] S. Midha, R. Vijay, S. Kumari. Analysis of RGB and YCbCr color spaces using wavelet transform. IEEE International Advance Computing Conference (IACC), DOI: 10.1109/IAdCC.2014.6776461, 2014.

[107] Georges Valensi et al. HSL (hue, saturation, lightness) and HSV (hue, saturation, value). https://en.wikipedia.org/wiki/HSL_and_HSV, 2020 (accessed 13 March 2020).

[108] C. C. Fung and R. Chamchong. A Review of Evaluation of Optimal Binarization Technique for Character Segmentation in Historical Manuscripts. Third International conference on Knowledge Discovery and Data Mining, DOI : 10.1109/WKDD.2010.110, 2010.

[109] T. Velmurugan and T. Santhanam. Computational Complexity between K-Means and K-Medoids Clustering Algorithms for Normal and Uniform Distributions of Data Points. Journal of Computer Science 6 (3): pp. 363-368, 2010.

[110] Charu C. Aggarwal and Chandan K. Reddy. DATA CLUSTERING Algorithms and Applications. © 2014 by Taylor & Francis Group, LLC, 2014.

[111] Hans-Hermann Bock. Origins and extensions of the k-means algorithm in cluster analysis. Electronic Journ@l for History of Probability and Statistics, Vol 4, n°2, December 2008, 2008.

[112] W. R. Tan, C. S. Chan, P. Yogarajah and J. Condell. A Fusion Approach for Efficient Human Skin Detection. IEEE transactions on industrial informatics. vol.8(1), pp. 138-147, 2014.

[113] Tahirou Djara, Marc Kokou Assogba, Amine Naït-Ali and Antoine Vianou. Recalage D'images D'empreintes Digitales En Biometrie Sans Contact. J. Rech. Sci. Univ. Lomé (Togo), 2013, Série E, 15(2), pp. 133-145, 2013.

[114] Ceyhun Ozgur, Taylor Colliau, Grace Rogers, Zachariah Hughes, Elyse "Bennie" Myer-Tyson. MatLab vs. Python vs. R. Journal of Data Science Vol. 15 N° 3 (2017), pp. 355-372, 2017.

[115] Gregory Piatetsky. KDnuggets : Top Analytics, Data Science, Machine Learning Software. https://www.kdnuggets.com/2019/05/poll-top-data-science-machine-learning platforms.html, Visited on Saturday 28 september (2019).

[116] N. Ahuja, M. Yang and D. J. Kriegman. Detecting faces in images: A survey. Pattern Analysis and Machine Intelligence, IEEE, 2002.

[117] Din Siyi Xu Qing and Yang Jie. Texture segmentation using LBP embedded region competition. Electronic Letters on Computer Vision and Image Analysis, 2005.

[118] M. J. Jones and P. Viola. Bust real-time face detection. Journal of Computer, 2004.

[119] D. G. Lowe. Distinctive image features from scale-invariant keypoints. Journal Comput, 2004.

[120] D. J. Kriegman P. N. Belhumeur, J. P. Hespanha. Eigenfaces vs fisherfaces: Recognition using class specific linear projection. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 1997.

[121] A. P. Pentland and M. Turk. Eigenfaces for recognition. Journal of cognitive neuroscience, 1991.

[122] Lahiru Dinalankara. Face Detection and Face Recognition Using Open Computer Vision Classifires. Robotic Visual Perception and Autonomy, Faculty of Science and Engineering, Plymouth University, August 4, 2017.

[123] Abdenour Hadid, Timo Ahonen, Matti Pietikainen. Face description with local binary patterns: application to face recognition, 2006.

[124] Abdur Rahim, Najmul Hossain, Tanzillah Wahid and Shafiul Azam. Face recognition using local binary patterns (LBP). Global Journal of Computer Science and Technology Graphics & Vision, Volume 13 Issue 4 Version 1.0, 2013

[125] Geoffrey E. Hinton, Alex Krizhevskyn and Ilya Sutskever. Imagenet classification with deep convolutional neural networks, 2016.

[126] Andrew G. Howard, Menglong Zhu, Dmitry Kalenichenko, and Bo Chen. Mobilenets : Efficient convolutional neural networks for mobile vision applications, arXiv:1704.04861v1 [cs.CV], 2017.

Annexes

Publications scientifiques

Les travaux effectués ont fait l'objet de quatre publications d'articles à savoir :

1- *Practical Method for Evaluating the Performance of a Biometric Algorithm*

Editeur: SPRINGER
Conférence: ***AFRICATEK*** tenue à Cotonou en mai 2018
Référence DBLP: https://dblp.org/db/conf/africatek/africatek2018.html
DOI : https://doi.org/10.1007/978-3-030-05198-3_11

2- *A Framework For Combination Of Sequential Architecture And Soft Biometrics In Multibiometric Scores Fusion*

Editeur: IEEE
Conférence: ***BioSMART*** tenue à Paris en avril 2019
DOI (IEEE Xplore) : 10.1109/BIOSMART.2019.8734247

3- *Incorporating Metadata in Multibiometric Score-Level Fusion: an Optimized Architecture*

Journal : ***International Journal of Innovative Technology and Exploring Engineering (IJITEE)***; Novembre 2019
Référence SCOPUS: https://www.scopus.com/sourceid/21100889409
DOI: 10.35940/ijitee.A4118.119119

4- *Biometric System Vulnerabilities: A Typology of Metadata*

Journal : ***Advances in Science, Technology and Engineering Systems Journal (ASTESJ)***; Janvier 2020
Référence SCOPUS: https://www.scopus.com/sourceid/21100898760
DOI : https://dx.doi.org/10.25046/aj050125

Participations aux rencontres à caractère scientifique avec communication

1. Journées scientifiques organisées dans le cadre de la célébration du quarantenaire de l'Ecole Polytechnique d'Abomey-Calavi (EPAC), du 12 au 15 décembre 2017 : Communication sur le thème : *Reconnaissance des empreintes digitales en biometrie sans contact : mise en œuvre d'un systeme d'evaluation de performance* ;
2. Salon des TIC pour l'agriculture, Ecole de Sociologie Rurale et de Vulgarisation Agricole (Université Nationale d'Agriculture), Porto-Novo,

Benin, du 12 au 14 Novembre 2019: Communication sur le thème : *La stratégie nationale e-Agriculture du Bénin sur la période 2020-2024*.

3. Journée portes ouvertes de l'Ecole Doctorale des Sciences de l'Ingénieur (ED-SDI/UAC) dans le cadre de la 35ème édition de la journée mondiale de l'habitat sur le thème "Un logement pour tous : Amélioration de l'environnement urbain", lundi 05 octobre 2020 : Présentation d'un poster sur le sujet : *Mise en place d'un système d'authentification par la biométrie multimodale sans contact : Une approche efficace face aux épidémies*.

Table des matières