République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université Mohamed el-Bachir el-Ibrahimi Bordj Bou Arréridj
Faculté de Mathématique et Informatique

MEMOIRE

Présente en vue de l'obtention du diplôme
Master en informatique

Spécialité : Ingénierie de l'informatique décisionnelle.

THEME :

Validation d'un algorithme de protection de template biométrique

Présenté Par :

- LAGGOUNE Abdellatif
- BENAMARA Islam

Soutenue le : 2021

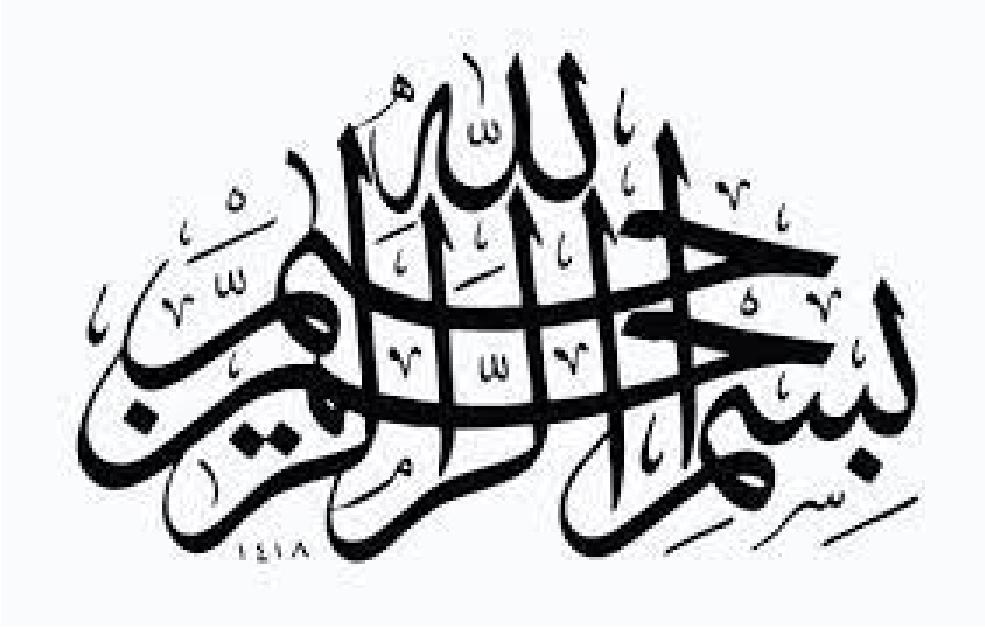
Devant le jury composé de :

Président Mr :

Examinateur Mr :

Examinateur Mr :

Encadreur Mr : Foudil Belhadj


بسم الله الرحمن الرحيم
١٤١٨

Résumé

La biométrie est largement utilisée dans les systèmes de sécurité . Sa principale modalité largement utilisée est l'empreinte digitale, en raison de son coût qui peut être relativement faible et de sa précision très optimale. Malgré le développement atteint par les systèmes d'identification par empreintes digitales, certaines problématiques persistent et nécessitent encore des recherches supplémentaires.

Dans ce mémoire nous avons abordé la problématique de la protection des données biométriques, pour le quel nous avons proposé un algorithme basé sur la biométrie résiliable. Cette méthode consiste à la réduction des attributs des minuties initialement extraits de la texture des empreintes digitales pour empêcher toute sorte d'attaque. L'opération de réduction d'attributs est basée sur la technique de décomposition d'une matrice en valeurs singulières. Les valeurs propres, résultats de la décomposition, sont utilisées comme attributs durant l'opération de matching.

Mots clés : Biométrie, Empreinte digitale, Systèmes de sécurités, Biométrie résiliable, SVD, Décomposition en valeurs Singulières.

Abstract

Biometrics is widely used in the most improved security systems. Its main widely used modality is the fingerprint, due to its high accuracy, and its cost that can be relatively low. In spite of the development achieved in the fingerprint identification systems, some problems persist and require further research.

In this thesis we have addressed the problem of biometric data protection, for which we have proposed an algorithm based on cancelable biometrics. It is based on the minutiae attributes reduction initially extracted from the texture around each minutia to prevent any attack. The attributes reduction step is based on the SVD technique (Singular Values decompostion). The obtained eigen values are used as matching attributes.

Keywords : Biometrics, fingerprint, security systems, cancelable biometrics, SVD.

ملخص

تستخدم القياسات الحيوية على نطاق واسع في أنظمة الأمان الأكثر تطورًا. الطريقة الرئيسية الاكثر استعمالا هي البصمة ، نظرًا لنهجها الرائع من قبل المستخدمين ، فيما يتعلق بالدقة ، والتكلفة التي يمكن أن تكون منخفضة نسبيًا. على الرغم من تطوير أنظمة التعرف على بصمات الأصابع ، لا تزال بعض المشكلات قائمة ولا تزال تتطلب بحثًا إضافيًا.

تناولنا في هذه الأطروحة مسألة حماية بيانات القياسات الحيوية ، والتي اقترحنا لها خوارزمية بناءً على القياسات . الحيوية القابلة للتعيين . تقوم هاته الطريقة على فكرة تخفيض عدد الخصائص المتاحة لكل نقطة مميزة (minutia) لمنع أي هجوم على البيانات. . أولا نقوم باستخراج بعض خصائص النسيج المحيط بكل نقطة مميزة ثم نستعمل طريقة التفكيك إلى قيم انفرادية (SVD) لخفض عدد الخصائص والاقتصار على القيم الذاتية الناتجة عن عملية التفكيك. مجموعة البيانات المتحصل عليها من القيم الذاتية تشكل الخصائص التي تستعمل لتحديد هوية الشخص.

كلمات مفتاحية :

القياسات الحيوية ، المطابقة ، بصمات الأصابع.

Dédicace

“

À mes chers parents Abdelmoumen et Bouchra, que nulle dédicace ne puisse exprimer mes sincères sentiments, pour leur patience illimitée et leur encouragement contenu, qu'Allah leur préserve une longue vie et bonne santé.

À mes soeurs Shahrazed et Alia, et ma nièce Naya pour leur grand amour et leur soutien moral qui m'a donné d'un grand souffle pour achever ce travail.,

À mes amis Boualem, Mohamed, Amine, Ilyes, Walid, Mehdi, Djalel et Maroua qui n'ont jamais cessés de me soutenir tout au long de mon parcours scolaire

À tous ceux qui me sont chers, à vous tous

Merci.

”

- Abdellatif

“

À mes très chers parents pour leur soutient durant tout mon cursus scolaire et qui m'ont permis de réussir dans mes études.

À tous mes amis Et mes frères
Abbas,Rachid,Farid,Abdelmalek,Abdelraouf,Youcef et mes sœurs,

À toute la famille Ben Amara.

À tous les professeurs et enseignants que j'ai eu durant tout mon cursus scolaire et qui m'ont permis de réussir dans mes études

À toute la promotion deuxième master Decisionnelle

À toute personne ayant contribué à ce travail de près ou de loin

Merci.

”

- Islam

Remerciements

Tout d'abord, ont remercient Allah le tout puissant de nous avoir donné le courage et la patience nécessaires à mener ce travail à son terme.

Ont tient à remercier tout particulièrement notre encadreur **Mr. BELHADJ Foudil**, pour l'aide compétente qu'il nous a apportée, pour sa patience et son encouragement. Son œil critique m'a été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections.

Ont tient à remercier également tous nos enseignants pour leur aide immense, la qualité de leur suivie ainsi que pour tous les conseils et les informations qu'ils nous ont prodigués avec un degré de patience et de professionnalisme sans égal.

Que les membres de jury trouvent, ici, l'expression de nos sincères remerciements pour l'honneur qu'ils nous font en prenant le temps de lire et d'évaluer ce travail.

Ont souhaite aussi remercier l'équipe pédagogique et administrative de l'Université El bachir El Ibrahimi pour leurs efforts dans le but de nos offrir une excellente formation.

Pour finir, je souhaite remercier toute personne ayant contribué de près ou de loin à la réalisation de ce travail.

Table des matières

Table des figures

Introduction générale

Contexte

De nos jours, nous vivons dans un monde extrêmement petit. Nous sommes constamment en mouvement, et connectés avec d'autres individus. Les technologies de l'information ont eu un très fort impact sur notre quotidien en particulier les appareils mobiles. Dans la plupart des sociétés comme les banques les hôtels ou même dans dans les administration scolaires ou de l'armée nationale, tous les service sont fournis électroniquement par des machines intelligentes maniable à distance.

Problématique

Une des conséquences directe du monde numérique est la sécurité des systèmes informatiques. Comme solution impérative, la biométrie est utilisée pour faire face aux attaques ciblant le fonctionnement de ces systèmes ou les données manipulées. Cependant, les systèmes biométriques ne font pas l'exception des attaques. Les données biométriques peuvent être violées pour usurper l'identité de leurs porteurs.

Objectifs

La protection de template biométrique est une étape primordiale dans un système d'identification biométrique. Elle permet non seulement de sécuriser les données biométriques d'un individu et d'empêcher de les reproduire, mais aussi de rendre les traits biométriques résiliables.

Il s'agit dans ce projet de proposer et de valider un algorithme de sécurisation des données biométriques en démontrant sa robustesse contre les multiples attaques.

Organisation du mémoire

Ce mémoire est organisé en trois chapitres :

Dans le premier chapitre "**Les généralités**" on découvre tout ce qui concerne la biométrie ainsi que ses multiples modalités, les systèmes d'identification et authentification, on a aussi vu les points vulnérables de ces systèmes biométriques, ainsi que leurs inconvénients, après on va mettre l'accent sur une modalité bien précise qui est l'empreinte digitale cette dernière nous sera bien utile pour la suite, et on clôture ce chapitre avec une conclusion.

Le deuxième chapitre "**Protection des données biométriques**" commence avec une petite introduction qui va nous guider vers une potentielle solution qui est la biométrie résiliable qui sera notre méthode proposée avec tous les défis à relever avant de pouvoir l'exécuter.

Le troisième chapitre "**Tests et résultats**" dans ce chapitre nous allons tout d'abord voir l'environnement les outils tel que le langage, l'éditeur et la base de données utilisés, après ça on pourra faire une analyse explicative avec des statistiques et des données.

Chapitre 1

Généralités

1.1 La biométrie

En tant qu'humains, nous utilisons tous nos aptitudes naturelles pour reconnaître les gens à travers leurs voix, leurs visages et d'autres caractéristiques. Les machines, d'autre part, peuvent être programmées et ordonnées (instruites) pour indiquer comment utiliser la même information observable pour effectuer la reconnaissance humaine. Les progrès technologiques, en particulier en biométrie, contribuent à combler l'écart entre la perception humaine et la reconnaissance de la machine [1].

1.1.1 Définition

C'est une méthode automatisée de reconnaissance d'une personne basée sur une caractéristique physiologique ou comportementale [2]. Elle consiste à vérifier ou déterminer l'identité d'un individu à partir de ses caractéristiques biologiques (comme l'ADN), comportementales (comme la voix) ou morphologiques (comme l'empreinte digitale) [3]. La biométrie permet l'identification ou l'authentification d'une personne sur des bases de données reconnaissables et vérifiables qui lui sont propres [4].

1.2 Les modalités biométriques

Les caractéristiques biométriques par lesquelles il est possible de vérifier l'identité d'un individu sont appelées modalités biométriques.

