Digital Fingerprint Quality Assessment

Zhigang Yao

UNIVERSITÉ de CAEN/BASSE-NORMANDIE

U.F.R. de Sciences

École doctorale S.I.M.E.M

Université de Caen Basse-Normandie

École doctorale SIMEM

Thèse de doctorat

présentée et soutenue le : 21/07/2015

par

Zhigang Yao

pour obtenir le

Doctorat de l'Université de Caen Basse-Normandie

Spécialité : Informatiques et Applications

Digital Fingerprint Quality Assessment

Directeur de thèse : Christophe Rosenberger

Co-directeur de thèse : Christophe Charrier

Jury

Jean-Luc Dugelay	Professeur à EURECOM	(Rapporteur)
Amine Nait-Ali	Professeur à l'Université Paris-Est Créteil	(Rapporteur)
Alan C. Bovik	Professor at the University of Texas at Austin (USA)	(Examiner)
William Puech	Professeur à l'Université de Montpellier	(Examinateur)
Jean-Marie Le Bars	Maître de conférences à l'Université de Caen Basse Normandie	(Examinateur)
Christophe Charrier	Maître de Conférences HDR à l'IUT Cherbourg Manche, antenne de Saint-Lô	(co-Directeur de thèse)
Christophe Rosenberger	Professeur à l'ENSICAEN	(Directeur de thèse)

For people who give me life and people who gave me a hand in life.

Acknowledgement

This thesis is finished by a full support from my two supervisors: Christophe Rosenberger and Christophe Charrier. These two scholars gave me an opportunity and a choice after I quit a PhD position at INSA de Rennes, where I failed to continue doing a thesis in Robotics because of a totally different background. Such an opportunity enables me to avoid the regret in pursuing a doctoral degree, and reaches my desire for having an academic career life. Both the failure and the opportunity are all significant changes in my life. Because of this, I have always been grateful to both scholars, and especially their patience and generosity in guiding me to enter the research field. Thanks to my supervisors at first.

Next, I want to thank Mr. Jean-Marie Le Bars, Dr. Mohamad El Abed at the Rafik Hariri University and Mr. Patrick Lacharme. Mr. Le Bars is my co-supervisor, and I am really glad to work with him. I can always feel his patience every time in revising the draft of my publications and always receive some pertinent comments from him, which are really helpful to my research. Dr. El-Abed is the senior PhD of our Lab and gave me a lot of opinions when I began working on my thesis. Mr. Lacharme helped me to prepare some documents required by the doctoral school in Caen, so that I can finish my PhD registration when I arrived in this new place.

It is a great honor that my PhD study and my life here in France are mainly supported by China Scholarship Council (CSC), where I obtained a scholarship for 42 months since later September 2011. Therefore, thanks to my country where I live, grew up and love with all my life. After my scholarship finished, my lab and the ENSICAEN provided me a funding of 3 months for my extension. Thus, thanks to my lab and the ENSICAEN, where I viewed as the second home in Caen. Rennes and Caen, the city I had lived there for one year and half a year since my arrival in France and the one I have been living since April 29, 2013. I have lots of memories in both cities, not only difficulties but also my sense of life. The familiar name of all these places will make me feel warm, and I am sure I will miss my life here in the future.

Once again, I don't know what I can say to my parents because everything they gave me is not measurable. The luckiest thing in my life is being a child of them.

Besides, I also want to thank the people who accepted my application of the first PhD position in France. Thanks very much for providing me an opportunity for an international study. Thanks very much for any help from them. They are Mrs. Marie Babel, Mr. Sylvain Guegan and Mrs. Muriel Pressigout.

Résumé

L'empreinte digitale est l'une des modalités les plus fiables en biométrie et donc a été largement étudié et déployés dans des applications réelles. La précision d'un système d'identification automatique d'empreintes digitales (AFIS) dépend largement de la qualité des échantillons d'empreintes digitales. La dégradation de la qualité d'empreinte digitales impacte le taux d'erreur lors de l'étape de vérification biométrique. Cette thèse se concentre principalement sur l'évaluation des mesures de qualité biométriques et plus précisément l'évaluation de la qualité des empreintes digitales (FQA), à partir d'une image en niveaux de gris et ou à partir de l'ensemble de minuties associées.

En faisant un examen à la fois raffinée des systèmes biométriques et des méthodes d'évaluation en préliminaire, cette thèse contribue tout d'abord par la proposition d'un nouveau cadre d'évaluation / de validation pour estimer la performance de métriques de qualité biométriques. Le cadre d'évaluation / validation est défini dans la phase d'enrôlement en utilisant des essais hors ligne. La validité d'une mesure de qualité biométrique peut être statistiquement mesurée par la dégradation du d'égale erreur (EER) et les intervalles de confiance (IC) associés.

Ensuite, cette thèse porte principalement sur l'évaluation de l'empreinte digitale de plusieurs façons, qui comprend trois parties dans le contexte de la FQA, où chacune d'entre elles est positionnée à partir d'une revue systématique de la littérature des études existantes. Tout d'abord, une approche d'évaluation de la qualité basée sur de multiples fonctionnalités et un avant-connaissance du rendement correspondant est proposé dans cette thèse, image d'empreinte digitale de qualification qui réalise avec des schémas de fusion et d'apprentissage et observe certains problèmes potentiels de ce type de solution. Deuxièmement, un nouvel algorithme FQA est proposé en utilisant uniquement modèle minuties image d'empreinte digitale de. Cette approche démontre la possibilité pour estimer la qualité d'image d'empreinte digitale avec le modèle de minuties seul. Troisièmement, un autre cadre FQA est réalisée via une approche d'élagage pixel de l'image d'empreinte digitale, ce qui donne une nouvelle solution de cette question de la. Pendant ce temps, toutes les approches FQA proposé dans cette thèse offrent une étude comparative de cette question, pour les algorithmes FQA proposées sont en mesure de représenter chaque solution représentant parmi les études existantes.

Summary

Digital fingerprint is one of the most reliable modality in modern biometrics and hence has been widely studied and deployed in real applications. The accuracy of one Automatic Fingerprint Identification System (AFIS) largely depends on the quality of fingerprint samples, as it has an important impact on the degradation of the matching (comparison) error rates. This thesis mainly focuses on the evaluation of biometric quality metrics and fingerprint quality assessment (FQA), particularly in estimating the quality of gray-level fingerprint images or represented by a minutiae set.

By making a refined review of both biometric systems and relevant evaluation techiniques, this thesis firstly contributes by the definition of a new evaluation/validation framework for estimating the performance of biometric quality metrics. The evaluation/validation framework is defined in the enrollment phase by using offline trials. The validity of a biometric quality metric can be statistically measured by the degradation of the global Equal Error Rates (EER) and the associated Confidence Intervals (CIs).

Next, this thesis makes effort mainly in assessing fingerprint image quality in several different ways which include three parts in the context of the FQA, where each of them is proposed in terms of a systematic literature review of the existing studies of this issue. First, a quality assessment approach based on multiple features and a prior-knowledge of matching performance is proposed in this thesis, which achieves qualifying fingerprint image with fusion and learning schemes and observes some potential problems of this kind of solution. Second, a new FQA approach is carried out via a pixel-pruning scheme of fingerprint image, which gives a new solution of this problem. Third, another FQA algorithm using the Delaunay triangulation is proposed to estimate the quality of a digital fingerprint via only its minutiae template. This approach demonstrates the possibility for estimating the quality of digital fingerprint with the minutiae template alone. Meanwhile, all the proposed FQA approaches in this thesis provide a comparative study of this issue, for the proposed FQA algorithms are able to represent each representative solution among the existing studies.

Contents

Introduction

Problem statement

Digital fingerprint, as one of the biometric modalities, has been the most widely studied object over the development of biometric technologies during the last few decades. Human society has long history in using fingerprint character, which denotes much experience that are considerable in biometric studies. Likewise, in contrast with the role that fingerprint played in later 1890s, its applications in biometrics nowadays is a modernization of what it used to be.

In the earlier years of the last century, the idea of using fingerprint was limited by scientific and technological status. Because of this, in that time, the application of fingerprint was mainly operated by manual approaches. Two cases can illustrate this problem, one is fingerprint for personal card files and another is collected by forensic experts: the former is generally obtained by rolling an inked fingertip surface across a contrasting background such as white card, and the latter is captured by several different means according to its situation such as chemical processing. The fingerprint on card file can be used for establishing a database of criminals, and those of forensic impressions can be used as an evidence to convict a criminal. However, while fingerprint experts processing the collected fingerprints, they have to face the difficulties raised by status of fingerprint impression, such as leftover pattern (fragmentary) and distorted impression [1]. This problem is referred to as fingerprint quality which can be described more simplistically as the clarity and visible region of the fingertip surface. More specifically, fingerprint quality involves several concrete situations, such as dry or wet cases which might be more easily happened for rolled impression, but it can be controlled by fingerprint experts while it is being captured.

Correspondingly, the quality problem exists as well in modern biometrics, and fingerprints either. The development of science and technology enables biometric characteristics to act somehow as traditional passwords, and to be used in a more widely range rather than the traditional forensic or official applications only. This variation could be represented by some civilian applications such as e-payment, smartphone, office building access, and so on. Without considering skin problems of individual subject, digital fingerprint is an image captured by specially made sensors. The associated quality problems are caused by the capture mode of the sensor in general. For instance, sensors of touching type capture fingerprints by putting a fingertip on the sensor (typically, a surface). This might results in wet fingerprint

or distorted impressions if the fingertip or sensor surface is a little humidity or the individual presses fingertip too much. These disadvantages would impact on the recognition operation of a biometric system. In this case, fingerprint quality has been introduced as an independent research topic just as the quality problem of other biometric modalities. The general purpose of fingerprint quality assessment is therefore to measure the captured fingerprint data and to achieve the goal of quality control. The quality control can be involved in several stages of a biometric system, which is presented in detail in this thesis. By considering this problem, the objective of fingerprint quality assessment is simply introduced next.

Objectives

As it is presented in the previous paragraph, fingerprint quality assessment is to obtain a measurement of sample quality. By doing so, it is able to reject the captured fingerprint data of poor quality and read the data one more time to ensure a good quality sample can be used for the biometric system (it is referred to as a fingerprint recognition system, in this thesis). This is the most direct purpose and the description of fingerprint quality assessment task. However, before diving into the details, a question can be asked that whether the system qualifies fingerprint quality with the same standard as human perception? Hence, several examples of some ink-rolled fingerprint and latent fingerprints are given to have a first look at the quality, depicted in Figure 0.1.

As it is shown in Figure 0.1, image 0.1(a) is a rolled fingerprint, and images 0.1(b) and 0.1(c) are latent impressions which need to be further processed by forensic experts. By having a glance at the image, it is easier to differentiate the visual information of these fingerprint impressions by human eyes and can give an estimation of their quality by brain. Image 0.1(a) is relatively clear among them undoubtedly, but how a fingerprint image quality can be measured by a biometric system. The objective of the thesis is to answer this question which involves in several aspects and is given in the following chapters in detail.

Organization of the Thesis

The content of this thesis is organized as:

Chapter 1 presents a simple introduction of biometrics which gives a general view of the research about biometrics and the essentials related to biometric sample quality assessment and its evaluation.

(a) Rolled fingerprint (b) Latent image (c) Latent image

Figure 0.1 – Fingerprint examples. 0.1(a) Rolled fingerprint, 0.1(b) and 0.1(c) Latent fingerprint impressions (collected from internet).

Chapter 2 presents an overview of the fingerprint modality with an introduction to digital fingerprint and a state-of-the-art of related studies for fingerprint quality assessment.

Chapter 3 mainly focuses on quality metric evaluation based on improving the overall performance. This chapter includes a review of most of the existing evaluation approaches, and a new validation framework is proposed.

Chapter 4 details a preliminary study of fingerprint quality assessment which is implemented by using a linear combination of multiple features. With a reference quality metric, this chapter also makes a comparative study of fingerprint metrics based on multiple features and learning approaches.

Chapter 5 defines a new quality framework of fingerprint with multiple segmentation approaches. This framework could be regarded as a fusion of multi-feature in segmentation phase which avoids of dealing with coefficients problem.

Chapter 6 proposes another new quality metric of fingerprint image by using minutiae template only. The proposed quality metric could also be viewed as an effective feature and reveals the possibility in estimating quality with no image information.

Chapter 1

Overview of Biometrics

This chapter presents a quick overview of biometrics in several aspects: a short introduction of the development of biometrics, a general view of biometric system including its modules, operation mode and performance evaluation.

Contents

1.1 Introduction

Biometrics is a scheme to perform user authentication by using one or several human characteristics for personal recognition instead of PIN codes [1]. The use of this kind of characteristics in human society can be tracked back to ancient time as early as in 6,000 BC, such as Qin and Han dynasties in China, Babylonia and Assyria time [2, 3]. The historical record also shows that people in ancient time have used handprints as evidence during burglary investigations. In a non-sophisticated way, comparing with the forensic applications of biometric data in 19^{th} century and the early 20^{th} century, human experts in this field are gradually replaced by biometric systems in most cases. The deployment of biometrics is still increasing, which conforms to today's complex and universally used information technologies. For example, the Automated Fingerprint Identification System (AFIS) built in a smart phone has been adopted in the last few years. Many similar innovations could be found during the last few decades as the study of biometrics is strongly promoted by the development of information technology. Meanwhile, some problems created as well several branches in biometrics, such as the encryption [4] of the biometric

sample. This thesis chiefly pays attention to fingerprint modality, especially its quality assessment problem.

This chapter is organized as several items: In Section 1.2, a summary of biometrics introduces an outline of this issue such as the modality, and its development is provided. Second, Section 1.3 addresses a general framework of the biometric systems illustrating the mechanism of this kind of system. Next, issues related to biometric system evaluation are given in Section 1.4. Section 1.5 presents the objective of this thesis. Finally, a summary of this chapter is given in Section 1.6.

1.2 Biometrics

The definition "Biometrics" or "Biometry" has been used in the early 20^{th} century for issues of data analysis in biological sciences. Currently, biometrics or biometric authentication is a computer science-based technique which is mainly used for human identification, access control and other similar applications. By comparison with traditional recognition methods, this kind of application is implemented by using physiological or behavioral information of human being, such as their face characteristics, signature dynamics, traits and so on. However, not every biological characteristic can be utilized as a biometric data. The selection of a biometric data depends on specifications of its application. Some researchers has classified rules for choosing a biometric data [5], in which several necessary conditions were presented to determine if a biological characteristic can be used as a biometric sample. In biometrics area, those human characteristics that can be regarded as biometric samples are known as biometric modalities.

1.2.1 Modalities

In general, biometric samples can be divided into several categories in terms of their physical properties [6], including biological analysis, behavioral analysis and morphological analysis. The human characteristics used mainly for biological analysis include odor, blood, DNA and so on. This kind of biometric data might be sensitive to a person, for it might involves in privacy and susceptible issues. Data forms used for behavioral analysis can be keystroke dynamics, human voice, gait, signature dynamics, etc. However, this kind of biometric data might vary along with body conditions, and sometimes might be not very distinctive. On the other hand, morphological analysis includes modalities such as fingerprint, iris, palm-print, finger veins, face and so on. This kind of biometric data is more acceptable to people and

more convenient for collection.

Each biometric data has its own strength or weakness; some are convenient, effective, invariant and invulnerable, while some others are not. The choosing of biometric sample(s) for a system hence depends on application requirement which can be a system with only one kind of biometric sample (unimodal) or a system using multiple biometric data (multi-modal). In addition, system complexity might also be a factor that should be considered for choosing biometric modality. Some commonly used biometric characteristics are shown in Figure 1.1.

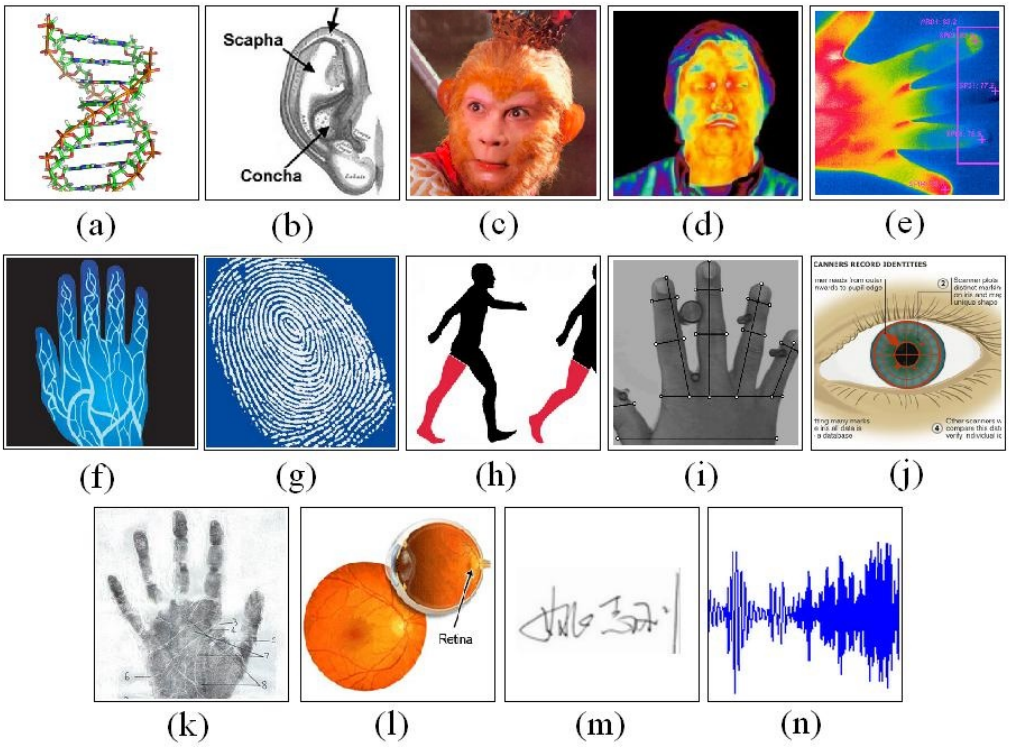


Figure 1.1 – Some representative biometric data forms: (a) DNA series, (b) ear, (c) face, (d) facial thermogram, (e) hand theomogram, (f) hand vein, (g) fingerprint, (h) gait, (i) hand geometry, (j) iris, (k) palmprint, (l) retina, (m) signature and (n) voice wave. (Image source: the internet.)

However, there is no ideal biometric modal that is able to satisfy all requirements as it has just been previously mentioned. According to existing studies in biometrics [7], fingerprint and face are commonly known as the most widely studied modalities [8] and iris comes second. Among these three biological characteristics, fingerprint is in a dominant position due to its advantages that are more in accordance with the standard of being a biometric modality. Generally, as it is mentioned before, a biometric sample should satisfy several requirements [8]:

- Universality: the biometric sample is available for all individuals.

- Distinctiveness: samples belong to two persons must be sufficiently different.

- Invariability: characteristics should be invariant or must be stable for a period of time.

- Acceptability: characteristics for which people should not show their reluctance, *i.e.* people would like to accept the use of characteristics.

In addition to those criteria for choosing biometric modalities, a characteristic should also be beneficial to system processing such as computation cost, storage consumption, etc. Furthermore, the impact of a biometric sample on recognition accuracy and the security consideration of the potential characteristics are also taken into account. These factors can be considered as the requirement related to a practical system.

1.2.2 Development and present situation

Biometric technologies have been developed over the last few decades. In early 90s, the use of biometrics can be found as a mystery high technology as they were planted in science-fiction movies or novels. However, currently, the application of biometrics has been widely deployed for commercial or civilian use rather than the traditionally official implementations only [5]. Deployments can be easily spotted in daily life, such as office building access control, electronic product authentication, E-payment which is being popular recently and so on. The official use, in addition to those forensic applications, has already been adopted for electronic identification recognition such as biometric passport and ID card. Up to now, most of countries around the world have began issuing the biometric passports [1] (Cf. Figure 1.2).