Ces modalités sont basées sur l'analyse des données liées à l'individu et sont généralement classées en trois catégories : biométrie biologique, comportementale, et morphologique.

- La biométrie biologique se base sur l'analyse des données biologiques liées à l'individu (ADN).

- La biométrie comportementale se base sur l'analyse de comportements d'un individu (manière de marcher, dynamique de frappe au clavier, etc.).

- La biométrie morphologique se base sur les traits physiques particuliers qui, pour toutes personnes, sont permanents et uniques (empreinte digitale, visage, etc.) [3]

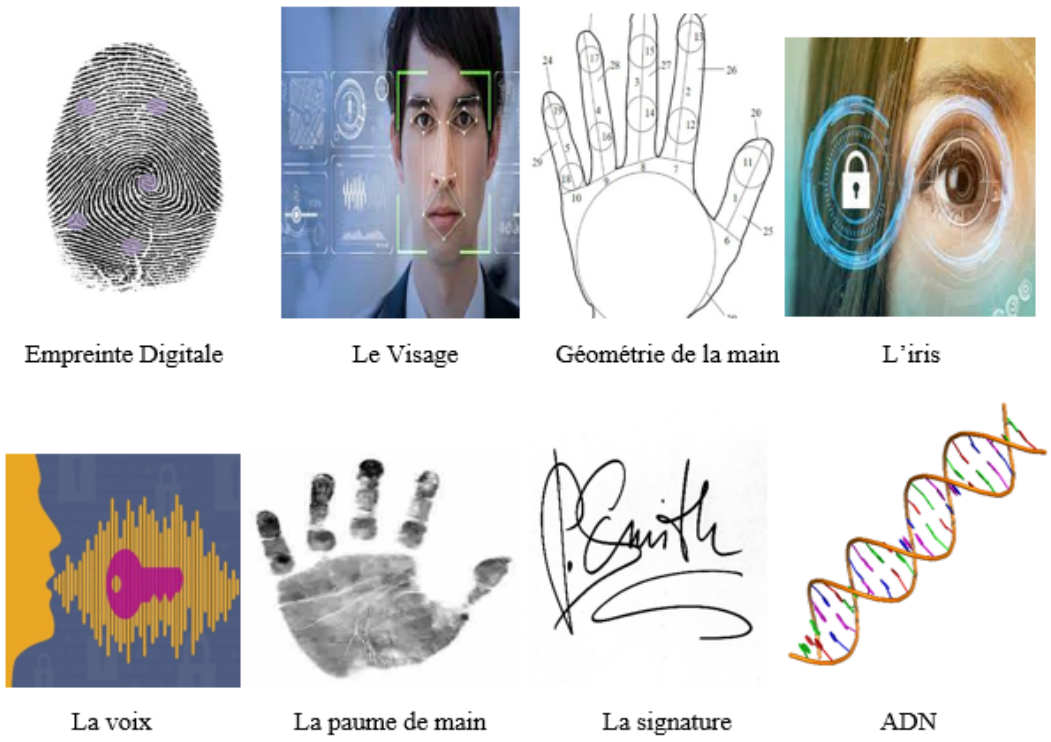


FIG. 1.1 : Quelques modalités biométriques

- **Le visage :** La mesure de cette modalité est non intrusive (pas de contact physique) et bien acceptée du grand public du fait de sa similarité avec le processus humain de reconnaissance des personnes. Pour la plupart des traitements une caméra standard suffit, cependant, le système est alors soumis à certaines limitations telles que les variations de luminosité, les variations du visage (vieillissement, barbe, lunettes, etc) [5].

- **La géométrie de la main :** Contrairement à d'autres modalités, la mesure de la géométrie de la main n'est pas perturbée par la chaleur, l'humidité ou les impuretés de la peau et peut donc être effectuée dans n'importe quel environnement. Malgré ça, cette modalité est peu précise et présente un fort risque de similitude entre deux personnes [5].

- **L'iris :** L'iris est présent sur la face antérieur du globe oculaire et s'apparente à un "diaphragme" contrôlant l'intensité de lumière captée. Un des avantages de la

mesure de l'iris est la grande unicité du motif, elle présente l'avantage de posséder des caractéristiques extrêmement différentes d'une personne à l'autre. Le port de lunettes ou de lentilles de contact n'affecte pas la mesure tandis que la lumière ambiante amène des perturbations visibles. Les principales difficultés sont donc liées à l'éclairage, la distance du sujet au système [5].

- **La voix :** Le son de notre voix est directement influencé par des facteurs physiques comme le nez, la bouche, les cordes vocales, ou par l'état émotionnel, la langue natale, les conditions médicales, et d'autres paramètres. La qualité de l'enregistrement dépend du capteur ou de l'environnement (écho, bruit), c'est pour cette raison que cette modalité ne s'est pas fortement répandue [5].

- **La paume de main :** Cette technique possède de nombreux points communs avec les empreintes digitales. Les motifs de surface d'un simple doigt fournissent déjà une grande quantité d'informations, la paume est constituée de diverses régions qui peuvent être scannées, stockées et traitées indépendamment selon les besoins [5].

- **La signature :** Il est question de mesurer la manière dont une personne signe à l'aide d'un stylet et d'une tablette graphique. En plus de l'aspect général de la signature, les caractéristiques mesurées incluent les pauses, les variations de rythme, la pression, la direction des traits et la vitesse. Cette signature change dans le temps et est influencée par la condition physique du sujet et le système d'écriture utilisé, ce qui rend cette modalité peu utilisée [5].

- **L'ADN :** Présent dans toutes les cellules, l'Acide Désoxyribonucléique (ADN) est une macromolécule biologique contenant toute l'information génétique permettant le développement, le fonctionnement et la reproduction des êtres vivants. Plus souvent utilisée en médecine légale, cette méthode isole et compare les séquences de segments d'ADN de différents individus, avec un risque de similitude entre deux personnes de moins d'un pour cent milliards.. Très coûteuse en temps de traitement et en matériel, cette méthode est peu commode dans le prélèvement des échantillons [5].

1.3 Système d'identification biométrique

C'est le procédé permettant de déterminer l'identité d'une personne. Il ne comprend qu'une étape : L'utilisateur fournit un échantillon biométrique qui va être comparé à tous les échantillons biométriques contenus dans la base de données biométriques du système. Si l'échantillon correspond à celui d'une personne de la base, on renvoie son numéro d'utilisateur. Sinon l'identification échoue.

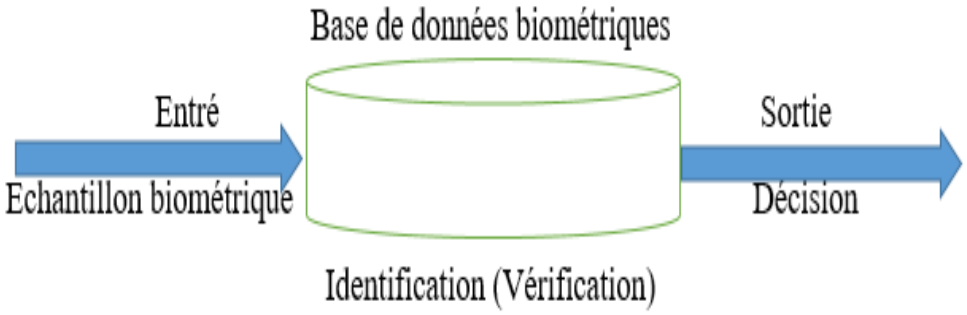


FIG. 1.2 : L'architecture d'un système d'identification biométrique

1.4 Système d'authentification biométrique

C'est le procédé permettant de vérifier l'identité d'une personne. Il comprend deux étapes : - L'utilisateur fournit un identifiant « Id » au système de reconnaissance (par exemple un numéro d'utilisateur). - L'utilisateur fournit ensuite un échantillon biométrique qui va être comparé à celui correspondant à l'utilisateur « Id » contenu dans la base de données biométrique du système. Si la comparaison correspond, l'utilisateur est authentifié [4].

1.5 Les points vulnérables d'un système biométrique

Un système biométrique peut être modélisé dans le cadre du système de reconnaissance d'un pattern. Les étapes d'un tel système sont illustrées avec les possibilités d'attaques dans la figure 1.7.

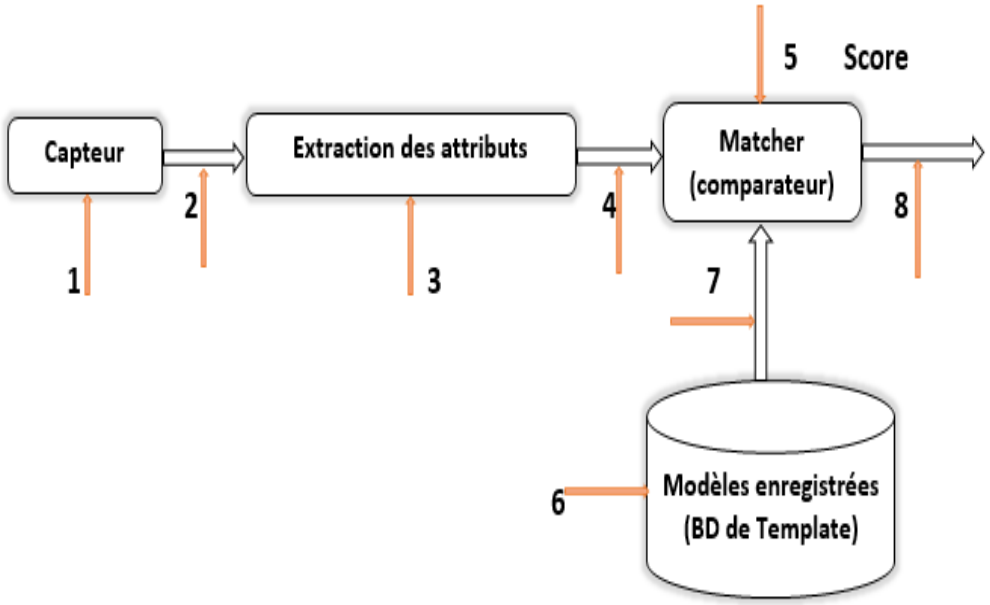


FIG. 1.3 : Points d'attaques possibles dans un système biométrique

Maintenant on va expliquer brièvement les huit cas d'attaques possibles représentées dans cette figure 1.7 :

1. Ce mode d'attaque consiste à présenter de fausses données biométriques au capteur, une reproduction possible de la caractéristique biométrique est présentée en entrée du système. Il s'agit par exemple d'un faux doigt, d'un masque facial ou même d'une copie de la signature.

2. Ce mode d'attaque resoumet des signaux biométriques numérisés enregistrés précédemment : un signal enregistré est rejoué au système en évitant le capteur. Par exemple présenter une ancienne copie d'une image d'empreinte digitale, ou d'un signal audio précédemment enregistré.

3. Remplacer le processus d'extraction d'attributs : l'extracteur d'attributs est attaqué à l'aide d'un Cheval de Troie (logiciel en apparence légitime, mais qui contient une fonctionnalité malveillante), de façon à ce qu'il produise des attributs présélectionnés par l'intrus.

4. Falsifier la représentation des attributs biométriques : les attributs extraits du signal d'entré sont remplacées par d'autres attributs différents et frauduleux. Dans la plupart des cas, les deux étapes d'extraction des attributs et d'appariement (Matcher) sont inséparables ce qui fait que ce mode d'attaque est extrêmement difficile. Cependant, si les minuties sont transmises à un comparateur par exemple sur internet, cette menace devient donc très réelle. Un Cloud pourrait espionner dans la pile TCP/IP (Transmission Control Protocol/Internet Protocol) et modifier certains paquets.