1. http://en.wikipedia.org/wiki/Biometric_passport

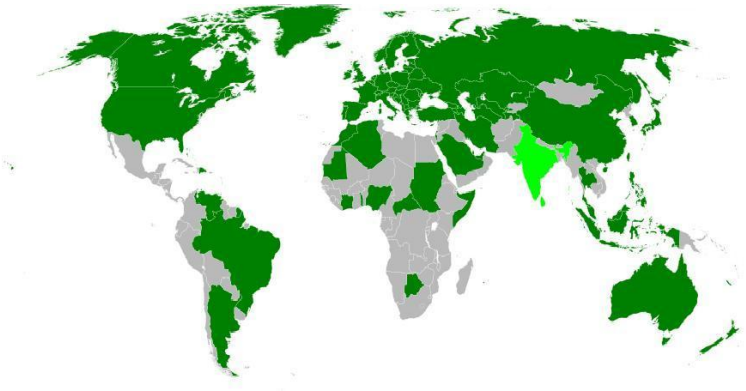


Figure 1.2 – Countries with biometric passport.

In Figure 1.2, highlight green is the area where has announced future availability of biometric passports, gray indicates area that biometric passport is not adopted and deep green represents the area where biometric passport is in use.

Currently, a typical symbol of biometric application is the biometric identification certificate which is usually planted with an electronic chip that contains one type of biometric data. This kind of documents are typically marked with the international symbol, as illustrated in Figure 1.3.

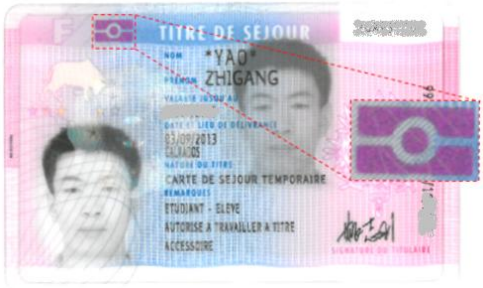


Figure 1.3 – Biometric symbol on one identification card.

Meanwhile, biometric sample or template protection [9] is a consequential problem which becomes one research branch of biometrics due to the potential risk raised by biometric attacks or fraudulent operations. In particular, face template and fingerprint template protection are widely studied among all the biometric modalities [10, 11].

On the other hand, in research field, multimodal biometrics and 3D biometric modalities [12] have already become a trend of biometrics study. Unimodal (also

known as monomodal in some publications) biometric system is the typical application and had been developed at first, but existing studies have found that biometric systems using a single biometric data is sometimes not sufficient to solve the problem due to its limitations [5]. In this case, researchers choose to study multi-modal biometric systems for improving system performance in multiple aspects [13]. Multimodal biometrics generally involves in 3 levels of fusion, including feature level, matching score level and decision level [14]. Fusion of feature level can be carried out either by the biometric data or by the features extracted from the sample. Matching score level fusion is a combination of matching scores obtained by different matching algorithms or those of different modalities. Decision level fusion is a scheme combining the outputs of different classifiers in terms of specified criteria, such as classifier selection [15]. The fusion at an early stage is generally considered to be more effective, for biometric data are more informative than those processed by later stages.

The recognition rate of biometric samples in 3D are believed to be more accurate than 2D and are not sensitive to the variation of illumination and viewpoint [16]. Studies in 3D biometric modalities mainly focus on face, ear [17] and finger surface, among which face is the most investigated object thanks to the advantage of capturing features. In addition, contact-less and behavioral biometric modalities might be numerously explored in future researches, and cloud biometrics has also been presented by researchers as the next generation applications for large-scale biometric data and mobile devices [18, 19] .

1.3 Biometric systems

In comparison with traditional PIN-based authentication schemes, a biometric system is generally described as an authentication system in which one kind of the physiological or behavioral characteristics of an individual is used for recognition [20, 5]. A password could involves several problems such as remembering, potential risk of brute force attack or the use from non-actual user [21] when considering the security elements and disadvantages of traditional authentication system. Biometric samples or their templates, however, are not easy to be duplicated and cracked without sophisticated method. Intuitively, a biometric system achieves user authentication via a comparison between a captured biometric sample of one user and a previously stored template (or a cryptographic template) of one user which might be the same one or a different. In this case, a biometric system is basically composed of several processing sessions and involves two operation modes. The processing procedure or modules are generally divided into several parts, including data acquisition, data

processing (feature extraction), data matching and decision [3].

1.3.1 Common system structure

According to the description of a biometric system, the structure of one system can be composed by several modules [22]:

- ***Capture module*** is to read raw data of human characteristics with a sensing device.

- ***Processing module*** performs pre-processing to the acquired raw data and feature extraction for coming operations. In this session, the system should evaluate the quality of an acquired data, and therefore determine whether the captured biometric sample would be rejected or accepted. The accepted data might be used as either an enrollment or an authentication sample matching with an existing template. The extracted features for an enrollment sample can be saved into a central or local database, which would depend on the application. In this case, enrollment might involves in two sessions, one is quality assessment and another is feature extraction. For example, a qualified fingerprint image of one person would be enrolled into the system as a reference when the user first time accesses the system.

  In fact, the processing of biometric characteristic depends on signal type applied in the system. In most cases of fingerprint, biometric data or human characteristics are recorded as images, but it will be different as the signal type changed. For instance, human voice is recorded as a waveform, while the characteristic of a signature recognition system is recorded as time series data. Therefore, while applying them into the system, extracted features of the recorded data will be finally stored in the database rather than the original recorded characteristics. This transformation facilitates data storage and matching operation.

  Enrolled as reference

  As mentioned above, enrollment is one of the followed operations which occurs when the user's biometric data is captured for the first time or a better sample of one registered individual is obtained. The first accessing to the biometric system leads to an once-again capture of user's characteristic until a qualified sample could be registered as an enrollment. For the second case, a user provides the system with a biometric sample for authentication, during which the enrollment of this user could be

replaced by a newly captured sample when its quality is better than the existed enrollment. An enrollment operation can be demonstrated via a diagram, see Figure 1.4.

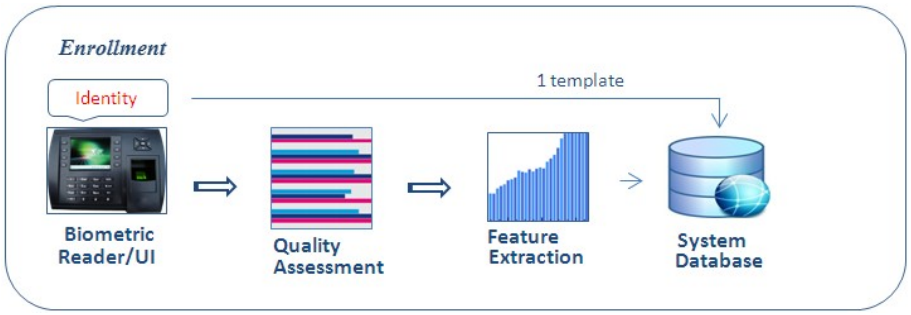


Figure 1.4 – Diagram of enrollment. This is a 4-step operation including reading (capture), quality control, feature extraction and saving.

- ***Matching module*** is to perform one-to-one or one-to-multiple comparison between authentication sample and enrollment sample. This processing depends on the operation mode of biometric system which includes verification and identification. This module is responsible for producing a matching result, based on which the decision module can determine to accept the authentication sample or not.

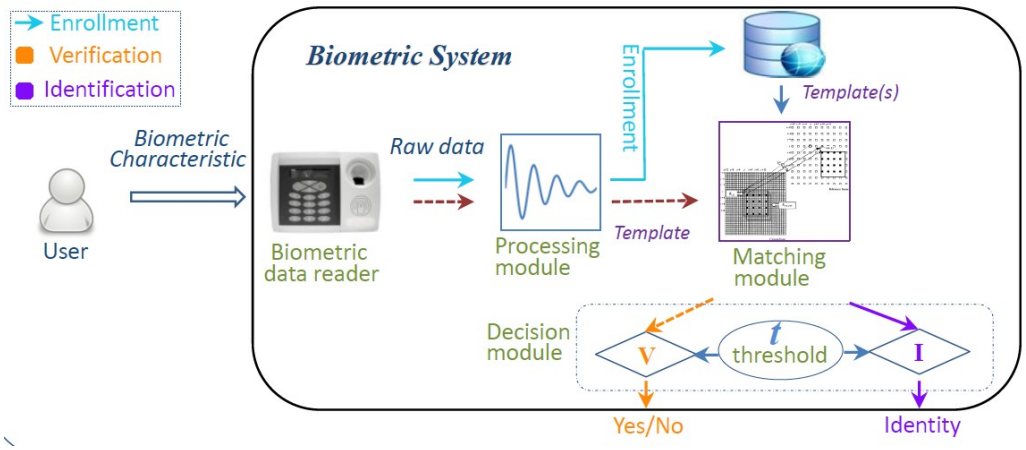


Figure 1.5 – System structure of a generic biometric system.

- ***Decision module*** makes a final decision to an authentication sample indicating whether it is a genuine sample or an impostor. This decision is usually performed by applying a threshold policy onto the matching result. For instance, matching results of some authentication samples are equal to or higher than a

threshold are considered to be matched samples. Figure 1.5 demonstrates a generic structure of a biometric system.

1.3.2 System mode

As mentioned above, a biometric system involves in different operation modes: 1) verification and 2) identification.

- **Verification**
  In this mode, a one-to-one matching is performed between an enrollment and an authentication sample corresponding to an identity claimed by a user. This one-to-one comparison determines the claim of the user is true or not [5], i.e. the user is who he (she) claims to be or not. It is also classified as an operation using for positive recognition which is to prohibit multiple users from using the same identity. Mathematically, the verification operation can be described as:

$$(I, X_Q) \in \begin{cases} w_1 & if \quad S(X_Q, X_I) \geq t, \\ w_2 & otherwise, \end{cases}$$

  where X_Q is an input feature vector extracted from the biometric data, I is the identity claimed by the user, S is the function for evaluating the similarity between feature vector X_Q and X_I, X_I is then the template corresponding to identity I, w_1 and w_2 indicate that the user's claim is true or not, respectively. Without losing generality, Figure 1.6 illustrates a diagram of verification mode.

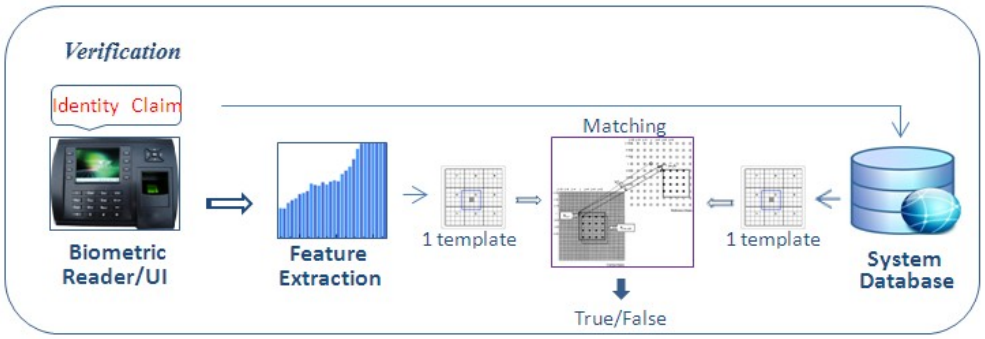


Figure 1.6 – Verification mode demonstration.

- **Identification**
  In identification mode, the user is recognized without the claim of identity. The system searches reference database to determine whether the captured

biometric characteristic matches one in the database. This leads to a one-to-many comparison which can figure out whether a user had already registered in the system or not. This mode is said to be critical to negative recognition applications. The aim of this mode is to prevent a user form using multiple identities. Similarly, identification mode can be formulated as:

$$X_Q \in \begin{cases} I_k & if \quad \max\limits_k \{S(X_Q, X_{I_k})\} \geq t,\ k = 1,\ 2,\ \ldots, N, \\ I_{N+1} & otherwise, \end{cases}$$

where I_k indicates k^{th} identity of totally N identities in the database, X_{I_k} is the template associated with I_k, and I_{N+1} is the reject case like an array bound. The identification mode is depicted by a diagram in Figure 1.7.

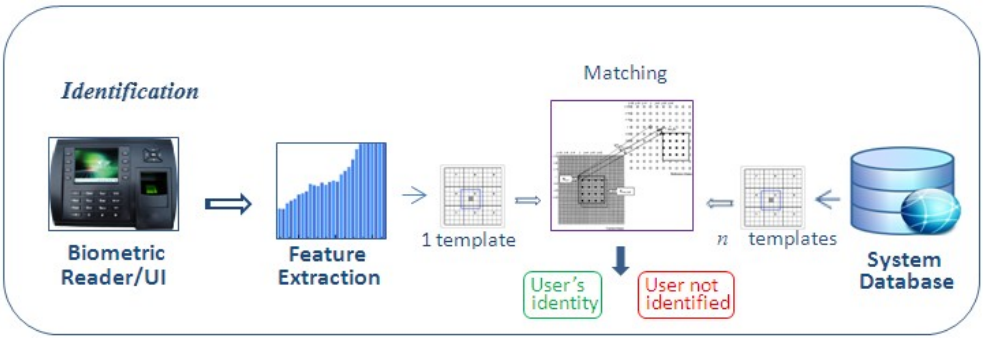


Figure 1.7 – Identification mode demonstration.

In this thesis, these two modes are not strictly emphasized due to the study mainly concentrates on biometric sample quality assessment. However, the matching mode involved in quality metric validation task can be considered as a verification mode because the users are known for the comparisons, for which the details are given in later chapters.

1.3.3 Advantages and Disadvantages

The features of biometric systems in both advantages and disadvantages have been analyzed via a comparison between traditional recognition or authentication techniques and two groups of biometric applications [5, 21]. Literally, these features can be stated as:

Security is the most important consideration for such a system. In this case, biometric system is more secure than traditional recognition technologies because duplicating biometric characteristics is not as easy as cracking traditional encryption

approaches based on alphanumeric and other characters. Nevertheless, the uniqueness of biometric samples reduces the possibility of unrestricted use of one identity by multiple users.

Usability or **convenience** in using biomeric modalities indicates that biometric systems are more convenient than traditional methods, for it is not necessary to remember a specific password and change it frequently, or keeping a card in the pocket.

Economical concern is an initial drawback for every new technology, on which biometric systems is not as good as traditional system but it is already not a problem for the cost of devices nowadays. Another shortage of biometric systems is technical problems, *i.e.* computation cost, especially storage requirement which is more considerable than traditional technologies. This kind of problem is still a limitation to those embedded applications such as the match-on-card (MOC) [23] systems of fingerprint.

1.4 Evaluation of biometric systems

The evaluation of a biometric system is important or necessary because the reliability of the system is crucial in real deployment. To measure the performance of a biometric system including its strength and weakness, some essential elements that affect or determine the behavior of the system should be figured out. Mohamad *et al.* [22] categorized factors of biometric system performance into 3 main aspects including usability, security and data quality.

First, usability is an aspect defined by ISO [24] which directly examines the performance of a human-machine interaction system including users' acceptance and satisfaction to a biometric system [25], efficiency and effectiveness of human-machine interactive operations. Second, security concern indicates the vulnerability of a biometric system when it is being used for specified applications, which could be threatened by potential attacks or deceiving operations [26]. Nevertheless, biometric data quality clearly impacts on the performance of a biometric system. A recorded biometric data with lower quality more probably leads to matching errors, especially false reject due to matching failure. Grother *et al.* [27] proposed biometric data quality is a predictor of biometric system performance. This property is known as the utility of biometric sample quality defined by ISO standard [28]. Studies about biometric sample quality assessment have proved that quality control approaches can greatly improve system performance [29].

In addition to these elements considered for evaluating the performance of a

biometric system, a well-designed evaluation approach is necessary to estimate system performance. Phillips *et al.* [30] categorized evaluation approaches to several cases including technology evaluation tests, scenario evaluations and operational evaluation. The protocol and database used for these types are different from each other. In this case, protocol and database became important for an evaluation approach. Cappelli *et al.* [31] presented that evaluation approaches based on self-collected databases are not comparable leading to meaningless results. In addition, both the evaluation type and protocol jointly determine how a biometric technology is quantified and how evaluation metric is defined.

1.4.1 Test scenario

In previous sections, elements relating to the performance of biometric systems are presented. In this section, the evaluation of a biometric system is discussed in two aspects which indicate how it is evaluated. The first one is the testing type of the biometric system which is introduced simply.

Among the three classes [30, 3, 32] of biometric evaluation tests, the technology test is viewed as the most general type in evaluating system performance, which could be done with either a sensing device or a large dataset (known as simulation). Mostly, to examine a completed software system or the prototype of an algorithm (which could be a component of a system), offline technology test is performed via simulation. In this case, given a specified protocol, the evaluation result is repeatable and comparable.

Scenario test is to determine the overall performance of the prototype of a system in a target environment, modeling a given application in the real world. Therefore, live operations are required when performing such a test. Generally, this is close to a real deployment of the tested system, and as a result, the purpose for verifying the maturation of a system is achieved.

The objective of an operational test is to measure the performance of a specified algorithm for a specific application, which is similar to the scenario evaluation. The main difference between them is the test scene, scenario test modeling the real world while operational test is performed at the real site and involved in the actual end users.

The study of this thesis pays attention to quality metric evaluation which is mostly done by using technology evaluation approaches. As a consequence, another requirement of this kind of evaluation is the statistic significance which quantitatively demonstrates the performance of the tested system [33]. Some commonly used metrics involved in estimating system performance are given in the following.

1.4.2 Evaluation metrics

The statistic measurements for indicating the performance are system errors and metrics presented by analyzing error types. As presented in Section 1.3, a biometric system makes the decision based on the generated matching score which indicates the similarity between the input biometric data and the template retrieved from system database. In addition, this decision is regulated by a specified threshold, t. The generated scores which are higher than or equal to t are decided as matched pairs, while scores lower than t are inferred as non-matched pairs.

Correspondingly, there are two distributions of each type of matching scores, one is genuine distribution generated from pairs of samples of the same person, and another is impostor distribution generated from pairs of samples of different users. As a result, there are two types of matching errors. One of them is known as false match indicating that two biometric samples from two different individuals are regarded as the samples from the same person, and another is false non-match meaning that biometric samples from one person are measured as two different samples from different users. The impostor matching error and genuine matching error are quantified by False Match Rate (FMR) and False Non-Match Rate (FNMR), respectively. As it is mentioned in the last paragraph, the threshold is one factor determining system decision. Obviously, both FMR and FNMR are also determined by this threshold, and they are functions of the threshold. In [5], these types of errors were concretely analyzed by considering both of verification mode and identification mode. The performance of a biometric system at all the operating points (thresholds) can be illustrated by a plot of FMR versus 1-FNMR which is called Receiver Operation Characteristic (ROC) curve. Figure 1.8 demonstrates these three metrics and the relationship between them.

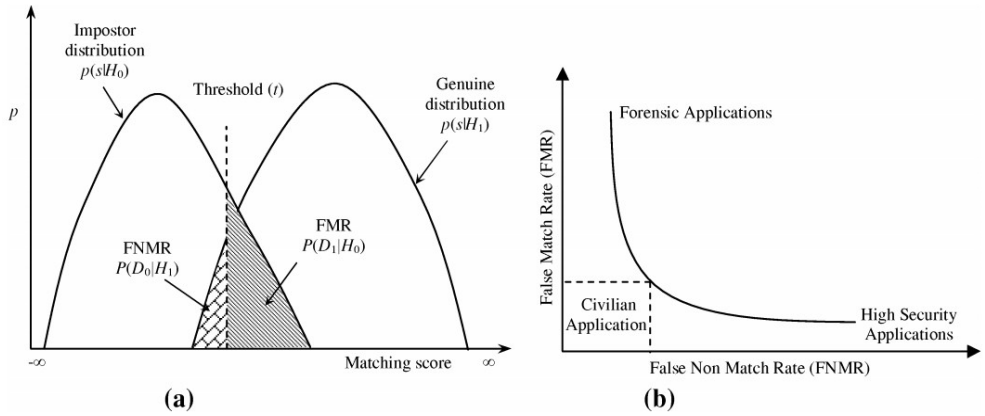


Figure 1.8 – (a) FMR and FNMR, (b) ROC curve. (Image source: the internet).