5. Corrompre le Matcher : le Matcher est attaqué et corrompu de manière qu'il produise des score de correspondance avec l'intrus.

6. Déformer les modèles stockés (Template) : sachant que la base de données peut être locale ou distante. Les données peuvent être réparties sur plusieurs serveurs. Dans ce cas, l'attaquant peut essayer de modifier un ou plusieurs modèles contenus dans la base de données ce qui amènerait soit à l'autorisation d'un individu frauduleux, ou bien le refus de service à la personne associée au modèle (Template) déformé. Un système basé sur carte à puce est particulièrement vulnérable face à ce type d'attaque.

7. Attaquer le canal entre les modèles stockés et le Matcher : Les modèles stockés sont envoyés au matcher via un canal de communication. Les données qui transitent par ce canal peuvent être interceptées et modifiées.

8. Annuler la décision finale : si la décision finale du match peut être annulée par le pirate, alors tout le système d'authentification est mis en cause. Même si le cadre de reconnaissance de Pattern présente d'excellentes performances, il a toutefois été rendu inutile par le simple fait de passer outre le résultat du match.

1.6 L'empreinte digitale

Une empreinte digitale est le dessin formé par les plans de la peau des doigts. Elles sont particulières et immuables, elles ne se modifient donc pas au cours du temps (sauf par accident) mis à part leur qualité qui peut se dégrader. Une empreinte digitale se constitue d'un ensemble de lignes régulièrement parallèles créant un motif unique pour chaque personne (pattern). On distingue les crêtes, ce sont les lignes en contact avec une surface au touché et les vallées, ce sont les creux entre deux crêtes. A l'intérieur de ce motif, il y a un très grand nombre d'éléments qui nous différencient les uns des autres. Ces caractéristiques sont formées par le flux des crêtes formant l'empreinte. La figure 3 illustre un exemple de ces caractéristiques.

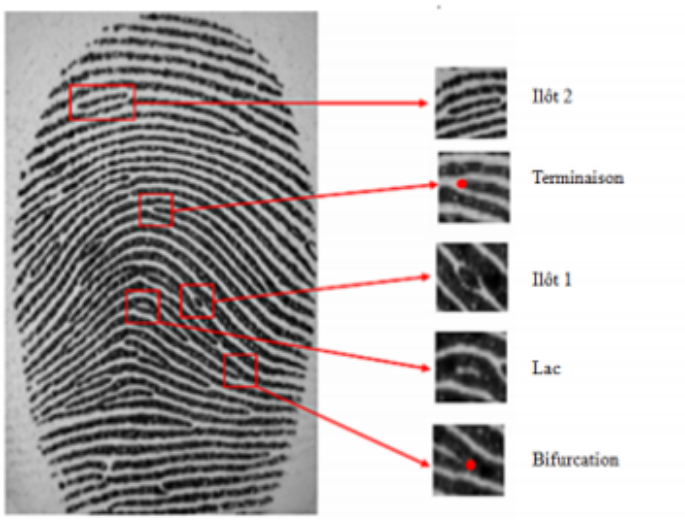


FIG. 1.4 : Quelques caractéristiques des empreintes digitales

[6]

Les empreintes digitales se distinguent en trois niveaux de précision telle que le montre la figure ci-dessous.

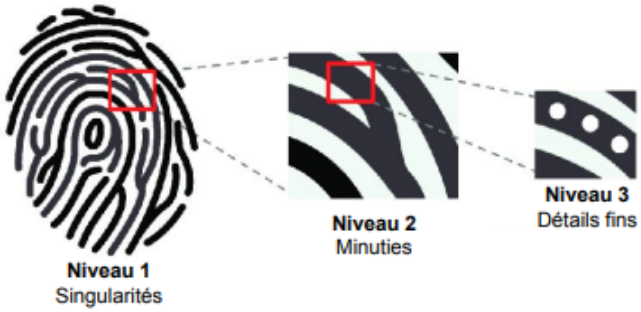


FIG. 1.5 : Les niveaux de caractéristiques fournis par les empreintes digitales [6]

Niveau 1: Le niveau 1 est le niveau le plus abstrait en caractéristiques. Il est seulement composé de formes générales, appelées singularités. On distingue trois grands groupes : les boucles, les verticilles et les arches. Chacun de ces groupes possède diverses variations. Ces singularités sont approximées par l'estimation de l'orientation locale des crêtes et peuvent être détectées à partir d'images basses résolution.

Niveau 2: Le niveau 2 regroupe des informations plus précises sur ces crêtes, il est composé de données précises telles que les débuts et les fins de lignes (points rouges), ainsi que les bifurcations et les croisements de lignes (points verts). Comme le montre la figure ci-dessous.

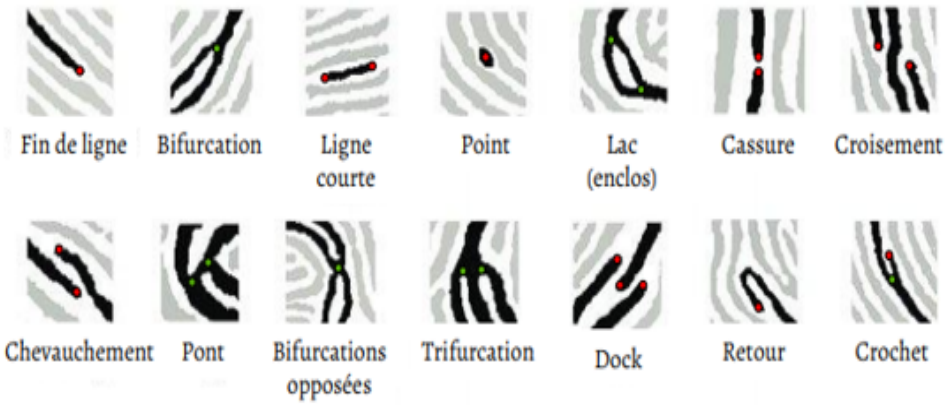


FIG. 1.6 : Caractéristiques de niveaux 2 - Principales minuties [6]

Niveau 3: Ce niveau comporte des données beaucoup plus précises et pour les exploiter on nécessite des images de très haute résolution et de haute qualité. Comme le montre la figure 1.6, on distingue les pores de la peau, la forme précise du contour des crêtes, et d'autres données précises propres à la peau.

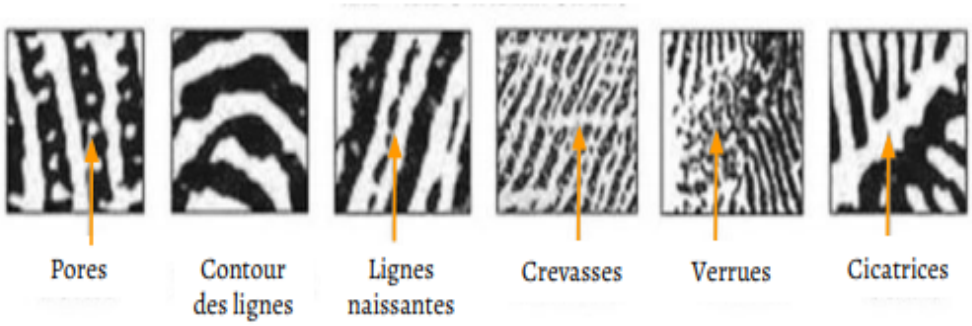


FIG. 1.7 : Caractéristiques de niveaux 3 - Principaux détails fins [6]

1.7 Les inconvénients de la biométrie

Malgré les avantages pertinents de l'utilisation de la biométrie, cette dernière soulève plusieurs préoccupations en termes de sécurité et de confidentialité, comme nous allons l'indiquer ci-dessous :

- La biométrie est authentique mais pas secrète : contrairement aux mots de passes et aux clés cryptées qui ne sont connus que de l'utilisateur, les données biométriques comme le visage, la signature, la voix ou même les empreintes digitales peuvent facilement être enregistrées et utilisées sans le consentement de leur utilisateur. Plusieurs fois des empreintes digitales artificielles ont été utilisées pour contourner les systèmes de sécurité biométriques.

- Les données biométriques ne peuvent pas être supprimées ou annulées : les mots de passes, les codes PIN, peuvent être réinitialisés s'ils sont compromis. Les cartes de crédit et les badges peuvent être remplacés en cas de vol. En revanche, les données biométriques sont associées de manière permanente à l'utilisateur et ne peuvent pas

être remplacées si elles sont corrompues. Bien qu'un utilisateur puisse enregistrer différentes empreintes digitales, le choix reste un choix limité de doigts, ce qui n'est toutefois pas possible pour les autres modalités biométriques

- Si un élément biométrique est perdu une fois, il est compromis pour toujours : certes, la biométrie évite d'avoir à mémoriser de multiples mots de passe. Mais cela signifie également que si un élément biométrique est compromis dans une application, toutes les applications où cette biométrie est utilisée sont compromises.

- Le jumelage peut être utilisé pour traquer des individus sans leur consentement : puisque la même biométrie peut être utilisées pour différentes applications et différents lieux, l'utilisateur peut potentiellement être traqué (suivi), si les organisations se réunissent et partagent leurs données biométriques respectives. Contrairement aux systèmes d'authentification traditionnels, l'utilisateur peut conserver différentes identités pour éviter cela. Donc le fait qu'une biométrie reste la même pose un problème de confidentialité.

1.8 Conclusion

La biométrie à pour but de limiter le processus de reconnaissance des formes mentales dans la mesure où elle identifie les personnes. Elle constitue une alternative plus sûre et plus fiable aux systèmes d'authentification classiques basés sur les schémas et les mots de passes. Les technologies biométriques utilisent les caractéristiques physiologiques et comportementales de l'homme pour reconnaître les individus.

Le processus de reconnaissance est basé sur deux étapes : la première, l'étape de recrutement, a pour objectif de permettre au système d'apprendre l'identité de la personne. Elle commence par l'extraction de certains attributs discriminants à partir des données détectées. Ces dernières seront compactées pour construire un modèle qui sera stocké dans une base de données. Ce modèle est une structure hautement représentative qui résume les caractéristiques biométriques individuelles. La deuxième étape, celle de la mise en correspondance (Matching), rappelle le modèle déjà stocké pour le comparer aux attributs

récemment extraits. En fonction des résultats de la comparaison, le système décide si l'individu est vraiment la personne qu'il revendique être ou non.

Le marché et l'industrie de la biométrie connaissent une forte accélération justifiée par la croissance des des services électroniques non supervisés qui nécessitent une authentification précise des individus, ainsi que l'augmentation des actes de fraude dans le monde entier. L'empreinte digitale est la modalité la plus dominante sur le marché, elle constitue un équilibre en termes de précision, de sécurité et de coût par rapport au autres modalités. [7]

Chapitre 2

Protection des données biométriques

2.1 Introduction

Après avoir vu les différents avantages ainsi que les inconvénients d'un système biométrique, on a constaté qu'il y'avait un sérieux problème lors de l'enregistrement des attributs biométriques dans la base de données, ces derniers pourrait être facilement usurpés où même falsifiés lors d'une attaque. Alors pour surmonter cette vulnérabilité, les deux communautés biométrique et cryptographique ont fait équipe dans le but de relever ce défi, plusieurs schémas de protection des modèles biométriques ont été proposés, notamment la biométrie résiliable qui a suscité énormément d'intérêt ces dernières années.

2.2 La biométrie résiliable

La biométrie résiliable utilise des données biométriques intentionnellement transformées au lieu des données biométriques originales pour l'identification. Sachant que la transformation est irréversible, les modèles biométriques originaux ne peuvent pas être récupérés à partir des Template transformés. Lorsqu'un ensemble de modèles biométriques est compromis, il peut être rejeté et il est possible d'en régénéré un autre [8].

Dans ce qui suit considérons M comme un algorithme de mise en correspondance d'empreintes digitales, qui est supposé être une boite de tolérance (Matcher), cette dernière agit sur deux biométries $x1$ et $x2$, et ainsi donne le score de correspondance ou de similitude $0 \leq M(x1, x2) \leq 1$.

Mais il y'a plusieurs défis à relever avant de réussir a concevoir une transformation résiliable C qui transforme la biométrie x en une biométrie C(x), nous allons présenter ces contraintes en termes d'exigences :

1. **L'enregistrement :** Pour que la transformation résiliable C soit reproductible d'une instance y1 d'un élément biométrique à l'instance suivante y2 de la même biométrie signaux biométriques y1 et y2 doivent être à chaque fois positionnés dans le même système de coordonnées. Pour les empreintes digitales, cela signifie que y1 et y2 doivent être tournées et traduites par une transformation de coordonnées T1 et T2, respectivement, de sorte que les signaux x1=T1(y1) et x2=T2(y2) se chevauchent. Par contre, il est nécessaire qu'avant de faire la transformation résiliable

C, les empreintes y1 et y2 soient enregistrées de sorte que les minuties correspondent le mieux possible.

2. **Tolérance à la variabilité intra-utilisateur :** un autre problème auquel on devra faire face, même après l'enregistrement, c'est la variabilité intra-utilisateur présente dans les signaux biométriques. Les attributs obtenus après la transformation doivent être robustes par rapport à cette variation, en ce sens la probabilité d'un faux rejet ne devrait pas augmenter dans le domaine transformé, c'est-à-dire lorsque x1 et x2 correspondent,

$$M(x1, x2) > t \Rightarrow M(C(x1), C(x2)) > t. \tag{2.1}$$

3. **Conservation de la mémoire :**la version transformée ne doit pas perdre son individualité. La force intrinsèque d'une biométrie transformée C(x) doit être comparable à l'élément biométrique original x, c'est-à-dire que la probabilité d'une fausse acceptation ne doit pas augmenter dans le domaine transformé. Idéalement ne devrions avoir que, si x1 et x2 ne correspondent pas,

$$M(x1, x2) \leq t \Rightarrow M(C(x1), C(x2)) \leq t, \tag{2.2}$$

où t est le seuil de décision.La combinaison de (1) et (2) est appelée préservation de la correspondance.

4. **Conception de la fonction de transformation :** la transformation C doit toutefois satisfaire les conditions suivantes :

   - La version transformée C(x) de doit pas correspondre à l'originale, c'est-à-dire,

$$M(C(x), x) \leq t, \tag{2.3}$$

     où encore une fois t est le seuil de décision pour l'algorithme de correspondance *M*.

   - Les multiples personnalités C1(x) et C2(x) générées à partir du même modèle (Template) x de la même biométrie ne doivent pas correspondre,

$$M(C1(x), C2(x)) \leq t, \tag{2.4}$$

Cette propriété de la transformation empêche la correspondance croisée entre les bases de données lorsque différentes transformations résiliables C1 et C2 sont utilisées.

- La biométrie ou le modèle d'origine x ne doit pas être récupérable à partir de celui transformé $C^{-}1$, c'est-à-dire que la transformation inverse C1 ne doit pas exister. Cette propriété préserve la confidentialité du modèle original x puisqu'il n'est pas stocké.

2.3 Méthode proposée

La méthode proposée passe par trois étapes principales :

1. Création de modèles basés sur la texture des empreintes digitales.
2. Création de modèles transformés.
3. Création du modèle résiliable.

La figure 1 montre un schéma de la construction de notre modèle protégé. L'empreinte digitale d'entrée doit tout d'abord être prétraitée pour extraire la liste des caractéristiques. Les sections suivantes résument chaque étape.

2.3.1 Création de modèles basés sur la texture des empreintes digitales :

Une empreinte digitale est une image texturée orientée en 2D avec des crêtes et des vallées entrelacées. Les minuties sont des discontinuités locales dans le motif des crêtes. Soit $M = mi(xi, yi, \alpha i)i = 1..N$. Une liste de minuties extraite de l'empreinte digitale d'entrée. (xi, yi) sont les coordonnées 2D et αi est son orientation.

Le descripteur associé a la minutie mi se base sur un ensemble de structure de points d'échantillonnage pris dans les circonférences d'un ensemble de L cercles . Le rayon du cercle Ci est ri et il contient Ki points d'échantillonnage pi,1, pi,2, ..., pi,Ki également répartis sur sa circonférence. Le point de départ pi,1 est situé sur l'axe des x en prenant

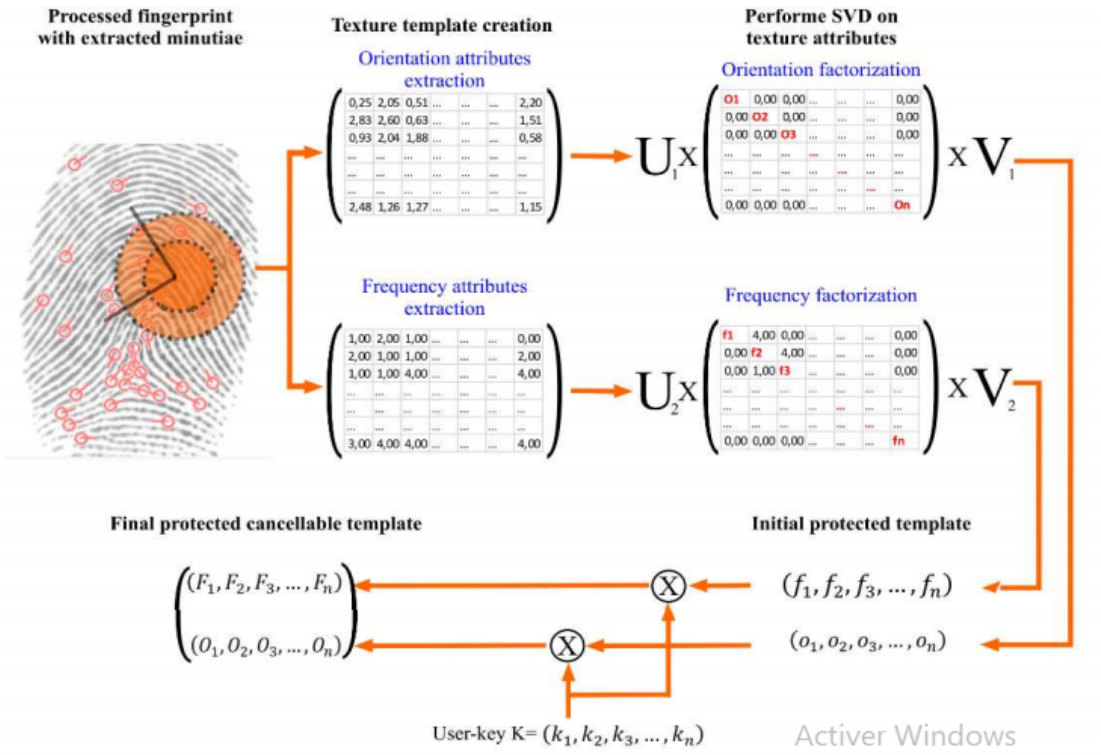


FIG. 2.1 : le bloc-diagramme pour construire le modèle résiliable protégé [9]

l'emplacement de la minutie comme origine et sa direction comme la direction positive de l'axe des x (voir la figure 2.1). Le nombre total de points d'échantillonnage est K tel que :

$$K = \sum_{i=1}^{L} Ki \tag{2.5}$$

Pour simplifier les notations suivantes, nous notons un point d'échantillonnage p_{ij} comme p_r, où r indique son indice dans la structure d'échantillonnage de la coque, vue comme un tableau arrangé en partant du cercle intérieur jusqu'à l'extérieur et du premier point d'échantillonnage au dernier dans le sens inverse des aiguilles d'une montre.

Les sous-sections suivantes décrivent comment extraire à la fois l'information d'orientation et de fréquence autour d'une minutie pour construire le descripteur de coque.

Extraction des attributs d'orientation :

Pour une minutie m donnée avec une structure d'échantillonnage consistant de K points d'échantillonnage, soit $\theta_i, (i = 1..K)$, est la valeur du champ d'orientation de l'empreinte digitale d'entrée au point d'échantillonnage p_i.Nous définissons l'orientation

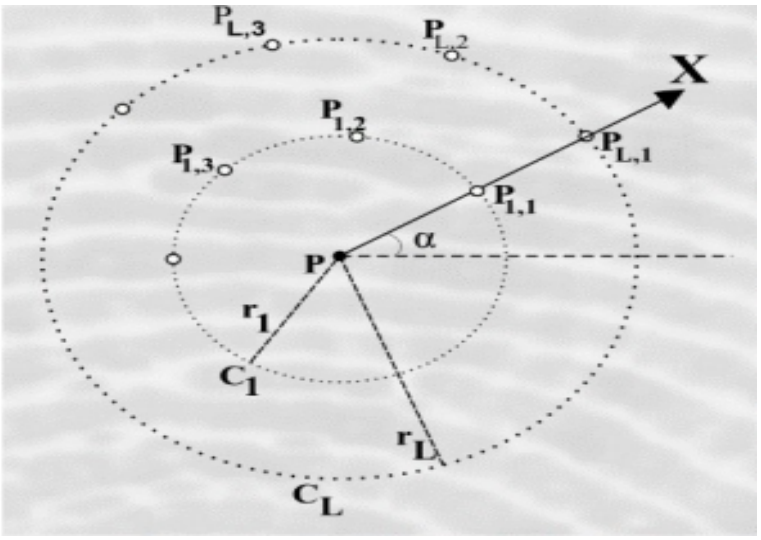


FIG. 2.2 : La structure d'échantillonnage utilisée pour extraire l'orientation et la fréquence des attributs

[9]

relative du point d'échantillonnage pi par rapport au point d'échantillonnage par rapport au point d'échantillonnage p_j noté $\theta_{i,j}$ comme suit :

$$\theta_{i,j} = \begin{cases} \theta_i - \theta_j - \pi & if(\theta_i - \theta_j) \geq \pi/2 \\ \theta_i - \theta_j + \pi & if(\theta_i - \theta_j) < -\pi/2 \\ \theta_i - \theta_j & Sinon \end{cases} \tag{2.6}$$

$\theta_{i,j}$ mesure le degré de concordance des orientations entre les points d'échantillonnage p_i et p_j.Il est invariant par rapport à la translation et à la rotation. Le descripteur d'orientation associé au minutieux m est défini sur tous les points d'échantillonnage les uns par rapport aux autres en prenant les valeurs absolues de leurs orientations relatives. Celles-ci définissent une matrice rectangulaire, appelée matrice d'orientation relative (ROM), dont la dimension est K.

$$ROM(m) = [(|\theta_{i,j}|)_{i=1}^{K}]_{j=1}^{K} \tag{2.7}$$

Notez que la ROM est une matrice symétrique dont la diagonale est nulle.