Mathematically, FMR and FNMR conform to the types of errors in hypothesis testing. These two metrics, by choosing verification task as an example, can be described as:

- Null hypothesis, H_0: input X_Q and its corresponding template X_I do not come from the same person;
- Alternative hypothesis, H_1: input X_Q and its corresponding template X_I belong to the same person.

And the corresponding decisions:

- D_0: person is not whom the user claimed;
- D_1: person is the one the user claimed.

Then, FMR belongs to type I error which means that H_0 is true but was rejected, and FNMR conforms to the type II error which represents that H_0 is not true but was accepted.

In addition to ROC curve, there is another error rate called Equal Error Rate or Crossover Error Rate (EER or CER), at which FMR and FNMR is equal to each other, as illustrated in Figure 1.9.

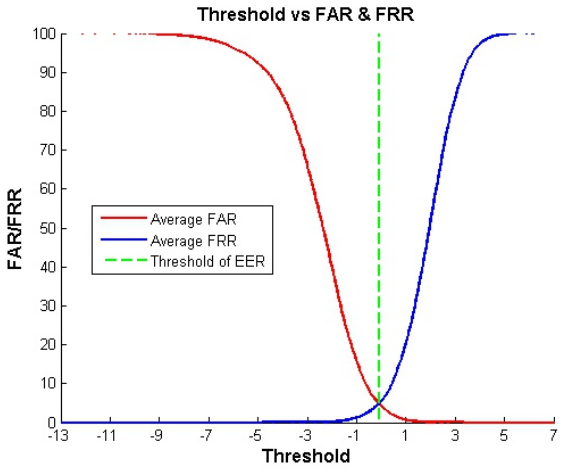


Figure 1.9 – The demonstration of the EER, FMR (sometimes known as FAR), and FNMR which is also denoted by FRR in some cases. (Image source: the internet.)

Another global measure is the Area Under Curve (AUC) value. The AUC is a quantitative index derived from ROC curve [34], which is defined in terms of the

Wilcoxon statistic when matching scores distributional assumption is involved, *i.e.*

$$W = \frac{1}{n_G \cdot n_I} \sum_{1}^{nG} \sum_{1}^{nI} S\left(x_G, x_I\right) \tag{1.1}$$

where $S\left(\cdot, \cdot\right)$ is a comparison rule

$$S\left(x_G, x_I\right) = \begin{cases} 1, \ x_G > x_I \\ \frac{1}{2}, \ x_G = x_I \left(discrete \ value \ only\right) \\ 0, \ x_G < x_I \end{cases} . \tag{1.2}$$

In this definition, x_G indicates the distribution of a set of genuine matching scores (GMS) of size n_G, and x_I denotes one of n_I impostor matching scores (IMS). The actual meaning of the AUC value is a probability of a correct ranking of a genuine and impostor pair. Consequently, the AUC value should to be much closer to 0 as the overall performance of the system getting much better.

Nevertheless, there are some other error rates, such as Failure to Enroll Rate (FER or FTE) and Failure to Capture (FTC).

- Failure to Acquire (FTA): Failure happened to sample capture session due to failures in detecting, identifying or acquiring a biometric sample of an adequate quality, failures related to user presentation, feature processing or quality control.

- Failure to Enroll (FTE): FTE occurs at enrollment phase caused by quality control or algorithm crash such as overtime or failure of feature extraction. It indicates the probability of failed enrollment attempts for either overall enrollment test or for a specific biometric instance.

- False Non-match Rate (FNMR): Sometimes this term is also defined as False Reject Rate (FRR) [35] which results in misleading. Generally, a FRR is a concept at system level which involves in FTAR and FTER. The FNMR is a probability indicating two samples of the same individual are incorrectly determined as samples of different individuals. In this study, it is defined as

$$FNMR\left(t\right) = \frac{card\{gms_{i,j,k} \mid gms_{i,j,k} < t\}}{card\{gms_{i,j,k} \mid gms_{i,j,k} < \infty\}} \tag{1.3}$$

- False Match Rate (FMR): This is an error corresponding to FNMR which is a proportion that an impostor sample was incorrectly matched to a genuine

sample. Likewise, it is often referred to as the False Accept Rate (FAR) in many cases. The FMR in this paper is defined as

$$FMR(t) = \frac{card\{ims_{i,j,l,k} \mid ims_{i,j,l,k} \geq t\}}{card\{ims_{i,j,l,k} \mid ims_{i,j,l,k} > -\infty\}} \tag{1.4}$$

In both equation 1.3 and 1.4, i and l are indices of the individual (user), while j and k represent indices of samples.

- False Accept Rate (FAR): A system level FAR is generally defined as

$$FAR = (1 - FTAR) \cdot FMR. \tag{1.5}$$

- False Reject Rate (FRR): A system level FRR is often defined as

$$FRR = FTAR + (1 - FTAR) \cdot FNMR. \tag{1.6}$$

  In this study, there is no ambiguous definition about FNMR, FMR, FRR, and FAR, for no FTAR is involved.

- Receiver Operation Characteristic (ROC) Curve: A plot of 1-FNMR against FMR.

With these evaluation metrics, one can deduce a numerical observation of the performance of a biometric system within a ruled circumstance simply as it has just been mentioned in Section 1.4.1.

1.5 Thesis Objective

As presented above, this thesis aims at considering the fingerprint modality, chiefly concentrating on its quality assessment and the evaluation of quality assessing approaches. The motivation is attributed to the widespread deployment of this modality. Although existing studies have found some solutions for this issues, challenges still exist in this field. For instance, the diversity of sensor settings makes difficulty in generating a commonly fitted feature or quality index and limited as well the performance of fusion-based approaches. On the other hand, some evaluation approaches focus only on the genuine error rates. In addition, the accuracy of some technology evaluation that relied on a global error rate are affected by impostor matchings, for the dimension of impostor matching are greatly larger than genuine matchings in general. In this case, this thesis makes effort in proposing new solutions for such a purpose.

1.6 Conclusion

This chapter makes a quick overview of biometrics in both the historical and technical manners. The general description of the biometric system is to illustrate how it is working and what the difference is between biometrics and traditional authentication techniques. The evaluation task estimates the performance of a biometric system, for none of biometric systems are perfect due to the specificity of biometric modalities. The evaluation is generally indicated by some statistical error rates proposed in the existing studies, which gives means for comparing the performance of different authentication systems. In addition to approaches based on error rates, there is also evaluation method relies on a statistic model [36] for predicting the authentication performance. Comparing with those error rates, the model requires some parameters and hence not adopted very often, especially in multi-vendor use. To this end, this chapter mainly provides a pre-requisite for our study in biometrics and the evaluation of quality metric.

In the next chapter, fingerprint and the existing studies for estimating the quality of a digital fingerprint are investigated.

Chapter 2

Fingerprint Modality

This chapter presents a brief overview of fingerprint modality including most of the relevant issues, and mainly focuses on fingerprint quality assessment (FQA). A literature review of quality assessment approaches is also explored in this chapter.

Contents

2.1 Introduction

As discussed in previous sections, fingerprint is one of the physiological biometric data which visually a pattern formed by parallel stretched lines and the slots between each pair of them. The lines are known as ridge lines and the slots are termed as valley. The ridge flow runs in parallel but smoothly changes its orientation and abruptly bifurcates or terminates somewhere hence forms some distinctive shapes or points which are known as features, such as core point and delta of the fingerprint, see Figure 2.1.

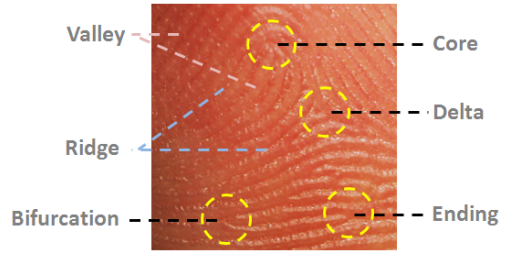


Figure 2.1 – Illustration of fingerprint pattern and minutiae features.

According to the record of archeology, the use of fingerprint in human society has been found on ceramics about six thousands years ago for marking the identity of the potter [37]. Fingerprint has also been used as a means of identity recognition during thousands years evolution of the human society. Computer aided fingerprint processing had occurred in the 1960s. Rudei *et al.* [38] had proposed to use the location of delta, core point and minutiae points in 2D coordinate system performing fingerprint classification. In the following decades, fingerprint has become a typical characteristic of modern biometrics and is known as the most popular modality for biometric systems. The reason that why fingerprint is widely adopted in biometrics could be attributed to several aspects, as addressed below:

1. Fingerprint characteristic is unique to each individual even for identical twins [39] and it is time invariant [40]. Studies shown that fingerprint of human has been formed with the first seven months during fetal development [5].

2. Fingerprint has obvious advantages in acceptability, which is known as a legal and mature biometric technique [41].

3. Most studies proved that fingerprint-based biometric systems perform well in accuracy. In FVC2006, the best accuracy was an Equal Error Rate (EER) of 2.15% in the open category, and 1.92% for a respectively light category [42]. In addition, it costs less than most of the biometric modalities in both external device development and storage requirement. The study of fingerprint occupied a large amount in biometrics field, which is beneficial for further research and development.

4. The physiological characteristic of human body provides more options for using fingerprint. Multiple fingerprint system, for instance, is also a means for ensuring the recognition accuracy and reliability [14].

This chapter mainly focuses on the use of fingerprint modality, especially the state-of-the-art in fingerprint image quality assessment studies. This chapter is organized as the follows. Section 2.2 gives a detailed description of the fingerprint modality and several short summaries with respect to fingerprint image processing and matching operations which provide a comprehensive view of fingerprint processing. Section 2.3 addresses a literature review of most of the existing studies in FQA, which details what the quality of fingerprint is and how it is estimated. At last, a summarization is presented to introduce the potential issue and existing difficulties.

2.2 Fingerprint processing

According to the literature mentioned above, the widespread deployment of fingerprint biometric systems could be attributed to the advantages of this modality. However, fingerprint recognition still cannot be considered as completely full-fledged solution because of some limitations of this characteristic such as intra-class variability and inter-class similarity [43]. The accuracy, computational cost and quality assessment of fingerprint system are still open issues. A fingerprint recognition system is generally viewed as a pattern recognition system which greatly depends on salient features extracted from fingerprint image and the correspondingly designed matching algorithm. Without considering the decision module, this section presents issues related to fingerprint processing.

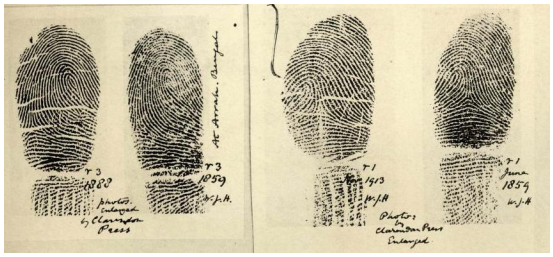
2.2.1 Acquisition

The acquisition of a fingerprint pattern is achieved by means of capture techniques, which could be divided into two categories: the first one is based on traditional inked fingerprint image known as off-line collection, and the other is on-line scanned image (also known as live-scan) digitized by special sensors [44].

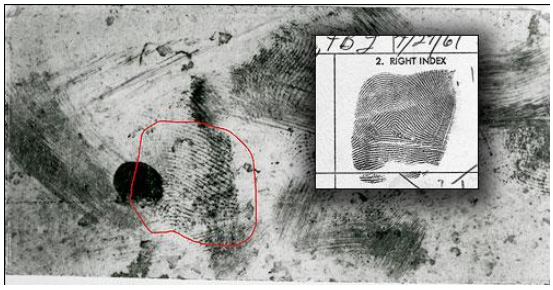
An off-line image can be viewed as a digitization of existed inked image or latent fingerprint observed at crime scenes. An inked image (Cf. 2.2(a)) [1] is typically an obtained impression by rolling a fingertip surface with dipped ink on a well contrast paper or card. This kind of impression is quality controllable but over-inked (wet) and dry patterns are easily created. The off-line collection of inked image could be either carried out by scanning the impression paper or taking high quality photo of the impression. The latent fingerprint (Cf. 2.2(b)) [2] is an impression left by an individual touching a surface with fingertip surface. The latent impression is impressed on the surface due to the human skin's oily property. This 'touch' mode leads to distortion of fingerprint impression and incomplete pattern of the print. In addition, the processing for obtaining the residual pattern needs to employ special means such as chemical techniques. Because of this, the digitized image is generally a bad quality sample due to small finger area and distorted ridge-valley pattern [45].

On-line scanned fingerprint images are captured by putting the fingertip surface onto live-scan device. The sensing device directly generates a digital image instead of performing a digitization process of inked fingerprint pattern. Correspondingly, sensor type determines image specification such as image size, resolution, contrast

1. http://en.wikipedia.org/wiki/Fingerprint
2. http://www.fbi.gov/news/stories/2011/october/print_101411



(a) Inked image



(b) Latent impression

Figure 2.2 – Examples of inked fingerprint image 2.2(a) and latent impression 2.2(b).

Figure 2.3 – Example of fingerprint images from the FVC databases (from the left to right): optical, capacitive and thermal sweeping. Resolutions are over 500-dpi.

and so on. A comprehensive introduction of fingerprint scanner in terms of sensing mechanism can be found in [43], among which three types are mainly used in current applications of fingerprint, *e.g.* optical, capacitive and thermal, where optical sensor is the most and commonly used type and the others are solid state sensors. Several examples formed by these sensors are shown in Figure 2.3.

In Figure 2.3, one can observe some differences among the 3 images, such as resolution and foreground area which indicates the valid area of fingerprint and is hence greatly related to fingerprint quality. However, one can not conclude that a specific type of sensor more probably generates bad quality images because the

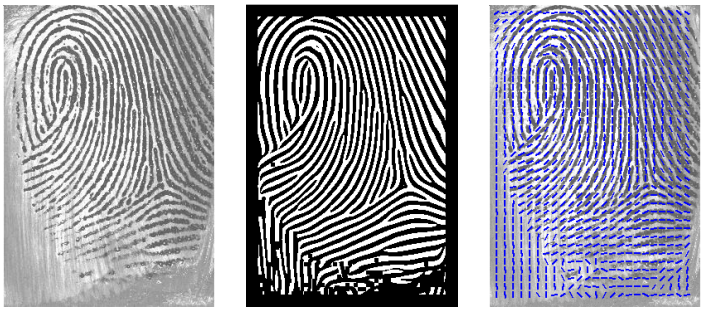
factors that constrain fingerprint quality are manifold. According to the literature [46], in addition to the sensor types introduced above, ultrasonic sensor is as well a well-known device, but it is relatively expensive in comparison with the others and hence rarely adopted. Nevertheless, mutli-resolution acquisition devices [47] and the built-in camera of smartphones are also studied to collect more informative fingerprint images [48].

2.2.2 Fingerprint pre-processing

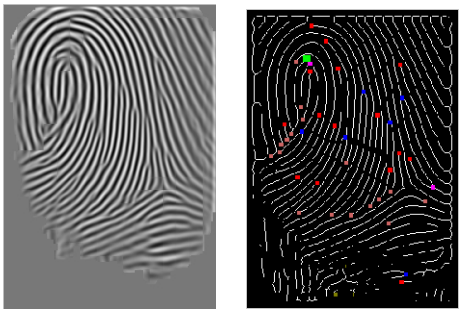
For a fingerprint biometric system, the raw data captured by the input sensor is a fingerprint image in general. But the captured signal is generally a gray scale image which is sometimes considered as an unstable means for fingerprint recognition [43]. In most cases, fingerprint recognition is carried out by using features extracted from the original image due to the consideration of accuracy, computation consumption and so on. In order to obtain effective and to achieve a good performance, the validity and reliability of features are crucial for a recognition system or matching algorithm. Because of this, several pre-processing operations are commonly performed for feature extraction to improve the usability of the features. The pre-processing of a fingerprint data generally involves in noise removing, enhancement, segmentation, binarization and so on. This part presents a brief introduction of some important pre-processing steps.

1. Enhancement of a fingerprint image is helpful to remove spurious characteristics of fingerprint and improve the clarity of ridge-valley pattern. An enhancement algorithm is generally carried out by applying several intermediate operations onto the captured fingerprint image [49]. The enhancement of fingerprint image could be performed on either a binary image (Figure 2.4(b)) or the original gray-scale image (Figure 2.4(a)).

   Binary image is obtained by using ridge extraction approach to the original gray level image. This representation of fingerprint is beneficial to eliminate spurious ridge configurations, but the effect depends on the performance of the extraction algorithm. Some algorithms also perform a noise removal step before going any further, such as Wiener filtering [50]. The algorithm proposed in [49] generates an enhanced image from the original fingerprint image by estimating the orientation (Figure 2.4(c)) and the frequency of ridge-valley pattern of fingerprint. Chikkerur *et al.* [51] proposed an enhancement approach based on Short Time Fourier Transform (STFT) analysis. Their approach



(a) Original image (b) Binary image (c) Orientation image



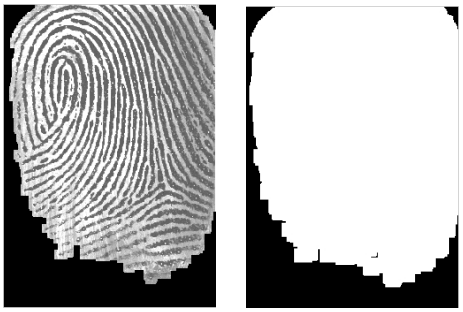
(d) Enhanced image (e) Skeleton image

Figure 2.4 – Examples of gray-scale fingerprint image 2.4(a) and its binary image 2.4(b), orientation field 2.4(c), enhanced image 2.4(d) and a skeleton image 2.4(e).

estimates block orientation via a marginal density function carried out by using the spectrum of Fourier Transform.

2. Segmentation (Figure 2.5) is generally an operation that separates the foreground area (ridge-valley pattern) from the background. This processing is able to reduce noise effect caused by fingerprint background, especially those on the border of ridge-valley pattern. As it is presented in the literature [52, 43], segmentation of fingerprint is also carried out by using features which differentiate foreground area from the background region. Because of this, some features are also used for assessing fingerprint quality, such as Gabor response [53] and statistic measures of image pixels [54].

Both two phase operations mentioned above are useful for feature extraction no matter what kind of matching mechanism adopted by the recognition system is. This advantage leads to the improvement of fingerprint image quality and hence indicates a difference between subjective measurement [55] of quality and the biometrical definition of the fingerprint quality.



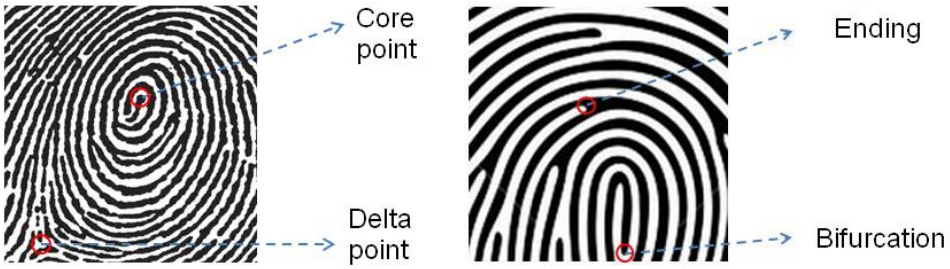
(a) Segmented image (b) Area contour

Figure 2.5 – Examples of fingerprint segmentation segmented image 2.5(a), and a contour illustration 2.5(b).

2.2.3 Feature Extraction

In section 2.1, a general view of several fingerprint features is given by Figure 2.1, in which the core point can be further classified as whorl and loop pattern of fingerprint, but it is generally described as "the top most point of the innermost ridge line" [56]. This kind of feature is known as level-1 feature and is widely used for matching operation due to its distinguishing property, *e.g.* correlation-based matching algorithm uses core point as the registration point [57], and some minutiae template alignment algorithms also rely on singularity points [58].

In addition, bifurcation and ending are known as minutiae points of fingerprint and are categorized as level-2 features, see Figure 2.6(b).

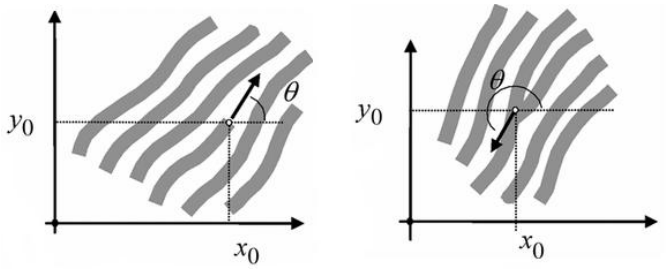


(a) Level 1 (b) Level 2

Figure 2.6 – Example of fingerprint features.

According to the literature [59], level-2 features are viewed as local features which involve over 150 different types of characteristics of local ridges. For instance, in addition to the marked points in Figure 2.6, some special characteristics are also categorized as minutiae, such as island or dot which are generally looked as a small short ridge.

However, in most cases, fingerprint recognition system based on minutiae template mainly use ridge ending and ridge bifurcation. These two kinds of ridge characteristics are said to be sufficient to determine the uniqueness of a fingerprint, and hence are useful for matching operations. Moreover, these features are easy to be captured and relatively stable. They also occupy the largest amount of fingerprint characteristics. Therefore, minutiae extraction algorithm mainly focuses on detecting these two kinds of points. A detected minutia point is generally represented by at least two components, location (x, y) and orientation θ. Nevertheless, a specific minutiae extractor also provides some additional information of minutiae, such as type and quality score. This is basically dependent on the requirements of the application, such as the ISO compact card format which records location, orientation and type of the minutiae [60]. Figure 2.7 illustrates a representation of minutia point.



(a) Ending (b) Bifurcation

Figure 2.7 – Example of minutia's representation. Figure 2.7(a) is the angle of ending and Figure 2.7(b) corresponds to bifurcation. (Image source: the internet.)

2.2.4 Matching

Matching (or comparison [61]) module compares a template acquired from a processed raw data with the reference template picked up from the system database. In most cases, data for matching operation are two templates acquired by feature extraction. Thus, the system determines whether two fingerprints are matched in terms of the similarity between their templates.

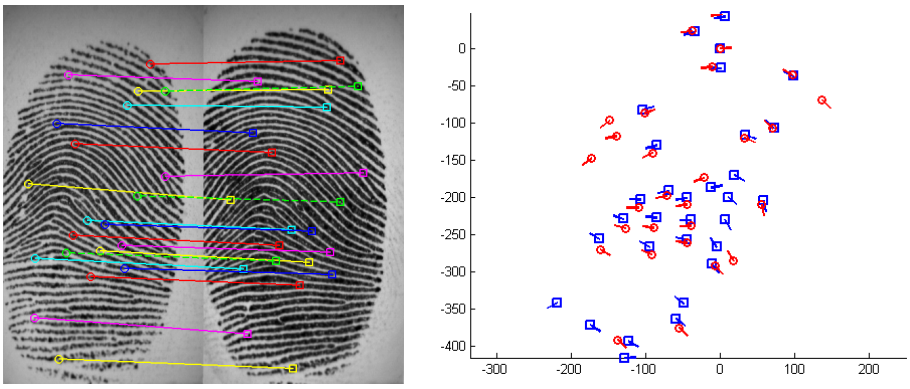
In general, fingerprint matching approaches are broadly divided into several categories, including minutiae-based approach, correlation-based methods and image-based solutions or other feature-relied matching schemes.

Minutiae-based approach is the mostly used scheme [62] due to low computation cost and good performance as it has been introduced in section 2.1. This kind of approach is to find alignment [63] of minutiae points between two pairs of fingerprint, and figure out the number of matched or correctly aligned minutiae points. A general

scheme is firstly to find a reference minutia point that generates the highest matching scores among all the combinations of two minutiae picked up from one template and another, and then two templates are aligned with this reference point. Such a transformation could be formulated as

$$\begin{bmatrix} X_1' \\ Y_1' \\ \theta_1' \end{bmatrix} = \begin{bmatrix} cos(\triangle\theta) & sin(\triangle\theta) & 0 \\ -sin(\triangle\theta) & cos(\triangle\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X_1 \\ Y_1 \\ \theta_1 \end{bmatrix} + \begin{bmatrix} \triangle X \\ \triangle Y \\ 0 \end{bmatrix} \tag{2.1}$$

where (X_1, Y_1, θ_1) and $(X_1', Y_1', \theta_1)'$ respectively represent the coordinates of a point of one template and its transformed coordinate, $(\triangle X, \triangle Y, 0)$ is the offset of translation and $\triangle\theta$ is the offset of rotation. Figure 2.8 gives examples of matched minutiae points and the transformed alignment.



(a) Matched points. (b) Plots of transformed minutiae.

Figure 2.8 – Illustration of minutiae-based matching. Figure 2.8(a) shows matched points of two templates. Figure 2.8(b) shows plot of transformed templates.

Correlation-based approach computes the correlation for the corresponding pixels of the aligned fingerprint. Thus, this requires the alignment of two fingerprints in the same orientation. In order to do this, the correlation for all possible alignments is necessary, and singularity might have to be used. These constraints increase the complexity and other disadvantages of the approach, especially for global correlation method. In this case, mostly, local correlation [64] is applied for dealing with matching problem, which is to compute the correlation of local regions mainly.

Image-based matching schemes generally extract a set of features from the grey-level image [65] and generate matching result with the comparison of two sets of extracted features in terms of a specific measurement such as distance.

In the light of the literature presented above, minutiae-based matching has distinctive advantages and is the primary choice in fingerprint recognition systems,

while the other two schemes are generally used as auxiliary approaches in some particularly difficult cases. In this manuscript, hereinafter, the matching tools adopted are all dependent on minutiae-based algorithm unless otherwise stated.

2.3 Fingerprint Quality Assessment (FQA)

Previous sections explored the general framework and procedures for processing a digital fingerprint with a refined overview of this modality. In this case, one can note the most important factor is matching accuracy for a biometric recognition system. What affect the matching result most is the robustness of the matching algorithm. However, in addition, reliable features or minutiae points are the origin of matching performance, which fully rely on the quality of the fingerprint image. In this section, fingerprint data quality is introduced in detail which is also the primary topic of this study.

2.3.1 Introduction

The FQA has been studied since the end of the last century due to its effect to the overall matching performance. According to the literature [66, 29, 67], existing studies in qualifying fingerprint image could be classified into three categories, including latent image, digitized patterns and camera photographs. In another word, this is involved in the use at both indoor environment and outdoor situation [62]. In this case, for example, luminance and temperature can be potential factors that affect the fingerprint quality while reading a sample with a camera or thermal sweeping sensor. However, this is mostly determined by the sensor type adopted by the biometric system. In addition to these external effects, sample quality is primarily determined by internal factor comes from individual condition, *i.e.* individual fingerprint surface determined how well the quality is. Without any sophisticated statement, this study mainly focuses on gray-level fingerprint pattern that had been mostly studied so far. This kind of fingerprint image is generally captured via acquisition devices introduced in section 2.2.1. The qualification of gray-level image is initially inspired by some subjective criteria [55], such as ridge-valley contrast, clarity, foreground area, etc. Researchers have also classified unexpected fingerprint images into several cases, including some commonly poor quality patterns, such as dry, creased or wrinkled, abraded and so on. Figure 2.9 demonstrates several examples of the common conditions.

As it is shown in Figure 2.9, obviously, recognition result would not be guaranteed in some cases such as the wrinkled one. These problems could be regarded as

(a) Normal (b) Dry (c) Wet (d) Creased (e) Wrinkled

Figure 2.9 – Example of fingerprint samples: Figure 2.9(a): normal fingerprint, 2.9(b): dry fingerprint, 2.9(c): wet fingerprint, 2.9(d): creased fingerprint, 2.9(e): wrinkled fingerprint. (Image source: the internet.)

subjective factors which the perception tells us. However, in many cases, the intuitively judgment is different from the requirement of the matching system. For instance, the genuine examples depicted in Figure 2.10 illustrate such a dissent between subjective perception and matching accuracy. Visually, samples in Figure 2.10(a) and 2.10(b) are relatively better than the other two samples in terms of clarity and available area. They should generate relatively high matching scores. However, the matching scores between the sample in Figure 2.10(a) and other three samples are 187, 241 and 81 respectively (computed via Bozorth3 [68]). Details of Bozorth3 can be found in Chapter 3, and a further discussion is given in section 2.3.3. Sample in Figure 2.10(c) seems less clear than sample in 2.10(b), yet generate a higher score. This might be due to the performance of the matching program, for the employed features in one of the sample is less than the other, which leads to low genuine match. This kind of situation demonstrates the difference between objective requirement and subjective qualification. Therefore, evaluating the quality

(a) Sample 1 (b) Sample 2 (c) Sample 3 (d) Sample 4

Figure 2.10 – Example of several genuine fingerprint samples. (Image source: FVC2002DB2A.)

of fingerprint sample is essential before proceeding to the next phase. Similarly, this problem exists as well for objective assessment algorithms since no one of them is perfect, which will be discussed later in this study.

To achieve such a goal, a good quality fingerprint sample should be not only a clear image that satisfies subjective assessment criteria, but also an image that is suitable for extracting sufficient and reliable features. According to this principle, a state-of-the-art literature defined their quality metric of a biometric sample as a predictor of matching performance [69]. However, this kind of mechanism largely depends on the employed matching scores which impacts on the efficiency of the quality metric. As a result, the standards [70] in biometrics propose to measure the quality of a biometric data considering three aspects, known as:

1. **Character**, which indicates the quality of sample's physical attribution;

2. **Fidelity**, which is the degree of similarity between a biometric sample and its source, and is attributed to each of sample processing stages;

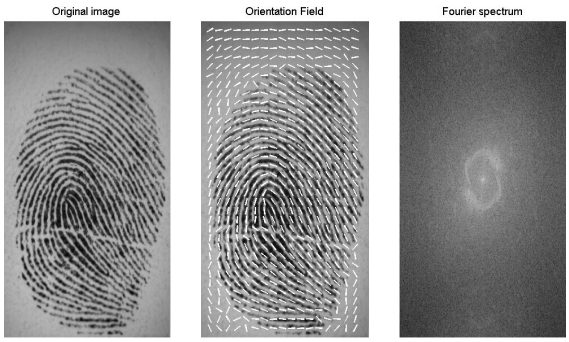
3. **Utility**, which refers to the impact of a biometric sample on the overall performance of a biometric system.

Among the three properties, utility is a function of both the character and fidelity of a sample and is generally presented as the contribution of a biometric sample to the performance in terms of recognition error rates. This property tells a truth that a good quality sample should be beneficial to matching operations.

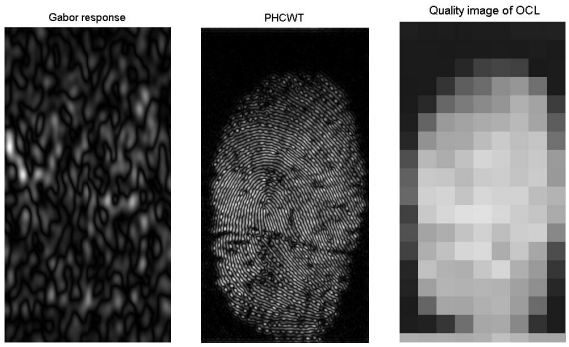
In the following, this chapter concretely discusses schemes about fingerprint data quality in terms of assessment approaches. In this study, the assessment approaches are divided into several categories by exploring most of the representative solutions in the existing studies, as given below.

2.3.2 Literature Review of FQA

Fernandez *et al.* [71] proposed a comparative study of the FQA prior to 2006, in which they categorized FQA algorithms into several classes known as local feature-based approaches, global feature-based methods and solutions with classifiers [71]. The assessing approaches reviewed in [71] can be simply summarized in several points: quality metrics based on the orientation of fingerprint pattern; algorithms depending on the variation of Gabor responses; approaches in frequency domain; measurements that are based on pixel information and quality indexes rely on classification with multi-features. In addition, that study also analyzed quality metrics in terms of linear



(a) Original (b) Orientation (c) Fourier-Spectrum



(d) Gabor (e) PHCWT (f) OCL

Figure 2.11 – Feature examples of fingerprint image: 2.11(a) is an original image; 2.11(b) is the orientation field; 2.11(c) is the Fourier spectrum; 2.11(d) is a Gabor response; 2.11(e) is the Pet Hat's wavelet and 2.11(f) is the quality image of OCL.

correlation between the metrics. Figure 2.11 shows several examples of commonly used features.

In this manuscript, we classify the existing studies into several frameworks in terms of their implementation to show their difference and some potential problems that need to consider. As it is mentioned above, the existing fingerprint quality metrics are all dependent on one or several features. According to how they are carried out, we propose the following categorization: 1) segmentation-based approaches; 2) a single feature-based quality index; 3) solutions rely on a combination of multi-features or indexes, which is further divided into methods based on linear fusion and classification. Table 2.1 shows a categorization of some of the representative studies in FQA.

In fact, FQA had been involved in other relevant studies of fingerprint, such as the contrast of image block [88] and region mask of ridge-valley pattern [49]. In table 2.1, work on the quality assessment of fingerprint image had been followed by an approach based on the segmentation of directional block and non-directional block

Table 2.1: Categories of existing fingerprint quality metrics.

Framework	Solution	Quality index or feature.
Segmentation	Image-based	Directional blocks area of foreground [72]. Count up poor foreground blocks [53] determined via Gabor features in 8 directions. Sum of directional contrast of local block [73].
	Template-based	Area of reasonable informative region [74] measured by Delaunay Triangulation of minutiae.
Single feature	Feature regularity	Cumulative total energy (CTE) in the initial few subbands of the wavelet transform of fingerprint [75]. Sum of pixels standard deviation (STD) of local block[73]. Sub-band energy of fingerprint Fourier spectrum [73]. Weight of block symmetry [76] generated by 2-order orientation tensor. Square root of the absolute value of the Pet's hat wavelet (PHCWT) coefficients [77]. Relative contrast index (CI) [78], a logrithmic ratio of the reflective intensity of the valleys to the one of the ridges. The shape of the probability distribution functions (PDFs) of ridge orientation and ridge-to-valley orientation [79].
Multi-feature	Linear fusion/regression	Orientation certainty level (OCL); Ridge frequency, ridge-to-valley thickness ratio and ridge thickness; Continuity and uniformity [80]. Global clarity score (GCS) and global orientation quality score (GOQS) [81]. Gabor feature in [53]; Ratio of foreground blocks to fingerprint image blocks; Central position [82]. Residual variances and manifold topology structure of the PCA of a block feature vector and block Harris-corner strength [83]. BLIINDS; SIFT; Root means square (RMS) of image block [84].
	Classification	Local SNR of the DFT of a signature (sine wave), unimformity and curvature [85]. 11-dimensional feature vector includes most of the existing features [69]. Auto-correlation and DCT-based features of image block [86]. A histogram of the unit activations of Self-organizing maps (SOM) obtained by training a SOM (Neural Network) with image block [87].

[72]. The quality index is represented via a ratio of the area of qualified blocks to image area. However, the determination of block prominent direction depends on the threshold which is a thorny problem for a common application. In practical, the segmentation-based measures for FQA are generally used in two ways, one is to represent quality via the qualified foreground area and another segments foreground from the image at first. For instance, Shen *et al.* [53] use the regularity of 8-direction Gabor features of image block to generate quality index. Their Gabor feature is initially used for segmenting foreground from the image, which is also involved in a threshold. In addition, some statistic measures such as standard deviation proposed for fingerprint segmentation [54] were also used for image quality assessment [73].

The use of segmentation criteria mentioned above are all associated with fingerprint image. In this thesis, we proposed an segmentation-like approach with minutiae template only, which is detailed in Chapter 6. According to these literature, one can note that foreground area is indeed a good factor for qualifying fingerprint image. In this case, multiple segmentation could be a potential solution to generate area-based quality metrics.

The second category discussed in this section are quality metrics that rely on a single feature which could be applied locally or globally to the image. For example, Nalini *et al.* [75] proposed to use cumulative energy of several subbands of the compressed image in the wavelet domain. Lee *et al.* [73] reviewed three approaches based on the fingerprint image, including local standard deviation [54], directional contrast of local block [89] and Gabor features [53]. They proposed a feature via observing the Fourier spectrum of the fingerprint image. Their quality metric depends on the pixels information of the Fourier spectrum image which is of course not always the same for different kinds of fingerprint images. Other quality metrics denoted by a single feature could also be found in [76, 78, 79], where the symmetry features decomposed via 2-order orientation tensor [76] depending on scale parameter and threshold, the contrast index (CI) relies on a mean spectrum of ridge-valley measurements and the difference of kurtosis value of the probability density functions (PDFs) [79] is not distinctive between some convex and concave shapes that are relatively smooth. Trial results in these literature show relatively good performance in comparison to baseline algorithm(s). However, threshold values and parameters are unavoidable for most of them and lead to difficulties when facing a difference scenario, for they are easily affected by image specification.

Many of the existing studies made effort in qualifying fingerprint image with

multiple features. This is generally carried out in two aspects: linear fusion with weighted coefficients and classification. Both could be associated with knowledge-based schemes [69, 84]. For instance, Lim *et al.* [80] proposed a quality metric through weighted combination of local and global quality scores that are estimated in terms of several features such as orientation certainty level (OCL) and so on. Their quality metric also involves in several thresholds to classify the local blocks into variant levels. Similarly, Chen *et al.* [81] proposed a metric by linearly combine the orientation flow (OF) and ridge-valley clarity features. Apparently, the weighted coefficients also have to be adjusted if a different image specification is involved.

In addition, this kind of approach can also be found in [82, 29, 83], where Chen *et al.* [29] estimates the power spectrum ring with Butterworth functions instead of observing directly the pixel information of the spectrum image, and Tao *et al.* [83] observed two regularities from the circle manifold topology of an order set of pixels from the image blocks and their principle component analysis (PCA). However, in addition to the coefficient problem, there are also constraints of the employed features. For example, the ridge frequency [29] depends on the resolution and image size.

Another solution with linear combination of multiple features could be illustrated by a regression-based approach [84] which adopts genetic algorithms (GA) optimizing the linear relationship between the quality value and genuine matching scores of training set samples. According to the literature [90], maximizing the correlation between the two measures is a solution for qualifying biometric sample. However, this is fully dependent on the genuine matching results. Likewise, it is also possible to question other quality metrics that are associated with a prior-knowledge of matching performance, for matching algorithms might be quite different. Similarly, to assess quality for a specific scenario, such a regression can also be implemented via other algorithms such as gradient descent [91] and Dempster-Shafer (DS) theory [92].

Quality assessment approaches with multi-feature carried out in another form is classification. Lim *et al.* [85] extended their work in [80] by classifying a certain amount of fingerprint samples with 3 different classifiers rather than calculating the quality metric. Later, the state-of-the art quality metric, NFIQ, employs 11-dimension feature to estimate a matching score and classify to five levels through a trained model of a neural network [69]. Further, in NFIQ 2.0, Olsen *et al.* [87] trained a two-layer self-organizing map (SOM neural network) to obtain a histogram of SOM unit activation with an intensity vector of image block. The histogram

is the frequency of the occurrence of the best-matching unit (with respect to the competitive layer) assigned to each block. The trained feature is then threw to a Random Forest (RF) to estimate the binned genuine matching scores (GMS). This is the first study of FQA to generate a learning-based feature by using unsupervised approach and quite large dataset. However, the quality feature is also a regularity of the proposed histogram, and the RF is to classify samples in terms of a prior-knowledge of matching score. So far, there is no studies that are able to conduct a perfect matching algorithm because the matching scores between two bad quality genuine or impostor samples are somehow unforeseeable [27].