Extraction des attributs de fréquence :

La fréquence des crêtes locales en un point p(x, y) est le nombre de crêtes par unité de longueur le long d'un segment centré sur p, et orthogonal à l'orientation de la crête locale en p. Il s'agit d'une autre propriété intrinsèque qui caractérise une empreinte digitale. La Fig. 3.c montre un exemple d'image de fréquence d'empreinte digitale.

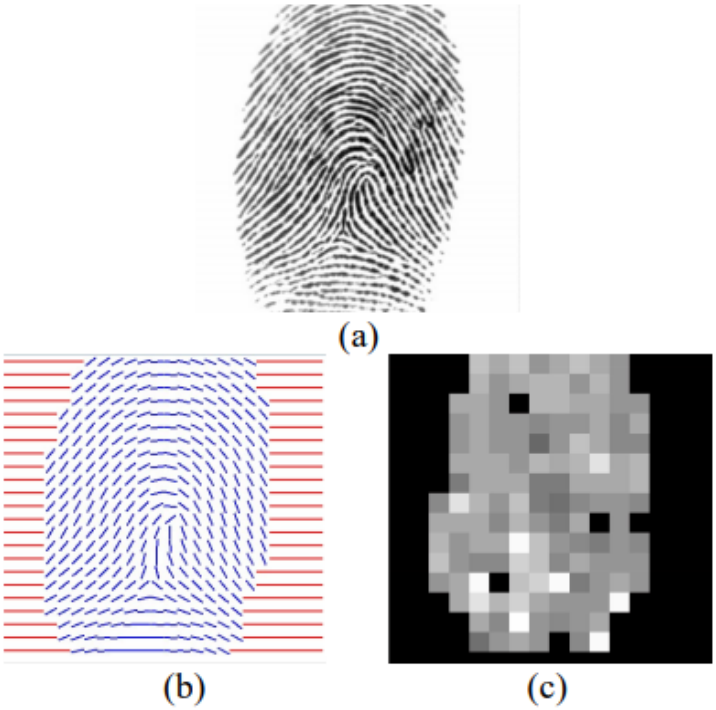


FIG. 2.3 : Une empreinte digitale (a) et son image de champ d'orientation (b) associée, et (c) l'image de fréquence

[9]

Soit f_i, (i =1..K), la fréquence de crête au point d'échantillonnage p_i. Nous définissons la fréquence de crête relative du point d'échantillonnage p_i par rapport au point d'échantillonnage p_j, appelée $F_{i,j}$, comme suit :

$$F_{i,j} = |f_i - f_j| \tag{2.8}$$

L'ensemble de toutes les fréquences d'arête relatives des points d'échantillonnage associés à un point caractéristique m détermine la matrice de fréquence relative (RFM) de dimensions $K * K$ qui constitue les attributs de fréquence du descripteur associé à la

minutie m.

$$RFM(m) = [(F_{i,j})_{i=1}^{K}]_{j=1}^{K} \tag{2.9}$$

Notez que la RFM est une matrice symétrique à diagonale nulle.

Descripteur de minutie combiné :

Le descripteur complet basé sur la texture associé à la minutie m, noté D(m), est la combinaison des attributs d'orientation et des attributs de fréquence.

$$D(m) = ROM(m), RFM(m) \tag{2.10}$$

Le descripteur combiné améliore la précision globale de la correspondance.

2.3.2 Création de modèles transformés :

Le modèle tel qu'il est défini par l'équation (2.10) est vulnérable car il peut facilement être attaqué. Une bonne idée pour protéger le modèle original est d'effectuer une réduction de dimensionnalité de ses attributs. Le fait que les attributs réduits soient des données déformées rend difficile, voire impossible, la récupération des attributs originaux. Cependant, la conception d'un tel schéma de réduction de la dimensionnalité est conditionné par la capacité des attributs réduits à réaliser une correspondance précise. Dans ce travail, nous proposons d'effectuer une décomposition en valeur singulière (SVD) sur les attributs d'orientation et de fréquence.

Récemment, la factorisation SVD a fait l'objet d'une grande attention dans la reconnaissance des formes.

La SVD permet de construire un modèle empirique, sansthéorie sous-jacente, d'autant plus précis qu'on y injecte de termes.

Soit D(m) un descripteur basé sur la texture associé à la minutie m. En effectuant une factorisation SVD des matrices d'orientation et de fréquence, on obtient :

$$ROM(m) = U_1 * \Sigma_1 * U_1^T \tag{2.11}$$

$$RFM(m) = U_2 * \Sigma_2 * U_2^T \tag{2.12}$$

Σ_1 et Σ_2 sont toutes deux des matrices diagonales constituées de valeurs propres positives ordonnées par ordre décroissant. Nous rangeons ces valeurs dans les vecteurs B1 et B2 comme suit :

$$B1 = diag(\Sigma_1) = (o_1, o_2, ..., o_K)^T \tag{2.13}$$

$$B1 = diag(\Sigma_2) = (f_1, f_2, ..., f_K)^T \tag{2.14}$$

Où $Oi_{i=1}^{K}$ et $Fi_{i=1}^{K}$ sont les valeurs absolues des valeurs propres de ROM et RFM respectivement.

Les vecteurs obtenus B1 et B2 constituent le modèle transformé (tD (m)) associé à la minutie m.

$$tD(m) = (B1, B2) = \begin{bmatrix} O_1 & F_1 \\ O_2 & F_2 \\ . & . \\ . & . \\ . & . \\ O_K & F_K \end{bmatrix} \tag{2.15}$$

Notez que le modèle transformé est indépendant de tout paramètre (clé) ; et ceci le rend résiliable.

La sous-section suivante explique comment rendre ce descripteur résiliable.

2.3.3 Création du modèle résiliable :

Pour rendre le modèle transformé obtenu résiliable capable de générer des pseudo-identités multiples à partir du même ensemble de minuties, nous faisons varier le modèle transformé avec une clé aléatoire spécifique à l'utilisateur, le PIN, pour obtenir le modèle transformé final.

Formellement, laissons le PIN être une clé aléatoire constituée d'une matrice rectangulaire de dimension K. Le modèle transformé résiliable final associé aux minuties m, fD(m),

est associé à la minutie m, fD(m), est obtenu de la manière suivante :

$$fD(m) = PIN * tD(m) = \begin{bmatrix} O_1 & F_1 \\ O_2 & F_2 \\ . & . \\ . & . \\ . & . \\ O_K & F_K \end{bmatrix} \tag{2.16}$$

2.3.4 Algorithme de Matching :

La tâche de comparaison d'empreintes digitales consiste à comparer deux empreintes digitales représentées par leurs descripteurs. Le résultat de la comparaison est un est un score d'appariement compris entre 0 et 1 qui indique le degré de similarité entre les deux empreintes. Soit m et m' deux minuties, leurs descripteurs résiliables respectifs sont les suivants :

$[Oi_{i=1}^{K}, Fi_{i=1}^{K}]$ and $[O'i_{i=1}^{K}, F'i_{i=1}^{K}]$.

Nous définissons la mesure de similarité S(m, m') entre m et m' comme suit :

$$S(m, m') = \alpha S_o(m, m') + (1 - \alpha) S_f(m, m') \tag{2.17}$$

Où S_o(m, m') et S_f(m, m') sont les mesures de similarité de la partie orientation et de la partie fréquence respectivement et $\alpha \in [0..1]$ est un facteur d'équilibre qui pondère la contribution de chaque partie.

$$S_o(m, m') = \pi_{i=1}^{K} exp(-|O_i - O'_i|/\mu_o) \tag{2.18}$$

$$S_f(m, m') = \pi_{i=1}^{K} exp(-|F_i - F'_i|/\mu_f) \tag{2.19}$$

μ_o et μ_f sont des facteurs de régularisation. La valeur de S(m,m') exprime le score de similarité obtenu appartenant à l'intervalle [0..1] dont 0 indique une totale inadéquation entre les deux minuties, et 1 indique une correspondance parfaite.

Étant donné deux empreintes digitales avec des minuties extraites $P1 = mi_{i=1..N1}$ et $P2 = m'j_{j=1..N2}$, chaque minutie des deux empreintes digitales est définie en termes de

modèle annulable. Pour établir la correspondance entre les deux empreintes digitales, nous devons comparer chaque point caractéristique de P1 avec chaque point caractéristique de P2. Il en résulte une matrice de similarité dont les éléments expriment les scores obtenus. Comme la correspondance entre les minuties est limitée par une relation de correspondance un-à-un ou un-à-zéro, la matrice de similarité doit être affinée pour éliminer les fausses paires de sorte que tous les éléments qui sont inférieurs au seuil prédéfini s sont mis à 0. En outre, si une minutie a plus d'une minutie correspondante, c'est-à-dire qu'il y a plus d'une valeur non nulle sur la même ligne ou colonne, nous ne gardons que celle qui a le score le plus élevé et nous mettons les autres à 0.

Par conséquent, nous définissons la similarité globale entre les deux empreintes digitales P1 et P2 comme suit :

$$S(P1, P2) = \frac{\Sigma_{i=1}^{N1} \Sigma_{j=1}^{N2} S(m_i, m_j)}{min(N_1, N_2)} \tag{2.20}$$

La valeur de S(P_1, P_2) est comprise entre 0 (les deux empreintes sont de doigt différent) à 1 (les deux empreintes sont des empreintes parfaites du même doigt).

2.4 Conclusion

La création d'un modèle biométrique résiliable est une solution adressée aux quelques attaques sur un système biométrique. Cette méthode nous permet de générer des pseudo-identités à partir du même ensemble de minuties, tout cela en variant le modèle transformé. Dans ce projet, nous avons suivis un algorithme pour réaliser cette transformation, et obtenir des résultats concluants.

Chapitre 3

Tests et résultats

3.1 Introduction

Dans ce chapitre, nous allons étudier l'implémentation et la validation de la technique présentée dans le chapitre précédent.

Nous allons d'abord voire les tests faits en appliquant la méthode étudiée, et les résultats obtenus. Nous allons présenter par la suite, les tests faits ainsi que la courbe obtenue.

Pour réaliser les tests, nous avons utilisés comme un environnement de programmation Matlab.R2015. Nous avons eu besoin d'un petit ensemble des images hôtes. Donc, Nous avons utilisés des empreintes digitales, relevées de la base FVC2002 [10], Les empreintes utilisées se divisent en 21 personnes avec 8 empreintes pour chacun. Nous avons extraits les minuties à l'aide d'un code source écrit en Matlab qu'on a appliqué sur les d'empreintes digitales relevées de la base elle-même. Les minuties extraites peuvent être sauvegardées dans un fichier avec l'extension '.m' et lus directement.

3.2 Bibliothèques tierces utilisées

3.2.1 Sc minutia

Nous avons utilisé le code Matlab « SSSC Minutiae »

https://www.mathworks.com/matlabcentral/fileexchange/29280-fingerprint-matching-al

pour extraire les attributs des empreintes digitales.