According to such a statement, one can note that approaches with a single feature is limited to a specified image type and knowledge-based solutions are not absolutely appropriate to cross-use. Besides, it is also possible to consider whether a quality metric based-on multi-feature really makes a robust criterion or takes the advantages of them.

2.3.3 Discussion

In the previous sections, we investigate most of the literature of FQA to illustrate the solutions that had been proposed so far. A common fact is that biometric quality of fingerprint sample is not completely the same as it is estimated by subjective criteria [55, 27]. The biometric definition should be related to the matching performance which is expected to be benefited from the qualified fingerprint samples. This problem could be simply illustrated by some examples, as given in Figure 2.12. Figure 2.12

Figure 2.12 – Example of fingerprint samples that are visually different. From left (S1) to right (S4): 73_2, 7, 5, 8 of FVC2002DB2A.

shows several genuine fingerprint samples that are visually different, among which we can simply determine the leftmost one is relatively clear and complete, followed by an image with a little bit translation, a fragment print and a scattered-looking image. According to some subjective assessment criteria [55], we believe the quality

of them are different as well. There is no ground-truth of them so that we chose several metrics to generate quality values of them, as shown in table 2.2.

Table 2.2: Quality values of the samples in Figure 2.12. Metrics are from Section 2.3.2.

QM \ Sample	S1	S2	S3	S4
STD	72	74	33	31
OCL	0.66	0.63	0.41	0.40
NFIQ	1	2	2	2

The smaller value of NFIQ represents good quality, while other two metrics give an inverse representation. Apparently, sample S3 is only a partial image and sample S4 is not a clear image with relatively small foreground area. In this case, these two samples should have relatively bad quality values. However, NFIQ gives both a quality value of level 2, indicating their qualities are better than average level. As one cannot assert that this result is not reasonable without the ground-truth of them, we simply calculate the GMS by assigning every of them as the enrollment, results are given in table 2.3.

Table 2.3: Genuine matching scores calculate by using Bozorth3.

Enroll \ Sample	S1	S2	S3	S4
S1 enroll		266	72	230
S2 enroll	263		95	231
S3 enroll	72	95		43
S4 enroll	228	230	43	

According to the matching scores, one can note that S3 leads to relatively lower values when it is used as the enrollment. Although the matching is fully dependent on minutiae including minutiae number but good quality sample should be suitable for matching [27], particularly genuine matching.

2.4 Conclusion

According to the discussions given in previous sections, a preliminary body of knowledge about fingerprint modality and its qualification could be acquired. Meanwhile,

one can note that FQA is still an open issue due to several problems:

1. Some assessment approaches based on multi-feature and a prior-knowledge of matching performance is not absolutely appropriate for multi-vendor applications, for the intra variability and inter similarity could be quite different when an external matcher is employed. Furthermore, multi-feature fusion does not always generate a more robust quality index than some metrics that rely on a single feature. A comparative study could demonstrate such an observation and is given in later chapters.

2. Recent studies began focusing on the fingerprint modality for mobile devices which make high resolution image and camera photograph available, and hence requires qualifying image via different solutions.

3. Quality score normalization is also a challenge due to the variety of diversely collected samples, which limited the effect of normalization techniques [93, 94].

To this end, this chapter gives a longitudinal study of fingerprint modality including quality assessment. This manuscript hereinafter presents the study of the FQA and metric evaluation in details, which mainly includes a new evaluation framework and several quality metrics implemented in different ways.

One important question concerns the validation of such a fingerprint quality metric. We discuss this aspect in the next chapter.

Chapter 3

Validation of Biometric Quality Assessment

This chapter addresses the validation/evaluation approaches related to the quality metric of biometric modalities. The main contribution of this chapter is a new framework for validating quality metrics, which is defined in terms of the utility property of biometric sample and is carried out by considering the impact of the enrollment sample(s) on the overall performance.

Contents

3.1 Introduction

Quality control works as a toll-gate to guarantee the matching performance of a biometric system by forbidding bad quality samples. In this case, the validity of a quality metric should be verified via its contribution to the overall performance. This study firstly pays attention to the evaluation approaches with respect to fingerprint quality metrics. To evaluate quality metrics, existing studies have conducted evaluation with many different criteria, including subjective measurements and objective solutions. In this chapter, first of all, a quick review of some evaluation algorithms related to fingerprint quality metric is presented, and then a generic validation framework proposed in this study is given in detail.

3.2 Related Works

As it has been discussed in Chapter 1, Phillips et al. [30] pointed out the significance of protocols for system level evaluation. However, some earlier studies in evaluating fingerprint quality metrics do not provide totally open protocols such as the employed dataset [27]. In addition, some other studies directly relate the evaluation of quality metric to the elements of subjective assessment [81], and some approaches proposed later choose to evaluate their metrics with the existing ones [75].

Similarly, Shen *et al.* [53] divided fingerprint images into several classes according to samples' quality type and compared their approach with another via computing the proportion of correctly classified samples. Some others propose to compute a quality benchmark through the observation of automatically detected minutiae of fingerprint [80, 81]. These approaches failed to explicitly reflect the relation between quality metric and matching performance. In addition, these attempts are more or less related to subjective observations when validating quality metric, such as the manual classification of fingerprint quality type, differentiating spurious minutiae or missed minutiae from the template. However, this kind of operation can be easily achieved by employing a synthetic fingerprint generator such as SFinGe [95].

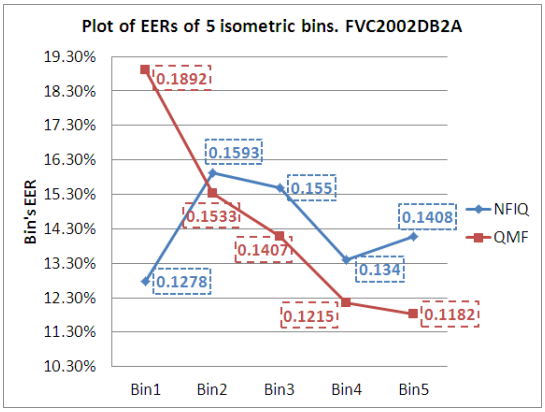


Figure 3.1 – Plots of 5 isometric bins' EER values, where blue points are obtained by NFIQ and red points correspond to QMF.

Tabassi *et al.* [69] defined biometric sample quality as a predictor of matching performance, for they observed that biometric samples of good quality should produce relatively high genuine matching scores (GMS) which are well separated from impostor matching scores (IMS). However, the prediction is totally dependent on the performance of the employed matching algorithm. Chen *et al.* [29] proposed that the EER should be monotonically decreased after a certain part of samples of bad qualities had been pruned. In addition, they also considered that the detected minu-

tiae of good quality samples should not greatly varied after enhancement. Another consideration of their study is that the detected minutiae of good quality samples should have relatively good consistency with manual groundtruth of minutiae. This is somehow a coherent principle when dealing with minutiae-based matching algorithms. Obviously, matching performance could be guaranteed if minutiae are precisely detected. This assumption is restricted by the underlying matching techniques, i.e. minutiae-based matching approaches. Figure 3.1 illustrates an example of Chen's approach obtained by using NFIQ and QMF [96] on one FVC database.

As it is illustrated in 3.1, this evaluation criterion requires an equivalency of sample number in each bin, otherwise the EER values of the bins could be affected. A valid quality metric should generate monotonically decreasing (or increasing) EER values of the bins. In addition, it is necessary to consider whether this approach is also appropriate to a quality metric that represents biometric sample qualities with several labels, because the difference between samples of the same label is unknown, and hence it is inappropriate to get samples of variant labels divided into the same bin, *i.e.* the EER values could be seriously contaminated by the outliers of each other. For example, the result of NFIQ in Figure 3.1 shows a disordered plot of the EER values of the bins. The details about 'isometric bins' could be found in the reference article.

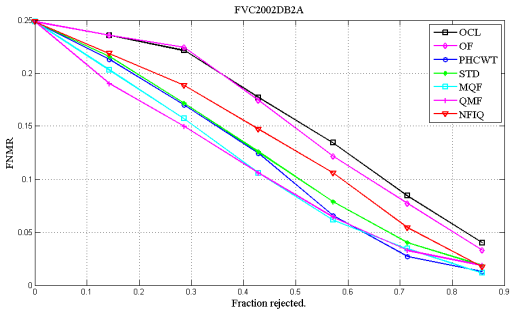


Figure 3.2 – An illustration of the error versus reject curve. Database is FVC2002DB2A, and selected metrics had been presented in Chapter 2.

Grother *et al.* [27] discussed that the quality measure of a biometric sample is generally employed within 3 different cases, including enrollment phase, verification task and identification. They proposed several evaluation approaches associated to matching threshold and quality levels, including rank-ordered DET curve, error versus reject curve and the approach based on the Kolmogorov Smirnov (KS) test. As they are closely related to the quality level, some of them may not be completely suitable for quality metrics with a larger range values, such as KS test-based approach.

The Figure 3.2 shows another form of the error versus reject curve which is independent to the quality level. For example, one database with M individuals and N samples per individual, the verification operation could be simulated via assigning a specified sample (for example, the first one) of each individual as the enrollment, while other $N-1$ samples act as authentication samples. Each time, one authentication sample with the lowest quality is eliminated for each individual, i.e. a fixed percentage of low quality genuine samples are removed each time. The FNMR then can be calculated with a threshold which is similar to the definition in [27].

3.3 Validation based on Enrollment Selection (ES)

According to the statement addressed in Chapter 2 and Section 3.2, one can observe that a quality metric acts like a black box producing a measure of one biometric sample. No matter how a quality metric is defined, a valid metric is expected to define good quality in terms of the positive contribution to the matching performance, which is known as a prediction of matching performance [27]. However, this is not an absolute assertion due to the limitation of matching algorithms. In this case, the evaluation of a quality metric might not be always similar when different comparison tools are involved. Therefore, the validity of quality metric proposed in this study does not emphasize what kind of criteria are used for making a definition of 'good' or 'bad' quality, such as clear ridge-valley pattern, good local orientation uniformity and so on. The 'good' quality in this study is simply considered as the capability of a biometric sample contributing to the degradation of matching error. Hence, the validation of quality metric in this study is totally indicated by the overall performance obtained by considering samples' quality.

According to Philips *et al.* [30], generally, a system level quality control is performed after the capture session. This process would be executed as a loop to acquire a qualified sample which determines the quality of the final enrollment sample of an individual. The re-capture at the same session might result in Failure to Enroll (FTE) and the Failure to Acquire (FTA) caused either by algorithm crash or sensor overtime.

However, to validate a quality metric, it is avoidable to consider this problem involved in the capture and enrollment sessions as it had just been mentioned. The contribution of this study, therefore, made an effort to provide a general validation approach of biometric sample quality metric by considering the impact of enrollment variations on matching performance. No matter under what framework (system

level or technology) the evaluation test is performed, enrollment can be a supervised process. In this case, without emphasizing the test type, the validation framework proposed in this paper is an algorithm level approach in the enrollment phase which ensures that there is no need to consider effects on FNMR and FMR raised by authentication mode, *i.e.* verification or identification. The proposed framework is denoted as the Enrollment Selection which relies on both the quality value and the EER (or AUC) value of a biometric sample, and the later is an objective measure representing the utility value of the sample, details are given in the following.

3.3.1 Algorithm description

The validation framework is consisted of an employed matcher $R(\cdot,\cdot)$, a quality metric $QI(x)$ that need to be evaluated and a trial dataset $D_{M\times N}$. The validation is totally dependent on the matching scores and quality values generated from $D_{M\times N}$.

3.3.1.1 Enrollment sample and matching score

In this validation framework, the trial dataset is supposed to have M individuals and N samples per individual. The enrollment sample is hence defined for each individual, as illustrated by Figure 3.3.

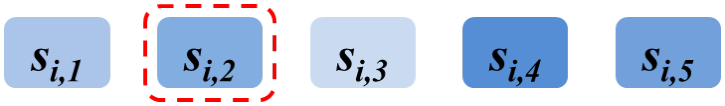


Figure 3.3 – Illustration of enrollment sample and authentication samples.

Figure 3.3 represents five samples of one individual of a database, where the sample marked by a rectangle in red dash lines represents the enrollment sample. Therefore, other four samples are used for authentications once the enrollment is determined.

In this validation framework, the intra-class and inter-class matching scores associated to an enrollment $(S_{i,j})$ is given by $N-1$ genuine matching scores (GMS) [97]

$$gms_{i,j,k} = R\left(S_{i,j}, S_{i,k}\right) \;\; j{\neq}k \tag{3.1}$$

and $N-1{\times}M-1$ impostor matching scores (IMS)

$$ims_{i,j,l,k} = R\left(S_{i,j}, S_{l,k}\right) \;\; i{\neq}l \;\; and \;\; j{\neq}k, \tag{3.2}$$

where i and l represent the subscript of individuals, j and k are sample subscripts, respectively.

3.3.1.2 Sample utility and quality

The proposed validation framework depends on several indexes for enrollment selection. The first one is an objective index represented by sample EER value. According to the definition in Section 3.3.1.1, one can calculate a FMR and a FNMR for each sample of each individual when it is specified as the enrollment. The FMR and FNMR of one sample are computed from a given set of threshold t:

$$\begin{aligned} FNMR_{i,j}(t) &= \frac{card\{gms_{i,j,k} | gms_{i,j,k} < t\}}{N-1}, \\ FMR_{i,j}(t) &= \frac{card\{ims_{i,j,l,k} | ims_{i,j,l,k} \geq t\}}{(N-1)\times(M-1)} \end{aligned} \tag{3.3}$$

where '*card*' denote the cardinality of a given set of matching scores. The sample EER of $S_{i,j}$, $SEER_{i,j}$, is simply computed as the point (of error rates) where $FNMR_{i,j}(t) = FMR_{i,j}(t)$.

Therefore, with a $SEER_{i,j}$ of one sample, one can have a measure that how much the contribution of a sample is within the experimental framework consisted of employed datasets and matching algorithms. In this case, the sample EER is regarded as a substitute of the ground-truth of a sample within the involved matching framework, which is also denoted as **sample utility** in this thesis.

Another measurement of each sample is simply the quality value generated by an involved $QI(x)$, which gives

$$q_{i,j} = QI(S_{i,j}) \tag{3.4}$$

for each sample.

3.3.1.3 Selection indexes

According to the description above, such a calculation results in a total of M-by-N sample EER values and sample qualities (in the same size) for one database, by which one can perform enrollment selection (ES) in three cases: The first case of ES is carried out by choosing the best sample of each individual as the enrollment. The 'best' here means that the $SEER$ of the selected sample is the smallest (best) among all the samples of one individual. This is defined by the consideration that one matcher cannot obtain better performance than this case from a given dataset. Likewise, one can achieve another ES in the opposite way which indicates the worst case. In this paper, the two cases are denoted as the **best utility** and the **worst utility**, respectively.

Figure 3.4 demonstrates the enrollment selection in terms of the best utility, in which the light color represents relative smaller sample EER value.

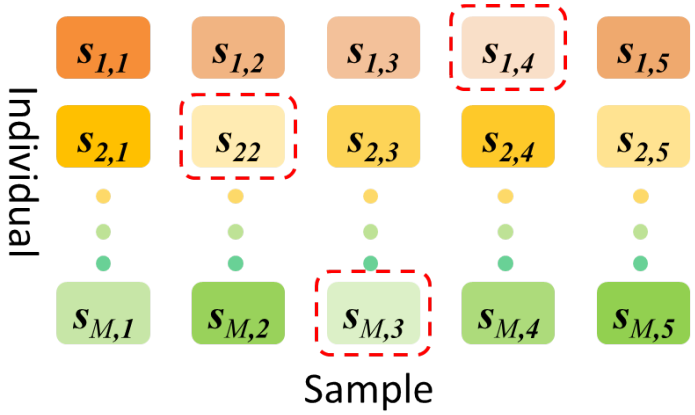


Figure 3.4 – Illustration of enrollment selection framework. $S_{2,N}$ is selected as enrollment in case $SEER_{2,N} = min\,(SEER_{2,1},\ \cdots,\ SEER_{2,N})$.

The third case is simply the enrollment selection in terms of the best quality for each individual. The enrollment sample of each individual is determined by doing so. In each case, according to equation 3.1 and 3.2 one can calculate again the intra-class and inter-class matching scores for the whole dataset by using the assigned enrollments. Finally, several global measurements can be figured out via the intra-class and inter-class matching scores of the trial dataset. The details are given in the following.

3.3.1.4 Global measures and AUC ratio

The intra-class and inter-class matching scores of one trial database $D_{M,N}$ are respectively consisted of $(N-1)\times M$ genuine matching scores

$$GMS = \{gms_{i,j,k} \mid i\in(1,\cdots,M),\ (j,k)\in(1,\cdots,N)\ and\ j\neq k\}, \tag{3.5}$$

and $(N-1)\times(M^2-M)$ impostor matching scores

$$IMS = \{ims_{i,j,l,k} \mid (i,l)\in(1,\cdots,M)\ and\ i\neq l,\ (j,k)\in(1,\cdots,N)\ and\ j\neq k\}. \tag{3.6}$$

Both expression 3.5 and 3.6 indicate that the comparison is only performed between the specified enrollment sample and other non-enrollment samples.

Further, by given a set a threshold t, one can calculate the FMR and FNMR of the trial dataset, respectively formulated as

$$\begin{aligned} FNMR(t) &= \frac{card\{gms_{i,j,k} | gms_{i,j,k} < t\}}{(N-1)\times M}, \\ FMR(t) &= \frac{card\{ims_{i,j,l,k} | ims_{i,j,l,k} \geq t\}}{(N-1)\times(M^2-M)} \end{aligned} \tag{3.7}$$

where '*card*' represents as well the cardinality as presented in expression 3.3.

In return, one can generate a global ROC curve via a plot of the FMR against the FNMR (in eq. 3.7). Meanwhile, one can calculate a global AUC value and a global EER at a point where $FMR(t) = FNMR(t)$.

By such a computation, one can have a graphical illustration and two quantitative measures for each of the three selection cases discussed in Section 3.3.1.3. According to the definitions above, for a relatively good metric, one can consider that the overall performance indicated by the ROC curve should be much close to the best case when choosing samples of good quality as the enrollment (the global EER and AUC either). A valid quality metric would satisfy this condition as much as possible.

In this case, we also define a global index namely **AUC ratio** by using three global AUC values obtained via the **best utility**, **worst utility** and the best quality, respectively. The definition is formulated as

$$r_{auc} = \frac{Q_{auc} - W_{auc}}{B_{auc} - W_{auc}}, \tag{3.8}$$

where Q_{auc} is the global AUC value of the ROC curve computed in terms of the best quality samples as it had just been mentioned, B_{auc} and W_{auc} are the global AUC values correspond to the ROC curves of the best and worst utilities. Obviously, the larger the value of r_{auc} , the better the performance of quality metric is.

Figure 3.5 demonstrates a result of these calculations.

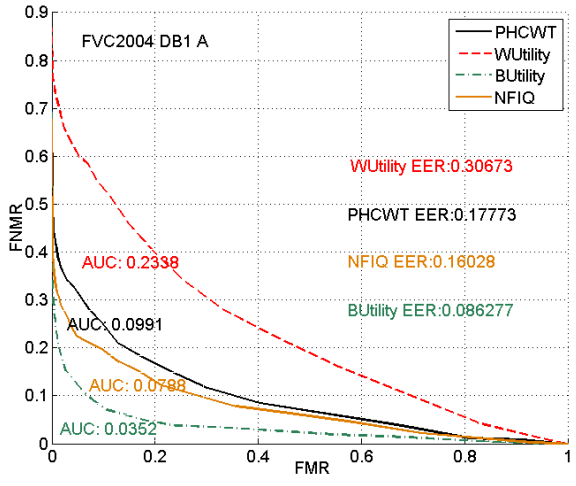


Figure 3.5 – Graphical illustration of the enrollment selection result.

The result in Figure 3.5 are obtained by using NBIS matching software Bozorth3 [68], and the **sample utility** is also calculated by using the matching scores of the Bozorth3. **Note** that the **sample utility** values must be calculated via the same

matching algorithm when using the proposed validation framework to estimate the performance of a quality metric.

In Figure 3.5, one can note that the NFIQ obtains a better result in comparison to the PHCWT (given in Chapter 2). The ROC curves are able to clearly show the difference between the performance of them, where the ROC curve of the NFIQ is closer to the best case than the one of the PHCWT. Both the global EER and AUC values of each case are given in the figure. Meanwhile, the **AUC ratio** of the NFIQ and PHCWT are 0.7805 and 0.6782, respectively.

Apparently, the proposed validation framework is fully able to estimate the performance of a biometric quality metric. However, as the dimension of the GMS and IMS are quite different from each other, *i.e.* the IMS could be much longer than the GMS when the size of the trial dataset is getting larger. Because of this, we also define the global measures with a bootstrapping method, which is detailed in the following section.

3.3.2 Global measures with the Confidence Interval (CI)

As it has just been discussed above, a Confidence Interval (CI) of each global measure is defined via the bootstrapping method [35]. The bootstrapping method enables us having a set of bootstrap samples of a global measure, and hence one can calculate the CI of the associated global measure. The bootstrap of the enrollment selection is dependent on random sampling of both the IMS and GMS with replacement, for the global measures are derived from the two sets of matching scores. In this case, the bootstrap is performed by

1 A set of L samples are randomly chosen from the $M\times(N-1)$ intra-class matching scores which are calculated after enrollment selection.

2 Similarly, another set of L samples are randomly chosen from the $(N-1)\times(M^2-M)$ inter-class matching scores.

3 The global measure given in Section 3.3.1.4 could be calculated from the two sets of randomly selected matching scores. For instance, a global EER can be figured out in a similar way to equation 3.7 (by using L instead of $(N-1)\times M$ for FNMR and FMR either).

4 Step 1 to 3 are performed for 1000 iterations, which achieves having 1000 bootstrap samples of a global measure and an average of the 1000 samples of the global measure are preserved at last.

Note that the size L of the matching score samples could be changed according to the dataset dimension. The changing is to ensure the diversity of the matching samples, especially for impostor matching scores.

According to the definition above, we can quantify the performance of a quality metric with **an average value** associated to a global measure. However, a further validation needs to be acquired for making a significant comparison between variant quality metrics, especially for comparing them in a statistical manner. To do so, the confidence interval (CI) [98, 99] at 95% level is defined for a global measure with those bootstrapping samples. In biometrics, the confidence intervals of two measurements are able to indicate the statistical difference between them, if their confidence intervals do not overlap each other [100]. By doing so, the difference between quality metrics could be determined statistically.

To define the CI of the global measure, the global EER is used **hereinafter** in this thesis, and the CI of other global measures can be calculated in the same way. The CI is formulated as the follows:

$\forall$ bootstrap samples of the global EER X_r.

$\exists$ a population $X = \{X_r | r = 1, \cdots, 1000\}$, and its CI at the confidence level of 95% is given by

$$CI = [\overline{X} - \frac{\sigma}{\sqrt{1000}}\mu_{\alpha/2},\ \overline{X} + \frac{\sigma}{\sqrt{1000}}\mu_{\alpha/2}], \tag{3.9}$$

where $\alpha = 100\% - 95\%$, $\overline{X}$ is the average of X, σ is the standard deviation of X and $\mu_{\alpha/2}$ is the $\alpha/2$ quantile. Details about CI and bootstrap could be found in any statistic book.

Note that the global EER value hereinafter is represented by the average value defined in this section unless otherwise stated. A full demonstration of the global EER and its CI is presented in the experimental section.

3.3.3 Monotonical global EERs

The description above is denoted as a simple enrollment selection. A more strict usage could be conducted with monotonically global EERs, *i.e.* the enrollment selection could be done by orderly choosing every of the samples that had been sorted in terms of quality as the enrollment. Correspondingly, one can obtain N global EER values and they should be in a monotonically increasing (or decreasing) order in an ideal case. However, it fully relies on the matching algorithm and the quality metric, and hence cannot be easily achieved.

3.4 Experimental demonstration

In this section, several experiments are performed to demonstrate how the proposed validation approach is used for estimating a quality metric and making comparison between variant metrics.

3.4.1 Protocol and database

Five datasets of the Fingerprint Verification Competition (FVC) are used in the experiment: FVC2000 DB2A, FVC2002DB2A and the first three databases of FVC2004 [101]. They are created by using different sensors, and the details are given in Table 3.1.

Table 3.1: Dataset specification.

DB	Sensor	Dim.	Resolution
00DB2A	Low-cost Capacitive	256×364	500dpi
02DB2A	Optical	296×560	569dpi
04DB1A	Optical	640×480	500dpi
04DB2A	Optical	328×364	500dpi
04DB3A	Thermal	300×480	512dpi

Each database contains 100 individuals and each of them has a total of 8 gray level samples, *i.e.* M=100, N=8 in the experiment. Note that these datasets are employed throughout all this thesis unless otherwise stated. A glance of the datasets are given by several samples in Figure 3.6.

The software used in the experiment includes minutiae detection and matching tools, and quality metrics. Grother *et al.* proposed that a quality algorithm is possible to show a generality or interoperability to multiple matchers [27]. They believed that the generality would not be certainly predicted as the behavior of matching algorithms is different. In order to test such a purpose, this study conducts validation experiments on two different matching algorithms, one is NIST matching approach namely Bozorth3 which involves in a minutiae detection application MINDTCT [68], and another set is implemented via a commercial fingerprint SDK. The MINDTCT is an extractor generating INCITS 378-2004 standard minutiae template, while the commercial SDK has 6 options of the existing minutiae template standards [60]. In the experiment, the minutiae templates of the ISO/IEC 19794-2:2005 standard have been extracted. Both two matching approaches produce an integral score for each comparison of two fingerprint minutiae templates. Correspondingly, the experiments

Figure 3.6 – Illustration of dataset samples.

involve 7 genuine matching operations and 693 impostor matchings for each sample of one individual. Totally, 5600 GMS and 554400 IMS need to compute for each database.

To demonstrate the validation approach of this study, two fingerprint quality metrics are adopted. The first quality metric is denoted anonymously as Q1 which represents sample quality via discrete labels from 1 to 5, where low value denotes good quality. Another one is a trial metric denoted as Q2 which generates continuous quality values and they are normalized into [0, 100] representing quality in an ascending order. **Note** that Q1 relies on multi-feature fusion and a prior-knowledge of matching performance, while Q2 depends on a single feature. This is able to conduct a comparative study as we had concluded in Chapter 2.

3.4.2 Results

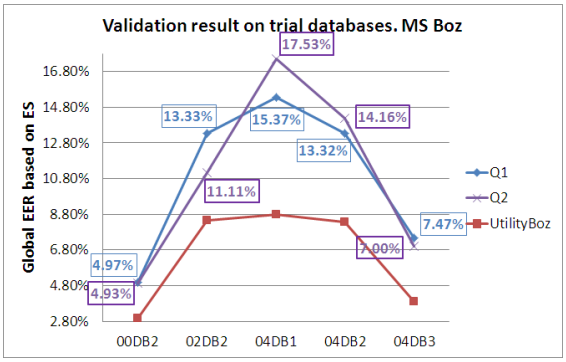
In this section, both the **sample utility** and quality measurements of fingerprint samples are used for demonstrating the proposed enrollment selection (ES) approach via an interoperate study.

First of all, according to Section 3.3.1.2, two sets of **sample utility** values are calculated by using the matching score of the Bozorth3 and the matching scores of the SDK, respectively. The ROC curves are not given anymore since this section. In this case, the enrollment selection based on the **sample utility** is performed

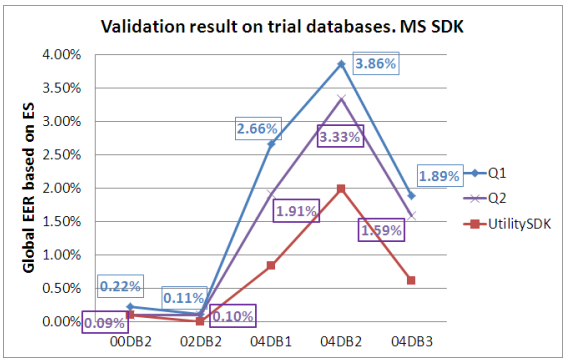
in terms of the '**best**' criterion defined in Section 3.3.1.3. Global EER values are computed in terms of the definition in Section 3.3.2.

3.4.2.1 Global EER

First, the ES is performed with two quality metrics. Results are given by several plots of the global EER values, see Figure 3.7.



(a) Result based on MS Boz



(b) Result based on MS SDK

Figure 3.7 – Plots of global EERs obtained by using quality-based ES. Figure 3.7(a) is the result based on MS Boz and Figure 3.7(b) corresponds to MS SDK.

Figure 3.7(a) and Figure 3.7(b) illustrate the results obtained with MS Boz and MS SDK, respectively. In each figure, the global EERs based on the associated **sample utility** (best case for ES) are plotted as a reference. **Note** that the **sample utility** values calculated from the matching scores of the Bozorth3 (MS Boz) and the matching scores of the SDK (MS SDK) are (hereinafter) denoted as 'UtilityBoz' and 'UtilitySDK', respectively.

Apparently, by observing the global EER values in Figure 3.7(a), Q2 performs relatively bad on 04DB1 and 04DB2 when MS Boz is involved. The global EER obtained by Q2 for these two datasets are 17.53% and 14.16%, while the counterparts

of Q1 are 15.37% and 13.32%, respectively. However, in Figure 3.7(b), one can found that Q2 generates better results than Q2 for each database when a vendor-free matcher (Q1 is related to MS Boz) is used in the experiment.

Table 3.2: The 95% confidence interval of EER of each quality metric.

QM / DB	Q1	Q2
00DB2A (NBIS)	[0.0492 0.0502]	[0.0488 0.0499]
02DB2A (NBIS)	[0.1326 0.1340]	[0.1103 0.1119]
04DB1A (NBIS)	[0.1529 0.1545]	[0.1744 0.1762]
04DB2A (NBIS)	[0.1321 0.1344]	[0.1407 0.1425]
04DB3A (NBIS)	[0.0741 0.0752]	[0.0694 0.0706]
00DB2A (SDK)	[0.0021 0.0023]	[0.0008 0.0009]
02DB2A (SDK)	[0.0011 0.0013]	[0.0010 0.0011]
04DB1A (SDK)	[0.0268 0.0276]	[0.0188 0.0194]
04DB2A (SDK)	[0.0390 0.0402]	[0.0327 0.0338]
04DB3A (SDK)	[0.0190 0.0195]	[0.0159 0.0164]

By doing so, from the figures, one can note that the shape of the curves are basically consistent with the plots of the **sample utility**, especially in Figure 3.7(b) meaning that a good matching algorithm may blurs a quality metric, *i.e.* it is easier to approach to a relatively better matching result if the matcher is relatively robust such as the result of Q2 on 00DB2 in Figure 3.7(b). In addition, in Figure 3.7(b), the difference between the two metrics on 02DB2 is not distinctive which could be observed by the CI, see Table 3.2.

3.4.2.2 AUC Ratio

Second, a demonstration of the AUC ratio addressed in Section 3.3.1.4 is presented as well. By using the AUC values (average values as well), the AUC ratio of each quality metric is calculated in terms of equation 3.8, see results obtained from two of the trial datasets in Table 3.3.

Table 3.3: AUC ratio of each quality metric based on two sets of matching scores.

AUC R. / DB	Q1	Q2
02DB2 Boz	0.8280	0.9020
02DB2 SDK	0.9910	0.9980
04DB1 Boz	0.7699	0.6676
04DB1 SDK	0.8431	0.8471

In Table 3.3, it shows the AUC ratios of the two quality metrics obtained from 02DB2A and 04DB1A, where 'Boz' and 'SDK' represent two groups of matching scores, respectively. The results are completely in line with the ones given in Figure 3.7.

3.4.2.3 Monotonic Global EER

As it is presented in Section 3.3.3, a more sophisticate case for validating a quality metric with the proposed approach is a set of monotonically varied global EER values. However, it is not easy to achieve this kind of result. We simply use one of the trial datasets to illustrate this scheme, see Figure 3.8. In the figure, the plots indicated

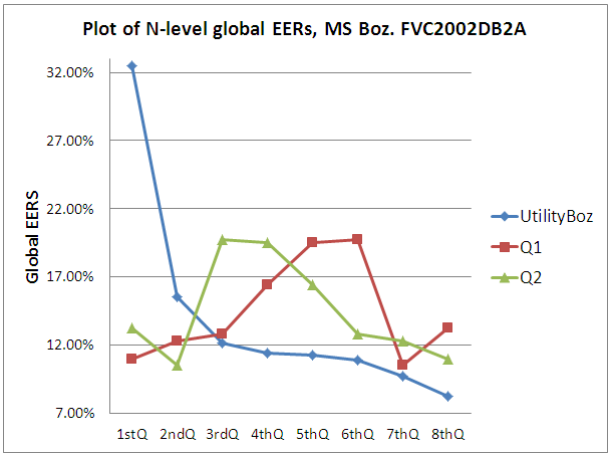


Figure 3.8 – N-level (N is 8 here) global EER values based on enrollment selection. The graph given here is only the result obtained from matching scores of Bozorth3.

by 'UtilityBoz' are global EERs (best case) obtained by using the **sample utility** values computed with MS Boz, while other two groups of plots are results of the quality metrics. The result based on **sample utility** is simply given as a reference. The first level ($1^{th}Q$) of each plot represents the case that the worst quality sample of each individual is used as enrollment, followed by the second and so on. In this case, loosely speaking, it should generate monotonically increasing global EER series if the quality metric is effective enough.

In Figure 3.8, one can note that sample **utility** satisfies the validation criterion, where other two metrics failed to generate monotonic global EER values. This result can demonstrate that it is quite difficult to observe similarity between sample quality and its utility unless the quality metric is fully relevant to the matching algorithm, *i.e.* quality metric is not an absolutely linear predictor of the matching performance, at least for the overall performance. The GMS in some cases can be predicted as the

genuine matchings mostly produce reasonable results. However, when bad quality samples are involved in a comparison, both GMS and IMS are unforeseen [27]. In this case, some studies estimate quality metrics with only genuine matchings [102]. In this study, we believe a good quality metric is helpful to the degradation of the global error rates.

Likewise, the result obtained by using MS SDK is also given, see Figure 3.9.

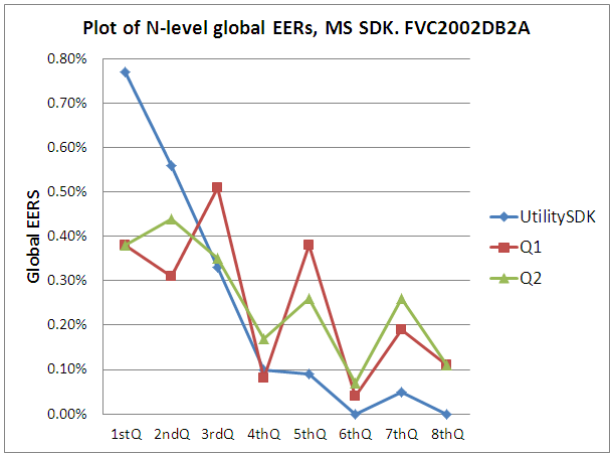
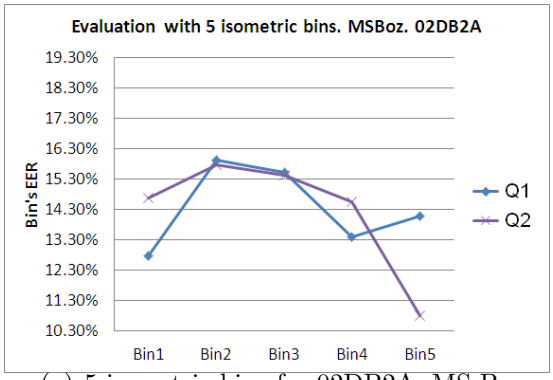


Figure 3.9 – N-level (N is 8 here) global EER values based on enrollment selection. The graph given here is only the result obtained from matching scores of the SDK.

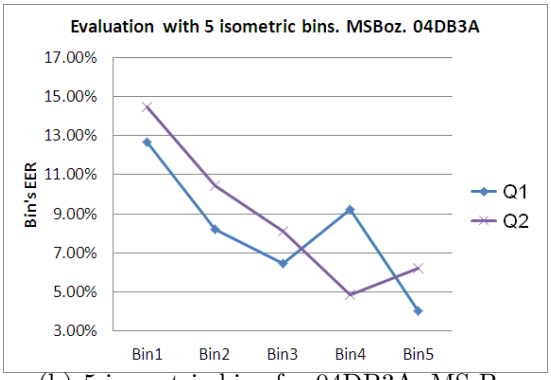
Apparently, according to the plots in Figure 3.9, one can note that matching algorithm significantly impacts on the evaluation results. Besides, each point is actually a global EER value computed in the way presented in Section 3.3.2. Thus, there is a very small error which becomes more distinctive as the overall matching performance approach to 0. In the experiment, this error is within 5 increment of the precision of the global EER value, which is calculated via observing the variation of this measure with 1000 sample. At this point, the FNMR given in Section 3.2 could be an auxiliary option for evaluating quality metric, for the impostor matching is not involved.

3.4.2.4 Pearson Correlation

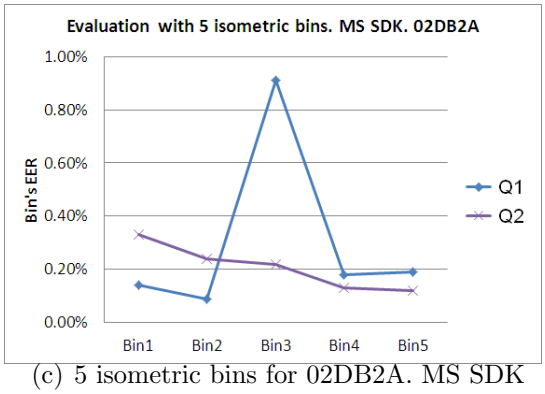
In Section 3.2, we reviewed an evaluation approach based on isometric bins [29]. The monotonically varied EER value of each bin demonstrates the similarity between sample GMS and its quality, which could be measured by the Pearson correlation coefficient [103] between the maximum GMS of the sample and the quality value simply as we mentioned in Section 3.2. In this part, we use 02DB2 and 04DB3 to show this property. First of all, the maximum value among the $N-1$ GMS of each



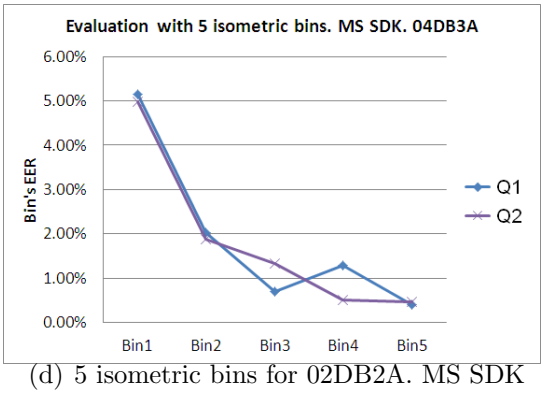
(a) 5 isometric bins for 02DB2A. MS Boz



(b) 5 isometric bins for 04DB3A. MS Boz



(c) 5 isometric bins for 02DB2A. MS SDK



(d) 5 isometric bins for 02DB2A. MS SDK

Figure 3.10 – Plots of bins EER values. 3.10(a) and 3.10(b) are results of 02DB2 and 04DB3 obtained by using MS Boz; 3.10(c) and 3.10(d) are counterparts based on MS SDK.

sample is figure out, which results in a $M \times N$ matrix of maximum value. Next, we calculate the Pearson correlation between the maximum GMS and quality values. A graphical illustration is given in Figure 3.10.

In Figure 3.10, one can note that two metrics satisfied the evaluation criterion for 04DB3 when the matching score of SDK is involved, and partially weird on this database when using MS Boz. However, Q2 shows a monotonically decreasing error rate on 02DB2 when MS Boz is used, but it is not a distinctive variation and hence means a low linearity. To show this problem, the Pearson correlation coefficients of the four cases are given in Table 3.4.