3.2.2 implémentation

1- Extractions des attributs :

En lançant le script extractdb, les minuties ainsi que leurs attributs d'orientation et de fréquences de toutes les empreintes seront extraites comme le montre la Fig.3.1

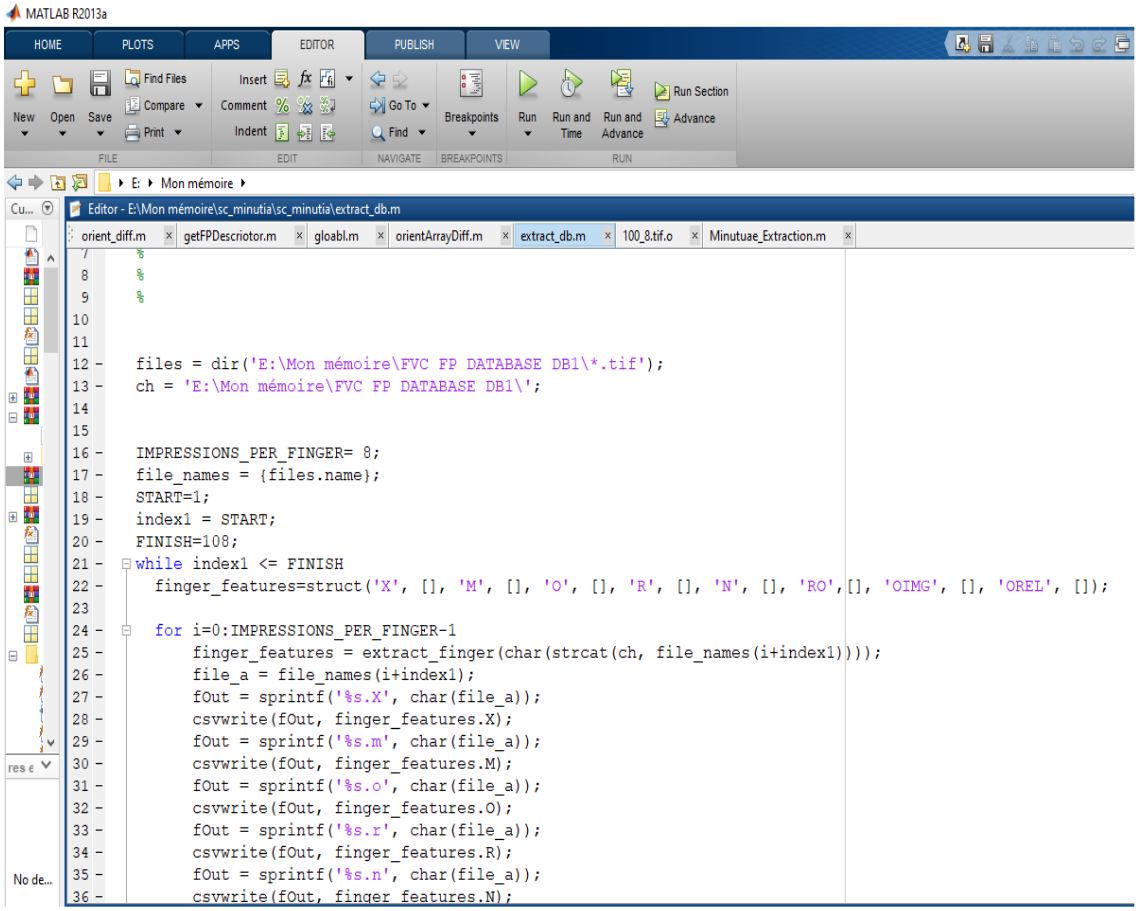


FIG. 3.1 : L'algorithme d'extraction des attributs

2- Chargement des attributs et création du descripteur biométrique :

Pour charger ces attributs pour une empreinte bien déterminée et lui créer un descripteur biométrique, nous avons écrit le script GetFPDescriptor qui aura comme entrées le numéro du doigt et le numéro d'impression et comme sortie le descripteur (template) d'orientation de chaque minutie dans l'empreinte comme le montre les figures Fig.3.2 et Fig.3.3.

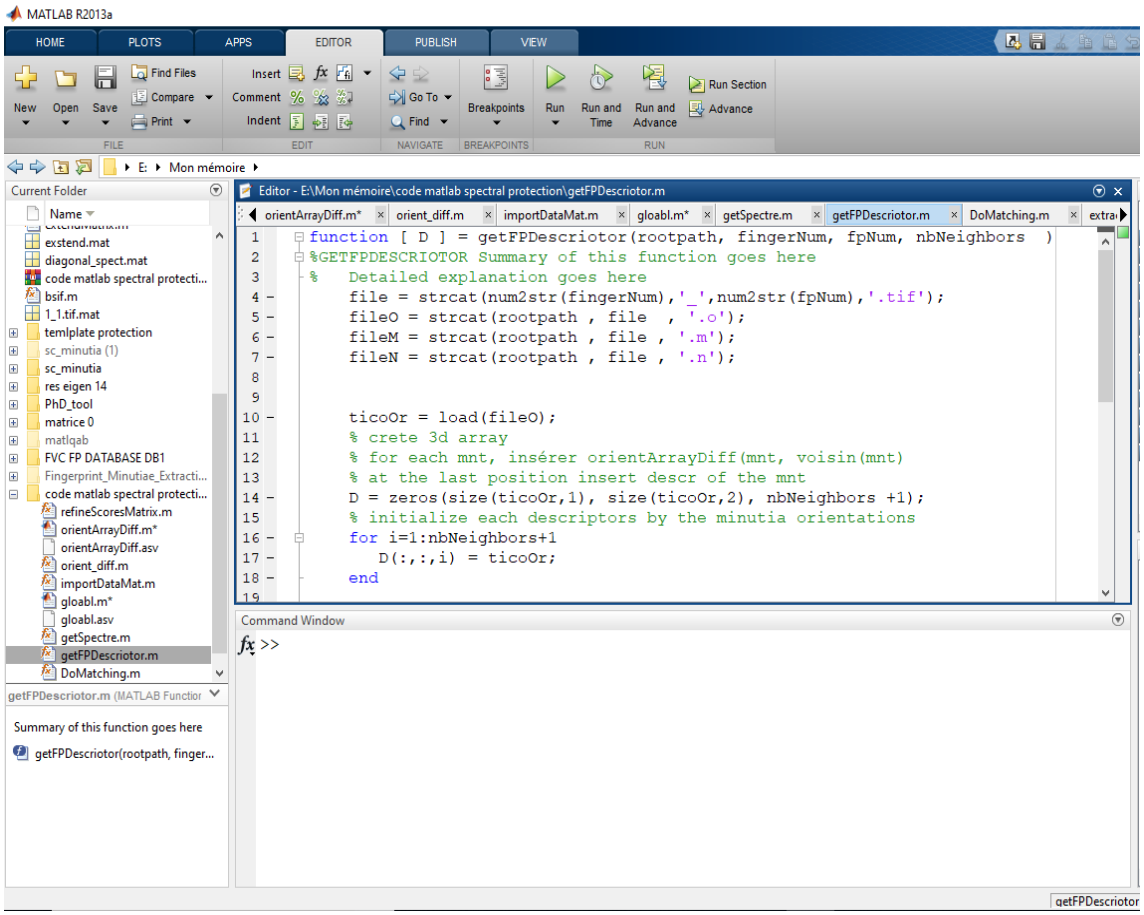


FIG. 3.2 : L'algorithme de chargement du descripteur d'empreinte

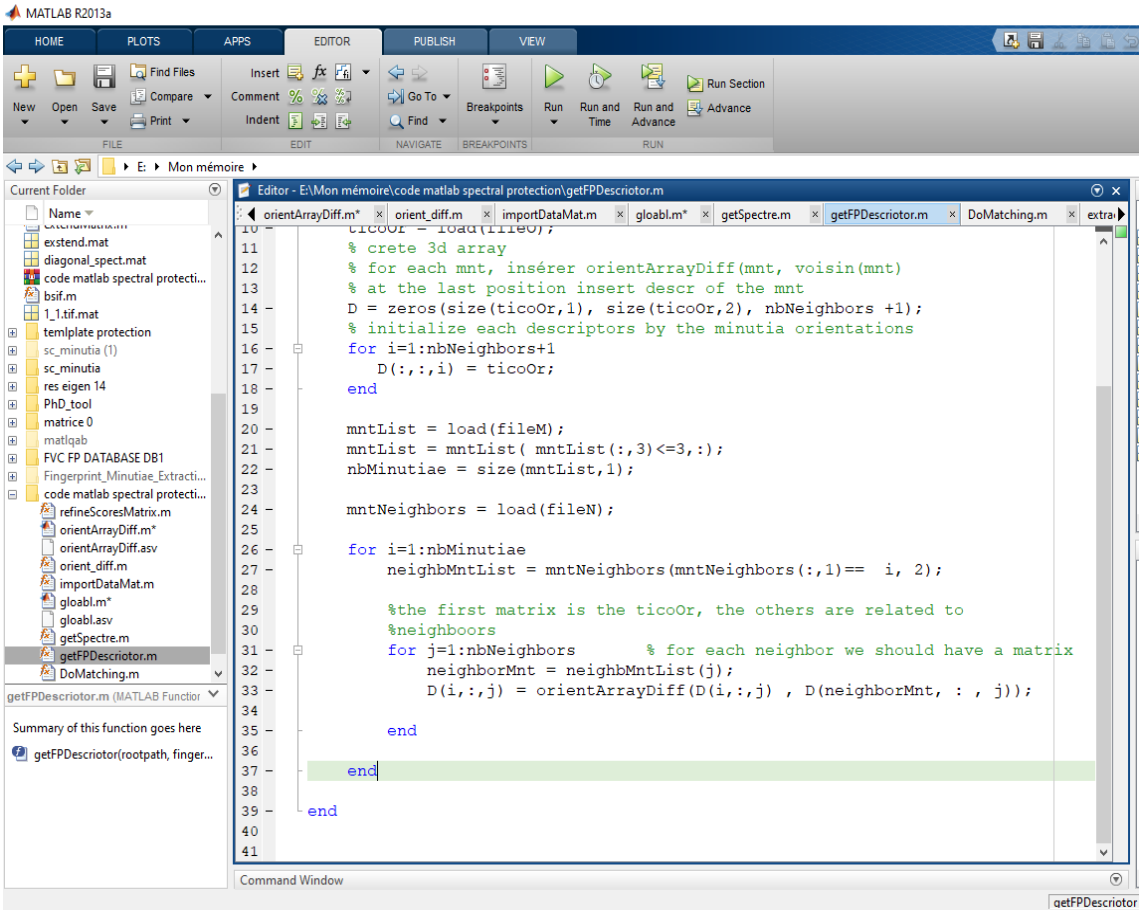


FIG. 3.3 : L'algorithme de chargement du descripteur d'empreinte

3- Création du descripteur biométrique résiliable :

Nous avons générer le template résiliables grâce à l'algorithme GetSpectre (Fig. 3.4) qui aura comme entrée le descripteur biométrique originale et comme sortie le descripteur résiliable. Cette transformation a été faite grâce à la fonction Matlab « SVD » qui permet de factoriser une matrice.