Table 3.4: AUC ratio of each quality metric based on two set of matching scores.

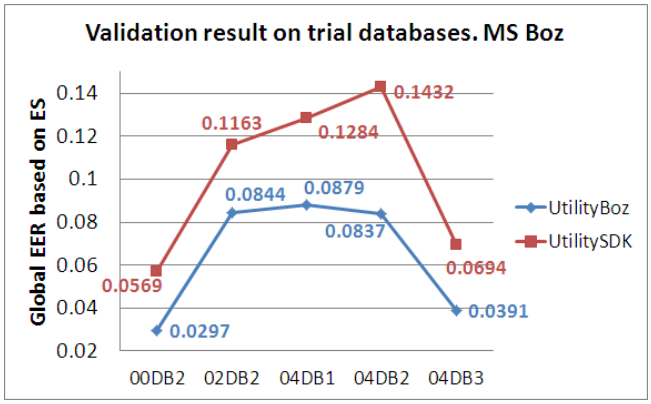
DB et MS / Metric Corr	02DB2 Boz	04DB3 Boz	02DB2 SDK	04DB3 SDK
corr Q1	-0.2728	-0.3357	-0.2505	-0.4521
corr Q2	0.2179	0.4851	0.2485	0.6188

Apparently, according to the correlation coefficients, quality metrics that obtained clearly decreasing EER values demonstrate higher similarity with the maximum GMS. The variation of the similarity (absolute coefficient higher than 0.3) also shows the effect of the matching algorithm.

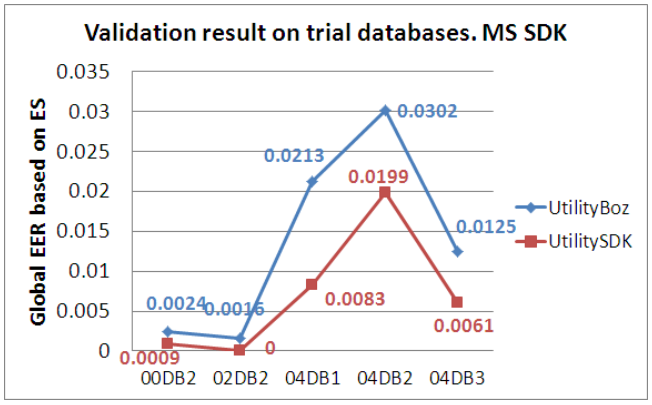
3.4.2.5 Discussion with sample utility

In this section, the two sets of the **sample utility** values are independently used as two groups of **quality** values, and substituted to the enrollment selection algorithm. The enrollment selection is also performed in terms of the 'best' criterion defined in Section 3.3.1.3. This is to demonstrate a limitation that some quality metrics based on learning a prior-knowledge of matching scores may not be an appropriate solution for a general application, for the results of the interoperate evaluations could be quite different even the prior-knowledge is only genuine matching scores.

The validation results by using sample utility are given in Figure 3.11.



(a) Matcher of NBIS



(b) Matcher of SDK

Figure 3.11 – Plots of global EERs obtained by using utility-based ES. Figure 3.11(a) is the result based on MS Boz and Figure 3.11(b) corresponds to MS SDK.

Figure 3.11 gives the global EERs obtained from the five trial datasets by using these two sets of **sample utility** values. According to Figure 3.11, both the UtilityBoz and UtilitySDK are all the approximations of the ground-truth of the samples but only valid to the matching algorithm of their own. Similarly, one can consider whether the evaluation result is robust or not when using different matchers to estimate the performance of a quality metric based on a prior-knowledge of matching scores. This limitation has already been demonstrated in Section 3.4.2.1, where the result of Q1 varies greatly when it is being evaluated with different matching algorithms. Similarly, next section gives a further investigation of this problem. **Note** that we don't assert reults of interoperate evaluations should be the same because the matching algorithms are different from each other. However, such a experimental result is able to show this problem clearly.

3.5 Case study: No-Image Minutiae Selection (NIMS)

The statement in previous section gives a complete definition of the enrollment selection. In this section, an auxiliary method is presented for validating the quality metric, by which enrollment selection is extended as dual-step operation. This part is carried out via No-image Minutiae Selection (NIMS) [104] for reducing the size of original minutiae templates.

3.5.1 Background

Due to the advantages in privacy and efficiency requirements, minutiae template-based matching is the dominant technique among the authentication approaches of fingerprint image [77]. The application of minutiae-based matching mainly involves two categories: resource-free systems and embedded employments such as smart-card [105]. The former has almost no limitation of the computing cost and the storage requirement. However, these factors are the prerequisites for the embedded applications, especially for those match-on-card (MOC) systems [106]. In this case, to satisfy such requirements, some on-card applications support only a minutiae template that contains a certain amount of minutiae such as 60 points. Generally, the amount of minutia points of one template could be less than 130 and it is fully dependent on the extractor and the image quality. Therefore, it is necessary to perform a removal or selection of minutiae from the original template before the templates can be used by those resource-limited applications. Meanwhile, such an operation should be able to guarantee the overall performance after a set of minutiae points had been pruned from the original template.

Existing studies of minutiae selection is generally implemented with an image-based quality value for each minutiae, and low quality points are pruned at first. The quality-based selection is followed by distance-based approach with respect to the image center if the remaining points still exceed the maximum minutiae number [107]. Vibert *et al.* [108] proposed several NIMS approaches to perform blind selection with non-compact templates. The kmeans and truncation proposed in [108] are used as the reference. The former is implemented with the Fuzzy c-means [109] algorithm clustering the minutiae of one template into several groups and the points are pruned in terms of their membership grade with regard to the associated cluster(s). However, this method is easy to undulate due to the c-means algorithm.

This study simply presents one of the NIMS criteria for validating quality metric, which is based on the vertices of the minutiae template [104], details are given in the following.

3.5.2 Vertex criterion

The selection criterion presented in this part is simply the distance between each vertex (Vert) minutia and the centroid of the convex hull (polygon) of the minutiae template. The minutiae template is of the international standard for the compact card application (ISO/IEC 19794-2:2005) [60]. First, the distance is simply calculated between each vertex and the centroid (pink star) of the polygon. We use the centroid of the polygon simply because the quality of an image is unknown and it is not appropriate to use the image center for some samples with only light translation of the foreground (even if the quality is not bad) (Cf. Figure 3.12).

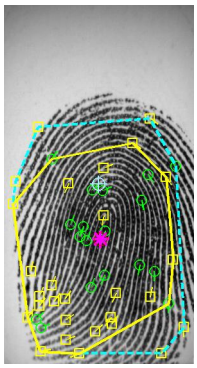


Figure 3.12 – Illustration of the disadvantage of using image center.

In Figure 3.12, one can observe that some minutiae are relatively far from the image center (marked by cross ring) and removing these minutiae can lead to low genuine matching if the translation of another template is tiny or relatively smaller.

For this measurement, with iteration operations, the vertices are pruned according to the largest value of the distances one by one. The desired number of selected minutiae ranges from 30 to 60 increased by 2.

3.5.3 Erollment selection for reduced template

The enrollment selection (ES) is to measure the performance of a quality metric in terms of the decrease of the error rate by choosing samples of relatively good quality. On the other hand, it is more probably to keep more reliable minutiae points for the enrollment sample if the quality of the original sample is better, and hence is beneficial to the matching accuracy of the reduced templates. In this case, the enrollment selection with the quality of the original sample could also be performed for the reduced template.

In this part, concretely, we firstly use the quality metrics calculating the quality value of each original sample, and then perform minutiae selection obtaining the reduced template and their matching scores. This operation generates 11 new datasets, where the template size of each dataset is fixed. After that, the ES operation is performed to each of the reduced dataset, detailed as:

1. The first sample of each individual is chosen as the enrollment to calculate a global EER for the original dataset. This global EER is denoted as 'NoSel'.

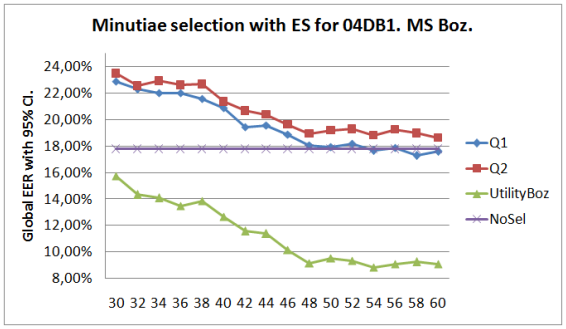
2. The ES operation is performed to each reduced dataset by using quality values of the original dataset. This operation hence computed another group of 11 global EER values for each reduced dataset. These values are able to measure the effect of quality to the NIMS. The effect is hence indicated by the difference between 'NoSel' and these 11 global EERs.

3. The ES is also performed to each reduced dataset by using the **utility** of the original database. This is to obtain a set of reference global EERs, for outlier is an unavoidable problem for the quality metrics.

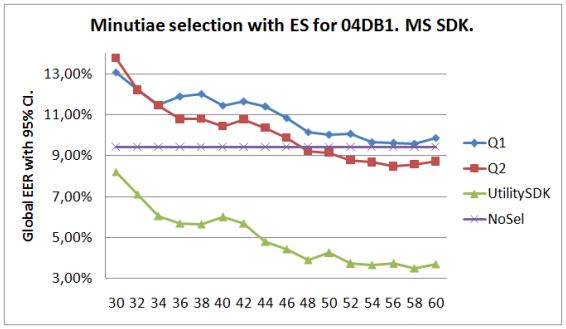
Finally, a quality metric is validated via a comparison between 'NoSel' and global EERs obtained by step 2.

3.5.4 Experimental demonstration

In this section, the ES is performed with NIMS as it has been mentioned in Section 3.5.3. This kind of operation needs to perform on each reduced dataset. Therefore, we simply choose one original database as an example, for which the 04DB1 is used



(a) Result based on MS Boz



(b) Result based on MS SDK

Figure 3.13 – Plots of global EERs obtained by applying the ES to reduced templates. Figure 3.13(a) is the result based on MS Boz and Figure 3.13(b) corresponds to MS SDK.

since there is a dissent between two matchers. In addition, the reduced datasets and their matching scores of 04DB1 are already available in [108]. The matching performance with the original template (NoSel) of this database is far from the utility-based EER (VertUtility), which makes a clear illustration, see plots of the global EERs given in Figure 3.13.

In Figure 3.13(a) and 3.13(b), 'UtilityBoz' and 'UtitlitySDK' respectively indicate the plots of the global EERs obtained via the utility of the 'MS Boz' and 'MS SDK' for each **reduced** template set. Obviously, by comparing the plots of the global EER values based on two metrics, Q1 (blue) and Q2 (red), the result is basically consistent with the one of the original ES operation (Cf. Figure 3.7). A little bit variation appears at over 30 of the desired number (x-axis) when calculating the global EER with 'MS SDK'. This is reasonable according to the study of NIMS [104]. In this case, with the reference quality metric and the objective measure (utility), one can found the difference between two quality metrics and the effect of matching algorithms on qualifying the metrics. The CI resutls of the EER values are not given for this case since most of them are clearly different from each other.

3.6 Conclusion

In order to validate a quality metric of biometric sample, this chapter makes a short review of evaluation approaches for performance assessment of fingerprint quality metric. The reviewed approaches are mostly common frameworks for various biometic modalities. By doing so, this study proposed an technology level validation approach by considering the **sample utility** within an employed matching algorithm. The proposed approach is defined as a criterion in term of the degradation of the global equal error rate (EER) by improving sample quality during the enrollment phase. This is due to the consideration that a valid quality metric is able to choose good quality sample for improving the matching accuracy.

The validation framework defines several global measurements to estimate the validity of a quality metric, including global EER, AUC and AUC ratio. The enrollment selection also provides a 95% confidence interval of each global measurement, by which a comparison between different quality metrics becomes statistically observable.

In addition, this offline framework enables the experiment repeatable, and achieves the purpose for a general application to multiple biometric modalities. Finally, with several experimental results, the effectiveness of the validation scheme is proved in measuring the effect of sample quality to the matching performance.

At last, Section 3.4 gives a quick comparative study between quality metrics carried out in different ways, and what's more, some observations are conducted by asking a few questions:

1) Are those fingerprint quality metrics based-on multi-feature really able to makes the fused metric complementary? and 2) To achieve a common solution, it is necessary to consider whether learning a prior-knowledge of matching performance such as GMS is reasonable or not? For instance, according to the literature [27], it is agnostic that whether two samples produce low impostor score when they are of low quality. Similarly, in Section 2.3.3, it is dubious as well for the genuine matching score between two genuine samples if one of them has an unexpected quality.

The experiment results do somehow answer such questions.

Chapter 4

FQA Combining Blind Image Quality, Texture and Minutiae Features

This chapter presents our first study of FQA by considering several complementary aspects: 1) Image quality assessment without any reference which estimates visual distortions of an image, 2) Textural features related to the fingerprint image and 3) minutiae features which correspond to the most used information for matching. The proposed quality metric is involved in both fusion and prior-knowledge of matching scores. Experiments performed on several trial databases illustrate the benefits of the proposed metric.

Contents

4.1 Introduction

A Biometric system essentially tends to process samples of good quality, for they are beneficial for matching operations and can improve system performance. As one can note in Chapter 3, features are very important to make a reliable judgment of the quality of a fingerprint. For instance, many of existing studies tried to combine multiple features to generate a more reliable metric, such as NFIQ. The general purpose of this chapter is to quantify the quality of digital fingerprint samples and to analyze whether it is possible to take advantages of multiple features for designing a quality metric. In this case, this chapter focuses on a fusion-based quality metric

by using a regression-based approach [84]. The solution in [84] evaluated altered fingerprint image quality with several kinds of quality features, one is a universal no-reference image quality assessment (NR-IQA) algorithm which is an objective measure for inspecting visual distortions of an image [110], and others are related to textural patterns and fingerprint minutiae template. The results presented in [84] demonstrated the effectiveness of this approach for altered fingerprint images. We extended this framework for estimating the quality of original gray-level fingerprints.

The metric given in this chapter is carried out by considering several different aspects: 1) the fingerprint image itself and 2) the associated minutiae template which is rarely taken into account in existing studies. The quality metric is hence implemented via a linear combination of these quality features. The validation of the proposed quality metric is performed by using the evaluation approach defined in Chapter 3. By doing so, an experimental study can be figured out to answer the questions addressed in Section 3.6. The remaining of this chapter is organized as follows: Section 4.2, 4.3 and 4.4 present the details of features used in the computation of the proposed quality metric. Section 4.5 addresses the proposed quality metric. Experimental results are given in 4.6. Section 4.7 concludes this chapter.

4.2 Features for alterated images

The quality metric in [84] assesses the altered fingerprint image which involves in several kinds of alterations: Gaussian noise, contrast, luminance, median blurring, rotation, scratches and occlusion. There are 11 features have been used for assessing the quality of altered fingerprint image, including one No-Reference Image Quality Assessment (NR-IQA) algorithm for estimating image distortions and some other image-based features. A general description of features in [84] is given in Table 4.1.

4.2.1 No reference image quality assessment

In image (or video) quality assessment (IQA), the ”quality” is subject to a wide range of distortions caused by acquisition, processing, compression and any of which may lead to a degradation of visual quality [111]. The IQA can be categorized as subjective evaluation by human expert and objective assessment via an algorithm. Existing studies in this field mainly focus on objective assessing approaches, for subjective evaluation relies on sophisticated tests given by human subjects [112].

Table 4.1: List of quality features in [84]

Feature	Description	NO.
NR-IQA	An image quality assessment approach called BLIINDS [113]	1-N1
SIFT point number	Number of SIFT keypoints	2-S1
SIFT DC coefficient	DC coefficient of SIFT features	3-S2
SIFT Mean	Mean of the scales related to SIFT keypoints	4-S3
SIFT STD	Standard deviation of the scales related to SIFT keypoints	5-S4
Block number	Number of blocks (17×17)	6-P1
Patch RMS Mean	Mean of blocks RMS[1] values.	7-P2
Patch RMS STD	Standard deviation of RMSs	8-P3
Patch RMS Median	Median of blocks RMSs.	9-P4
Patch RMS skewness	Skewness of blocks RMSs.	10-P5
Patch RMS kurtosis	Kurtosis of blocks RMSs.	11-P6

1. 'RMS' is the abbreviation of Root Mean Square.

The objective approaches are generally performed in 3 ways: 1) full-reference (FR) approaches rely on a complete reference image; 2) no-reference (NR) or "blind" assessment meaning no reference image is available and 3) reduced-reference (RR) assessing approaches which means reference image is partially available [114]. The full-reference quality assessment requires a comparison between the distorted image and an original image with no distortions. Among the existing approaches of 2D image, the Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE) and the Feature Similarity (FSIM) are known as the state-of-the-art approaches which show relatively good performance in taking advantages of human visual system (HVS) characteristics [115]. Also, a lot of work have been done in assessing image or video quality with RR-IQA approaches [116, 117]. However, in most cases, NR-IQA is required as the original image is not available, and so does biometric image quality. In this case, a NR-IQA algorithms is employed in the proposed quality framework, with which we generate a feature for distortions of fingerprint image.

The employed NR-IQA is known as BLIINDS [113] which is classified into non-distortion specific approaches, and it uses no distortion model and a different set of sample DCT statistics. This blind IQA method generally involves 3 kinds of DCT-based features: 1) DCT-based contrast feature v_1, 2) DCT-based structure feature v_2,

and 3) DCT-based anisotropy orientation features v_3 and v_4; v_1 is an average value of the contrast of k^{th} DCT path of an image, v_2 is a global image kurtosis based on the kurtosis value of each DCT patch, v_3 and v_4 are variance and max value of the mean value of each DCT patch's Renyi entropy in different orientations. A global quality score called BLIINDS is calculated by using a multi-scale approach as:

$$BLIINDS = \prod_{i}^{L} v_1^{\alpha_1^i} v_2^{\alpha_2^i} v_3^{\alpha_3^i} v_4^{\alpha_4^i} \ , \tag{4.1}$$

where α_j^i are calculated by using the correlation of v_i with the subjective notes given by human observers.

4.2.2 Salient feature and patch-based features

Salient features in Table 4.1 are extracted by using Scale Invariant Feature Transform (SIFT) [118, 119] operator. The SIFT algorithm generates keypoints from an image for object or scene matching. The most significant property of the SIFT is invariant to image scale in addition to rotation-invariant and other advantages such as the robustness to local geometric distortions. The scale invariant is achieved by generating a scale space via

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \tag{4.2}$$

where G is the Gaussian kernel, σ is the scale parameter of the Gaussian kernel and I is the image involved in this convolution $(*)$ operation. The scale space enables the SIFT to extract reliable features from images of different scales and even distorted images, and hence ensures the efficiency of matching.

In [84], the SIFT is used for generating a 128-descriptor of the detected keypoints of a fingerprint image. The number of detected keypoints (NB), DC coefficient of each descriptor matrix, mean values and standard deviation of scales associated to the keypoints are used as features.

For patched features, it firstly divides images into blocks of 17×17, and then the root mean square (RMS) value of each block is computed to obtain the statistic measures (see Table 4.1) of the patch.

4.3 Texture features

This section presents a set of image-based features which are classified into 4 classes in this study, including local binary pattern (LBP) based features [120, 121], Haralick features [122], Gabor wavelet features [123] and local relational string (LRS) feature

[124]. LBP-based features include the original LBP and its five transformations; Haralick features employed in this study involves eight default statistic measures; Gabor wavelet features are generated by four sets of Gabor filters in different scales and orientations which were derived form a filter bank of 16 scales and 16 orientations. Table 4.2 shows a preliminary description of the selected 11 image features.

Table 4.2: Texture features.