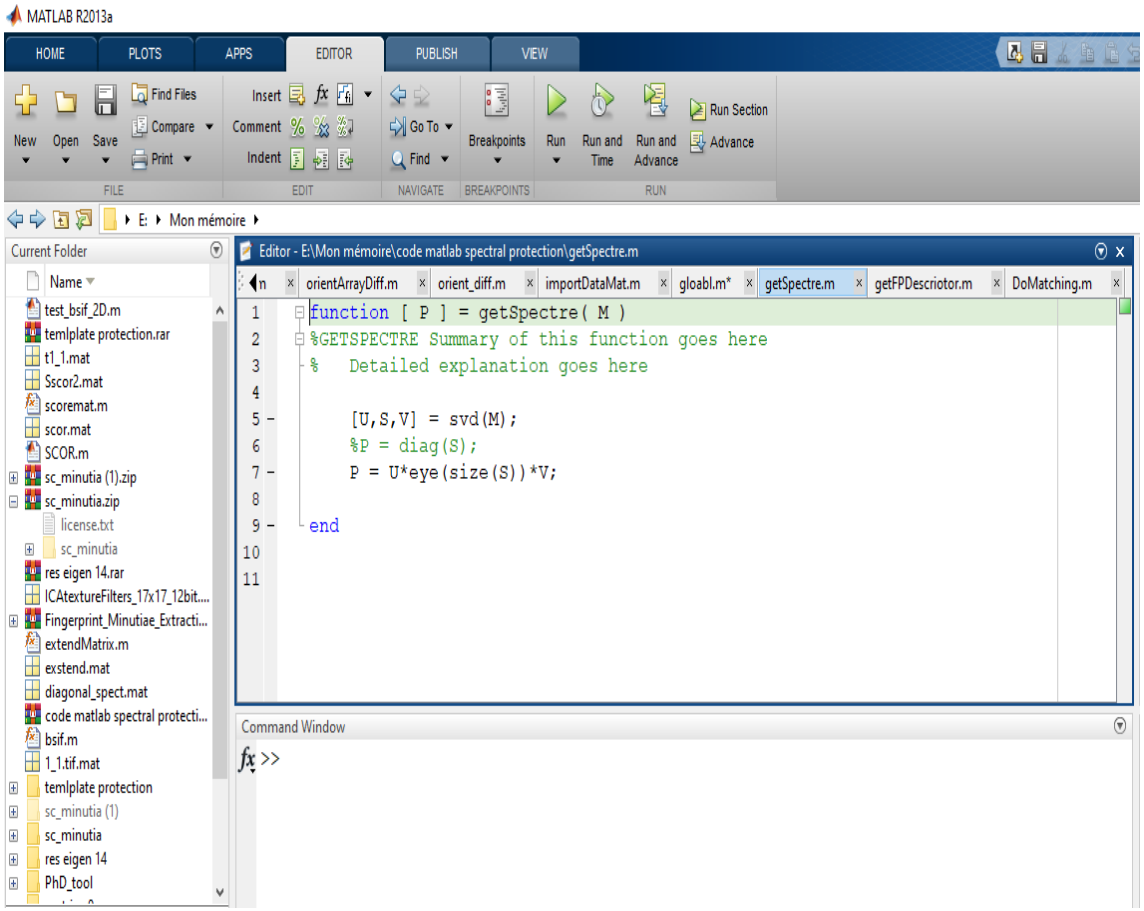


FIG. 3.4 : L'algorithme de génération du template résiliable

Matching :

Après avoir eu le template résiliable d'une empreinte digitale, on pourra entamera l'étape de matching. Cette dernière compare deux empreintes représentée par leur templates résiliables et établit un score de similarité comme le montre la Fig 3.5 et Fig3.6.

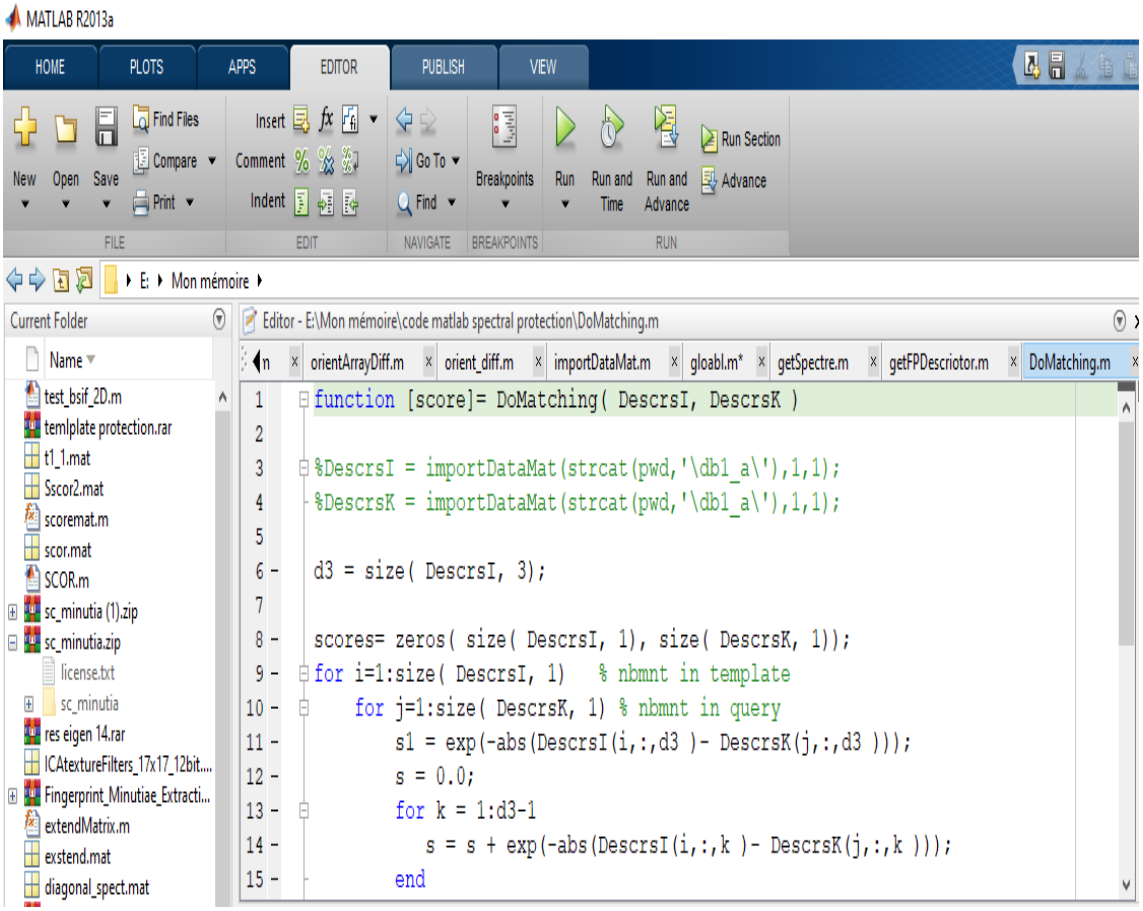


FIG. 3.5 : L'algorithme de matching entre deux descripteurs résiliables

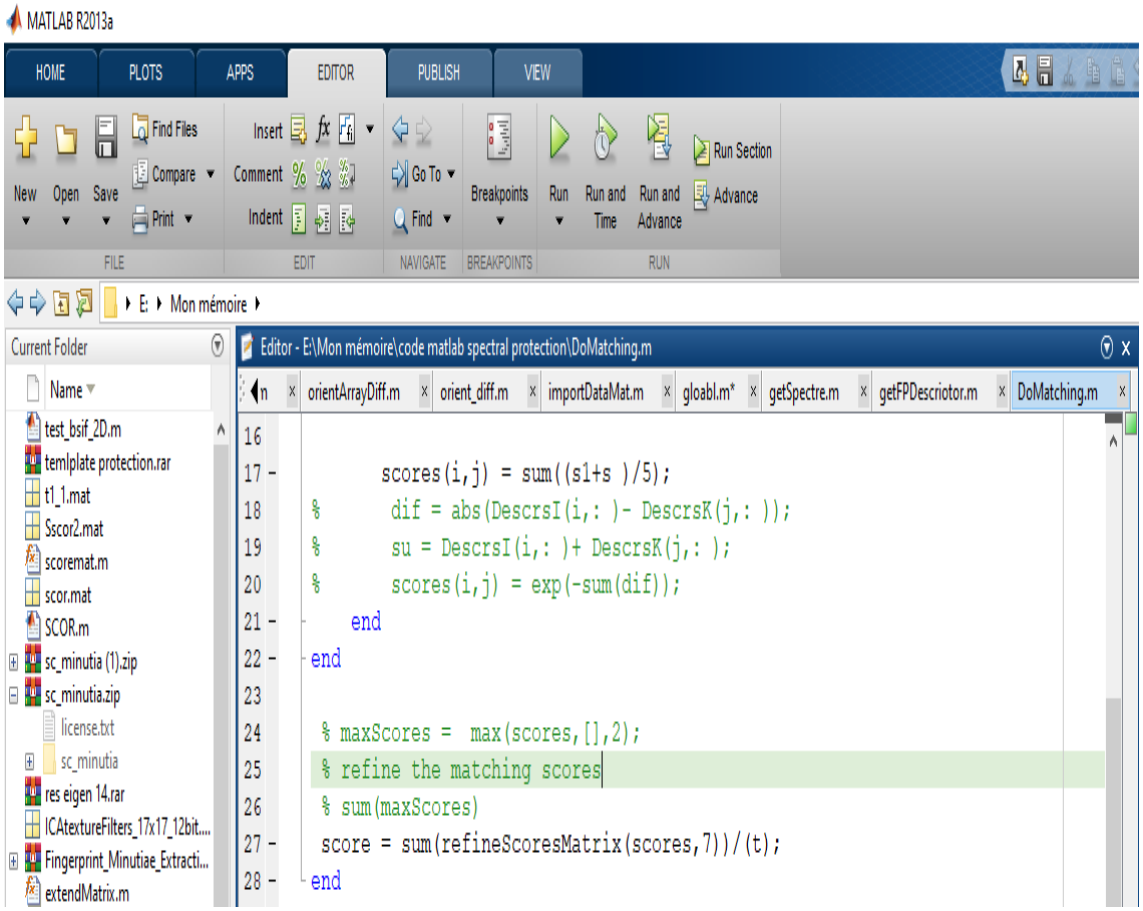


FIG. 3.6 : L'algorithme de matching entre deux descripteurs résiliables

3.2.3 Tests

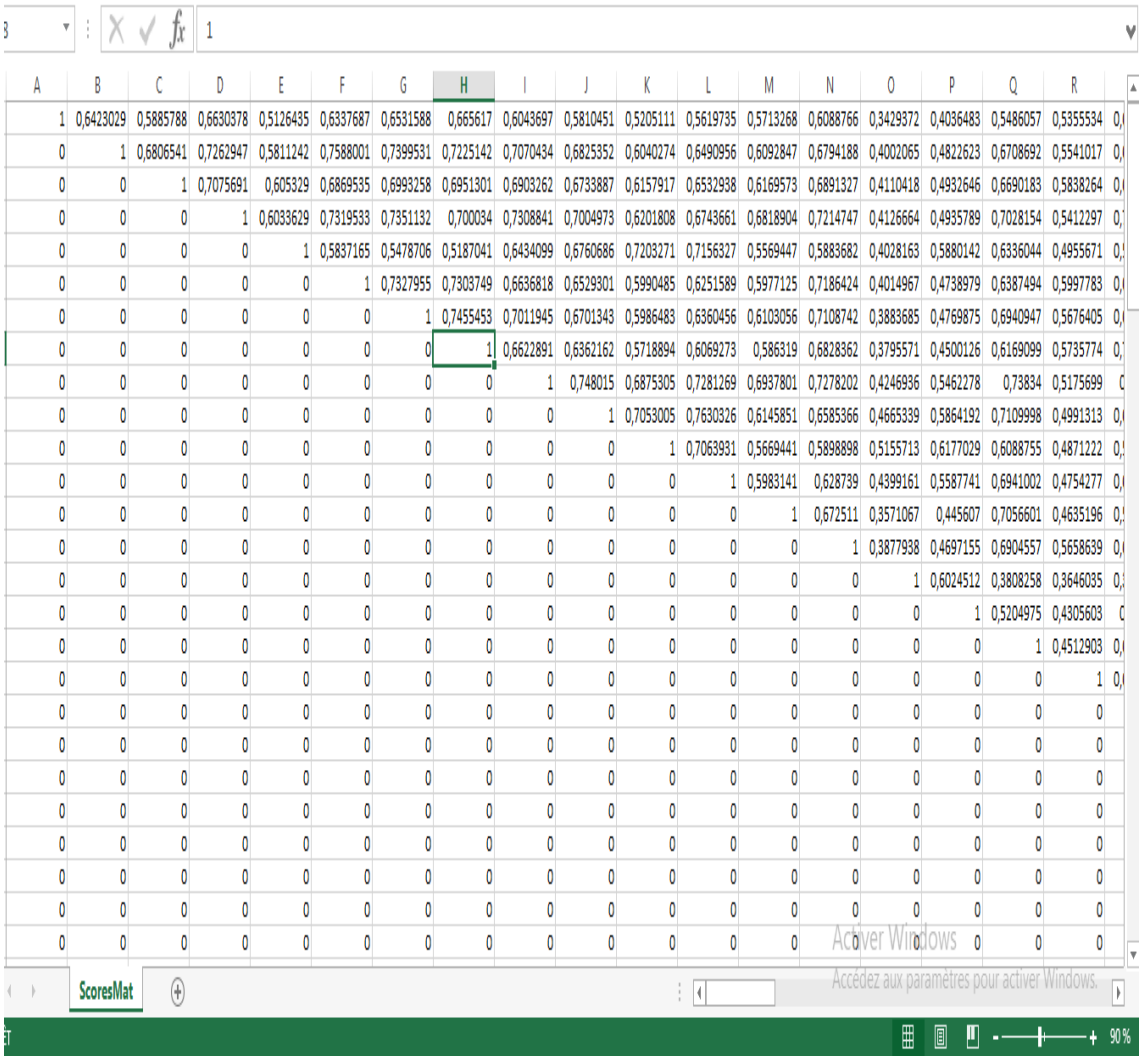
Test pour les personnes authentiques :

Ce test permet d'établir la distribution des scores des personnes authentiques. La procédure de test est la suivante : on compare chaque personne avec lui-même uniquement (chaque impression est comparée avec une autre de la même personne). Au fur et à mesure du matching, on enregistre les scores obtenus. Ces derniers seront triés pour construire une distribution des scores authentiques

Test pour les personnes imposteurs :

Ce test permet d'établir la distribution des scores des personnes imposteurs. La procédure de test est la suivante : on compare chaque personne avec les autres personnes (la première impression d'une personne est comparée avec la première impression de toutes les autres personnes). Au fur et à mesure du matching, on enregistre les scores obtenus. Ces derniers seront triés pour construire une distribution des scores imposteurs.

La figure 3.7 englobe quelques scores obtenus dans les deux cas.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	1	0,6423029	0,5885788	0,6630378	0,5126435	0,6337687	0,6531588	0,665617	0,6043697	0,5810451	0,5205111	0,5619735	0,5713268	0,6088766	0,3429372	0,4036483	0,5486057	0,5355534
2	0	1	0,6806541	0,7262947	0,5811242	0,7588001	0,7399531	0,7225142	0,7070434	0,6825352	0,6040274	0,6490956	0,6092847	0,6794188	0,4002065	0,4822623	0,6708692	0,5541017
3	0	0	1	0,7075691	0,605329	0,6869535	0,6993258	0,6951301	0,6903262	0,6733887	0,6157917	0,6532938	0,6169573	0,6891327	0,4110418	0,4932646	0,6690183	0,5838264
4	0	0	0	1	0,6033629	0,7319533	0,7351132	0,700034	0,7308841	0,7004973	0,6201808	0,6743661	0,6818904	0,7214747	0,4126664	0,4935789	0,7028154	0,5412297
5	0	0	0	0	1	0,5837165	0,5478706	0,5187041	0,6434099	0,6760686	0,7203271	0,7156327	0,5569447	0,5883682	0,4028163	0,5880142	0,6336044	0,4955671
6	0	0	0	0	0	1	0,7327955	0,7303749	0,6636818	0,6529301	0,5990485	0,6251589	0,5977125	0,7186424	0,4014967	0,4738979	0,6387494	0,5997783
7	0	0	0	0	0	0	1	0,7455453	0,7011945	0,6701343	0,5986483	0,6360456	0,6103056	0,7108742	0,3883685	0,4769875	0,6940947	0,5676405
8	0	0	0	0	0	0	0	1	0,6622891	0,6362162	0,5718894	0,6069273	0,586319	0,6828362	0,3795571	0,4500126	0,6169099	0,5735774
9	0	0	0	0	0	0	0	0	1	0,748015	0,6875305	0,7281269	0,6937801	0,7278202	0,4246936	0,5462278	0,73834	0,5175699
10	0	0	0	0	0	0	0	0	0	1	0,7053005	0,7630326	0,6145851	0,6585366	0,4665339	0,5864192	0,7109998	0,4991313
11	0	0	0	0	0	0	0	0	0	0	1	0,7063931	0,5669441	0,5898898	0,5155713	0,6177029	0,6088755	0,4871222
12	0	0	0	0	0	0	0	0	0	0	0	1	0,5983141	0,628739	0,4399161	0,5587741	0,6941002	0,4754277
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0,672511	0,3571067	0,445607	0,7056601	0,4635196
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0,3877938	0,4697155	0,6904557	0,5658639
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0,6024512	0,3808258	0,3646035
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0,5204975	0,4305603
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0,4512903
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

FIG. 3.7 : les scores du matching des empreintes digitales

Ce qui nous permet de dessiner notre courbe ROC comme illustrée par la Fig 3.8.

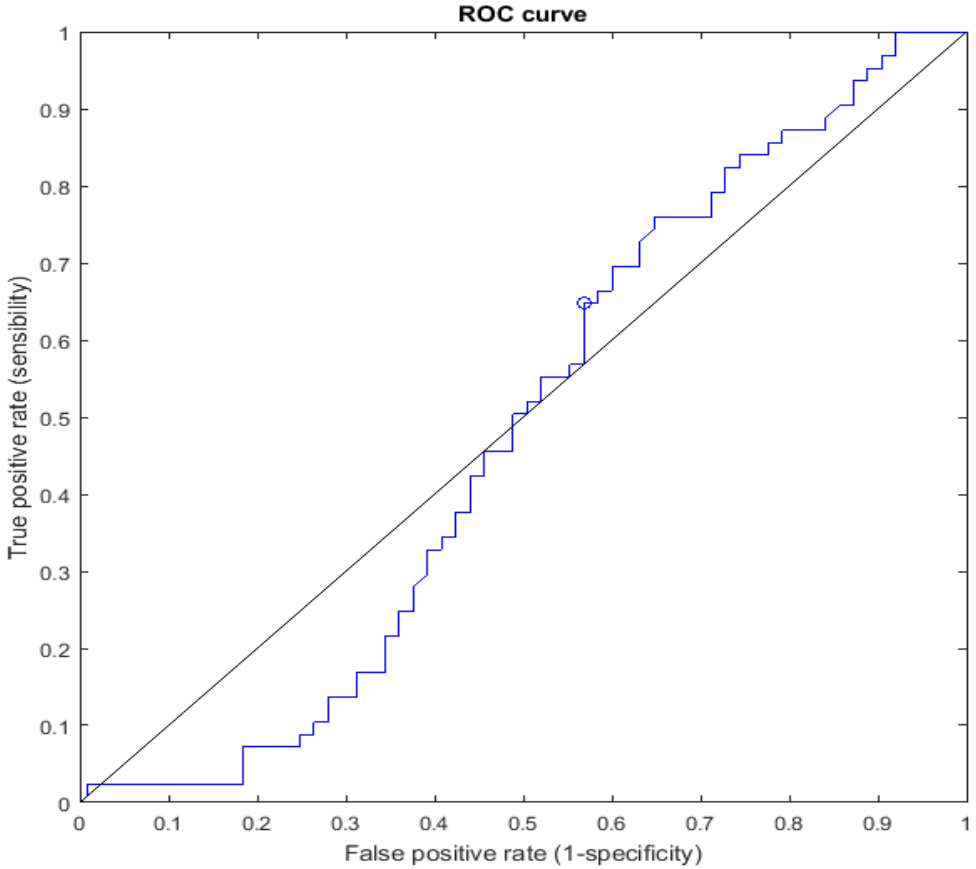


FIG. 3.8 : La courbe ROC.

Discussion des résultats

La courbe ROC obtenue montre des résultats acceptables mais ne sont pas performants. La réduction des attributs de chaque minutie à une valeur de 14, originalement étaient 176, a fait perdre beaucoup d'informations utiles durant le matching.

Conclusion et perspectives

Conclusion générale

La biométrie résiliable est une solution efficace au problème d'usurpation d'identité issu de la violation des données biométriques d'un individu.

Nous avons proposé un algorithme de protection de données biométriques basé sur la réduction des attributs de minuties. La réduction a été faite en appliquant la technique de la décomposition en valeur singulière.

Les contributions que notre projet a pu apporter peuvent se résumer dans les points suivants :

- Une analyse d'un système biométrique, et notamment la détection des multiples zones de vulnérabilités de ce dernier
- Un algorithme de transformation de modèle d'empreinte digitale en modèle résiliable.
- un algorithme de matching entre les modèles conçus et les modèles apportés.

En fait, c'était une occasion d'examiner l'exploitation d'un algorithme basé sur la biométrie résiliable pour améliorer la performance d'un système d'identification biométrique. Les résultats d'implémentation sont acceptables mais ne sont pas optimales.

Afin d'améliorer les résultats obtenus, nous proposons comme perspectives d'utiliser un plus grand nombre d'attributs, car nous avons utilisé quatorze minuties (14) au lieu de cent (176) disponibles.

Bibliographie

[1] J. D. Woodward Jr, "Nicholas M. Orlans Peter T. Higgins,"", *Biometrics", McGraw-Hill/Osborne*, 2003.

[2] J. R. Vacca, *Biometric technologies and verification systems*. Elsevier, 2007.

[3] M. El-Abed, "Évaluation de système biométrique", thèse de doct., Université de Caen, 2011.

[4] D. Guillerm. (). "Biometrics".

[5] P. Bonazza, "Système de sécurité biométrique multimodal par imagerie, dédié au contrôle d'accès", thèse de doct., Bourgogne Franche-Comté, 2019.

[6] A. S. Priya et R. Mukesh, "GA based Feature Selection for Multimodal Biometric Authentication",

[7] F. Belhadj, "Biometric system for identification and authentication", thèse de doct., Ecole nationale Supérieure en Informatique Alger, 2017.

[8] N. K. Ratha, J. H. Connell et R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems", *IBM systems Journal*, t. 40, n° 3, p. 614-634, 2001.

[9] P. Das, K. Karthik et B. C. Garai, "A robust alignment-free fingerprint hashing algorithm based on minimum distance graphs", *Pattern Recognition*, t. 45, n° 9, p. 3373-3388, 2012.

[10] (). "FVC 2002 Fingerprints databases".