Feature	Format	NO.
LBP	256-level LBP histogram vector	1-C1
Four-patch LBP	Descriptor code vector	2-C1
Completed LBP	512-bit 3D joint histogram vector	3-C1
GLCM measurements	8-bit GLCM vector	4-C2
LBP histogram FT	LBP histogram Fourier transform vector	5-C1
2-scale 16-orientation Gabor	64-bit Gabor response vector	6-C3
4-scale 16-orientation Gabor	128-bit Gabor response vector	7-C3
8-scale 16-orientation Gabor	256-bit Gabor response vector	8-C3
16-scale 16-orientation Gabor	512-bit Gabor response vector	9-C3
Local relational string (LRS)	81-bit LRS motif histogram vector	10-C4
Median LBP	256-level MBP histogram	11-C1

These features are widely used for texture classification and image retrieval applications [125, 121] due to good performance in texture representation. This study aims to use them for fingerprint image quality assessment. To generate a combined quality metric, a single measure of each kind of feature is calculated. A brief review of each feature is presented as follows.

4.3.1 LBP-based features (C1)

The local binary pattern feature is proposed by Ojala et al [120] for texture analysis. This feature is simple yet efficient so that it is widely used for relevant issues [126]. The LBP operator is based on the idea that the two-dimensional surface textures can be described by two complementary measures: local spatial pattern and gray scale contrast [127]. The original LBP is a light-weight operator, which generates a binary string by thresholding each 3-by-3 neighborhood of every pixel of the image, and the threshold is the central value itself. The basic operator also extended by many improvements such as uniformed LBP. Few of existing studies have made some

observations in using this feature for fingerprint applications, such as segmentation and matching [128, 129, 130]. The first class of features for generating a new quality metric is hence the basic form of LBP and several of its transformations.

4.3.1.1 LBP histogram (1-C1)

The LBP is a kind of local texture descriptor of an image [131], which reflects the relationship between pixels within a local window. The description of LBP is shown by Figure 4.1.

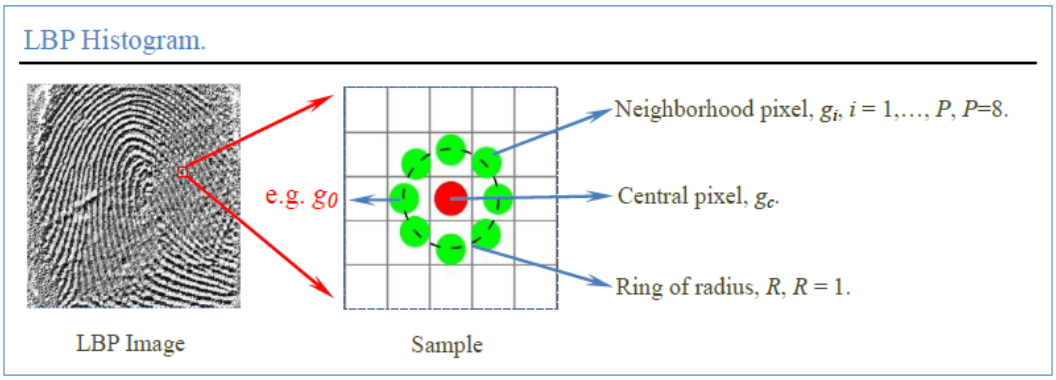


Figure 4.1 – Illustration of local binary pattern (LBP).

The basic form of the LBP calculates the relationship between the central value and its 8 neighborhood pixels along the ring of radius, R (R is 1 for basic LBP), for each pixel in the image. It follows the ring in counter-clockwise and hence g_0 (in Figure 4.1) is the first neighborhood pixel. Assume there is one pixel, $g_c = 54$, and its 8 neighbor pixels in the image, $g_p = [54\ 57\ 12\ 13\ 86\ 21\ 99\ 85]$. In a general form, the relationship between the central pixel and one neighbor pixel is represented by a binary bit,

$$a_i = \begin{cases} 1, & g_c - g_p < 0, \\ 0, & g_c - g_p \geq 0, \end{cases}$$

where $p = 0, \cdots, P-1$ and P is the binary pattern number defined as the Equation below.

The $LBP_{8,1}$ of g_c can be obtained by

$$LBP_{P,R} = \sum_{i=0}^{P-1} a_i 2^i = 1{\times}2^0{+}1{\times}2^1{+}1{\times}2^3{+}1{\times}2^6{+}1{\times}2^7 = 203,\ where\ P = 8,\ R = 1. \tag{4.3}$$

In this study, the LBP operator is applied to the fingerprint image in a global level, illustrated by Figure 4.2.

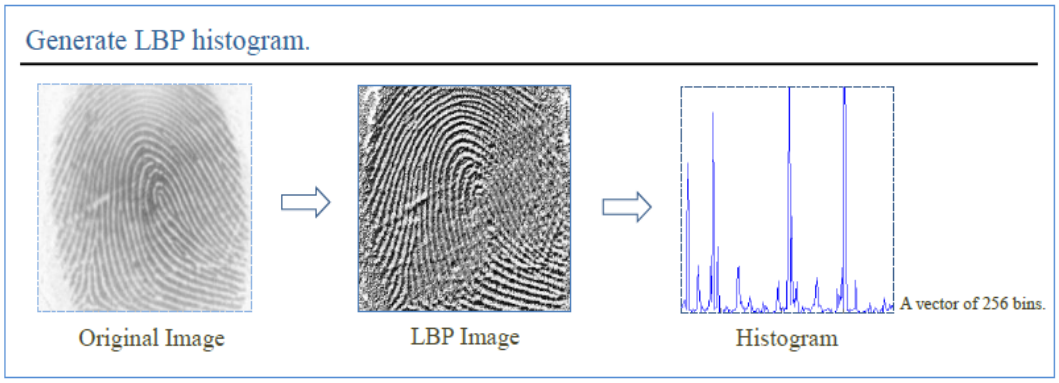


Figure 4.2 – Procedure of calculating LBP feature.

As it is shown in Figure 4.2, the LBP value of each pixel of the original image was calculated first, and then the LBP histogram can be obtained via a statistic of each LBP value. The LBP image in Figure 4.2 is constructed by the LBP value of each pixel.

4.3.1.2 Four-patch LBP (2-C1)

Four-patch LBP is one of the extensions of the LBP [132] operator. The calculation of this transform involves in several parameters, including radius of two different rings around a central pixel, the size of patch centered at a pixel on the ring, patch number and patch interval (unit in patch). An example for calculating FLBP of a pixel is illustrated by the description in Figure 4.3.

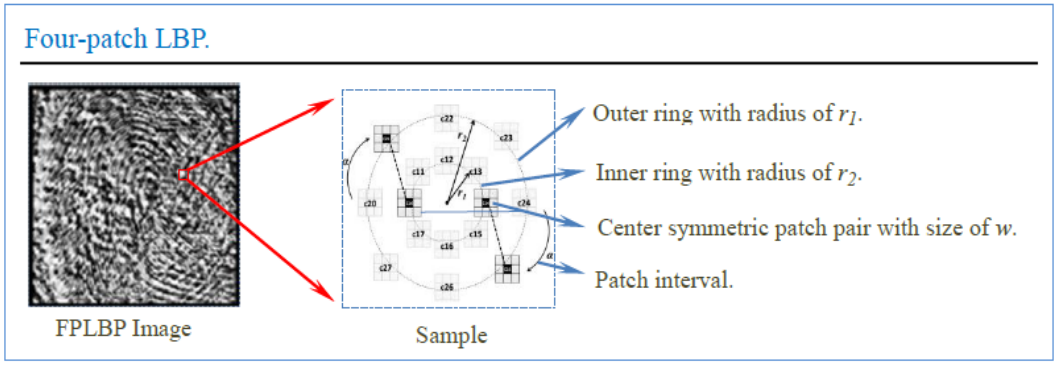


Figure 4.3 – Illustration of four-patch LBP (FPLBP).

Figure 4.3 demonstrates that the FPLBP operator follows patches evenly lied on two rings with the origin at one pixel. It chooses one patch on the outer ring and another patch which is one interval away on the inner ring as a patch pair, and

then another pair that is center symmetric to the selected one can be obtained to calculate the FPLBP code of the central pixel. In this example, α indicates patch distance, S represents patch number, ω determines patch size, and two radius are represented by r_1 and r_2, respectively.

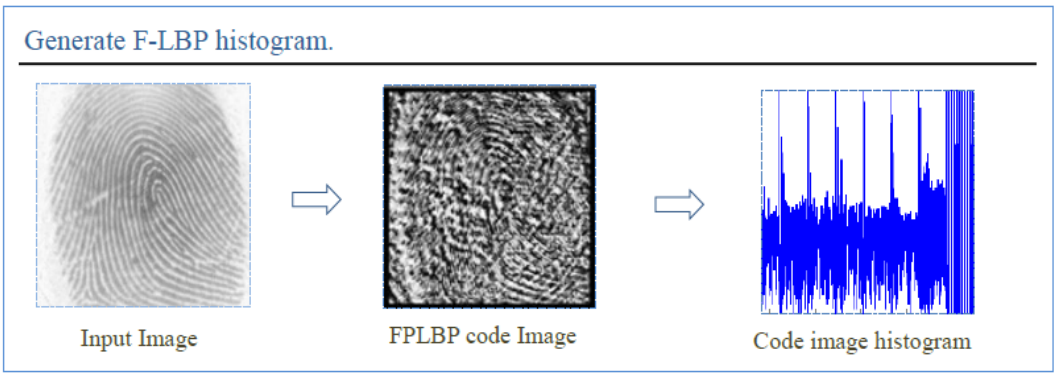


Figure 4.4 – Procedure of calculating FPLBP feature.

In this study, the FPLBP feature is represented by a code image histogram which is computed in terms of the procedure given in Figure 4.4. Each pixel of the code image (Cf. Figure 4.4) is the FPLBP value of the corresponding pixel in original image. The FPLBP value is calculated by:

$$FPLBP_{r_1,r_2,S,\omega,\alpha}(p) = \sum_{i}^{s/2} f(\, d(C_{1i}, C_{2,i+\alpha \bmod S}) - d(C_{1,i+S/2}, C_{2,i+S/2+\alpha \bmod S})\,)2^i, \tag{4.4}$$

where $C_{i,j}$ indicates a patch, $d(\cdot,\cdot)$ is any distance function between two patches and f is given by:

$$f(x) = \begin{cases} 1, & if\ x \geq \tau \\ 0, & if\ x < \tau \end{cases}. \tag{4.5}$$

4.3.1.3 Completed LBP (3-C1)

The completed LBP operator is also generated from the LBP for texture classification [133]. This operator defines 3 texture components: local difference sign component (S), magnitude of local difference (M) and center gray level (C). The sign component (S) is equivalent to the LBP operator which has been proved that it preserves more information of local difference than the magnitude component. These three components are illustrated in Figure 4.5.

I Description

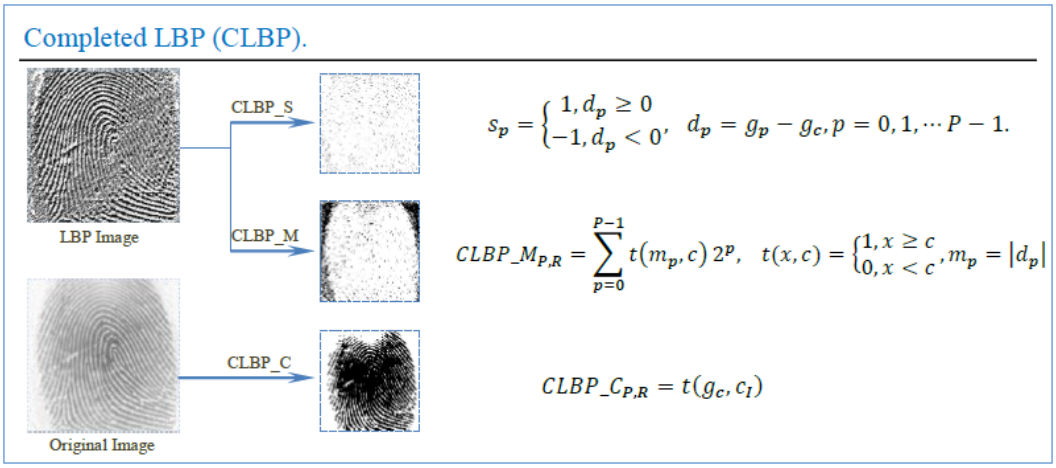


Figure 4.5 – Illustration of completed LBP (CLBP).

As it is depicted by Figure 4.5, d_p here is the difference between the central pixel g_c and one of its neighbor pixels g_p, s_p is the sign component where -1 is coded as 0, and m_p is the magnitude component. Besides, the CLBP_M code is determined by a threshold c which is the mean value of m_p in default. The central gray level component is also determined by the threshold function $t(\cdot, \cdot)$, and c_I is the threshold which is set as the average gray level of the image.

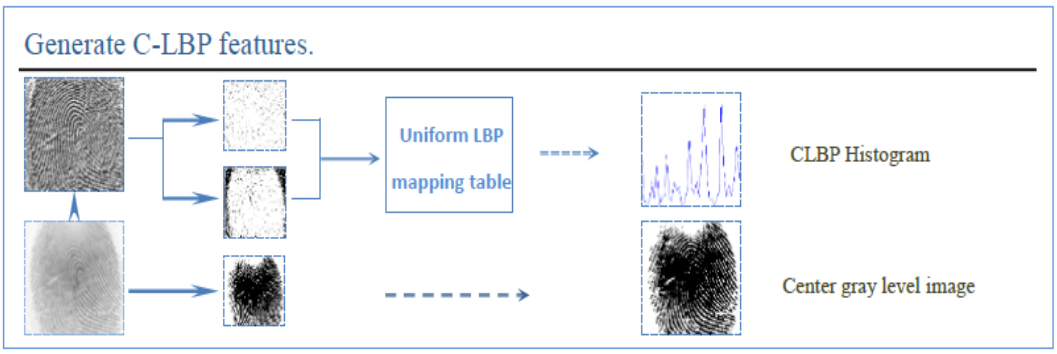


Figure 4.6 – Procedure of calculating CLBP feature.

The CLBP operator has several forms of fusions of the 3 components, such as joint 2-D histogram and concatenated histogram. Figure 4.6 illustrates the calculation of an example for generating a 2D joint histogram or a concatenated histogram of CLBP_S and CLBP_M, and a central gray level component. In this study, we use the 2D joint histogram for calculating the feature.

4.3.1.4 LBP Histogram Fourier Transform (LBPHF, 5-C1)

The LBP-HF [134] is a rotation invariant image descriptor based on the Fourier transform of the ULBP histogram of the original image. This operator considers complex conjugate of the ULBP histogram Fourier transform to deal with the cyclic shift in ULBP histogram caused by image rotation. An illustration concerning its computation is given in Figure 4.7.

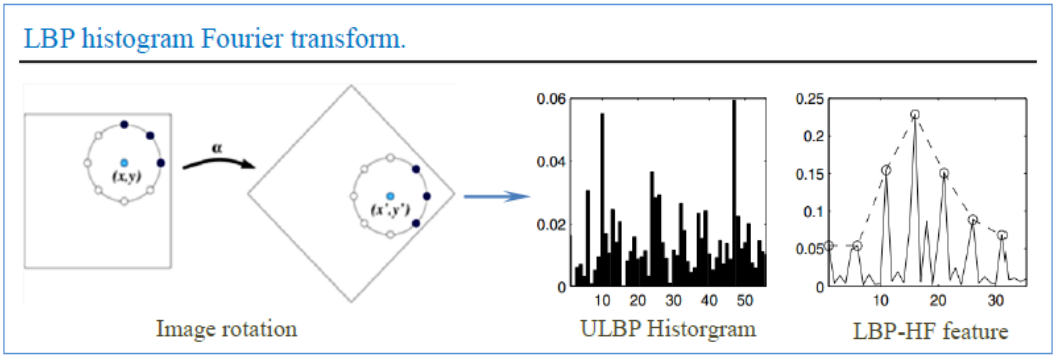


Figure 4.7 – Illustration of LBP Fourier transform (LBPFT). (Image source: [134])

In Figure 4.7, because of the image rotation, the uniform pattern of pixel (x, y) also rotated an angle of $\alpha = a\frac{360}{P}$ to a new location (x', y'). In [134], the Discrete Fourier Transform (DFT) of the uniform LBP histogram $h_I(U_P(n, r))$ of a P sampling points pattern is defined by

$$H\left(n,\ u\right) \ = \ \sum_{r=0}^{P-1} h_I\left(U_P\left(n,\ r\right)\right) e^{-2\pi ur/P}, \tag{4.6}$$

where n is the number of bit '1' in each ULBP pattern, and r is the rotation number.

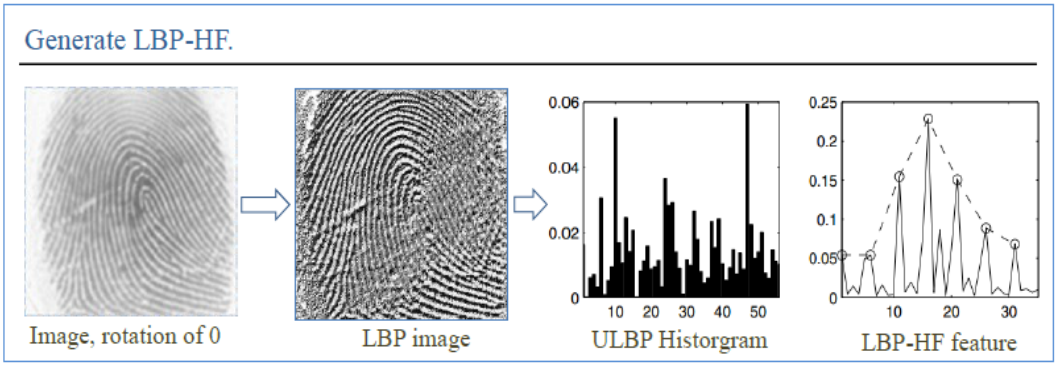


Figure 4.8 – Procedure of calculating LBPFT feature.

In [134], one feature can be represented by the magnitude spectrum of the $H(n,\ u)$, as

$$|H(n,\ u)| = \sqrt{H(n,\ u)\,\overline{H(n,\ u)}}, \tag{4.7}$$

where $\overline{H(n,\ u)}$ denotes the complex conjugate of H. The processing of this feature is illustrated by Figure 4.8, in which a LBP image is firstly derived from the original image. Then, a histogram with the '$u2$' constraint is obtained, and the feature can be calculated by using Equation 4.6.

4.3.1.5 Median LBP (11-C1)

Median LBP (MLBP) [135] is a transformation of LBP operator and invariant to monotonic change of gray-scale. Instead of using the central pixel, this operator generates the LBP code by thresholding each pixel in the 3×3 local region with the median value of the local pixels, as shown by Figure 4.9.

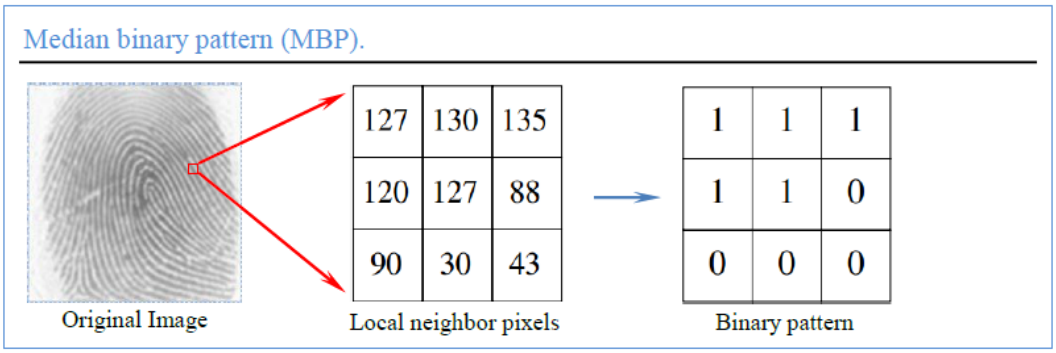


Figure 4.9 – Illustration of median LBP (MLBP).

After the binary patterns had been obtained, the MLBP code can be figured out via a similar way as the Equation 4.3.

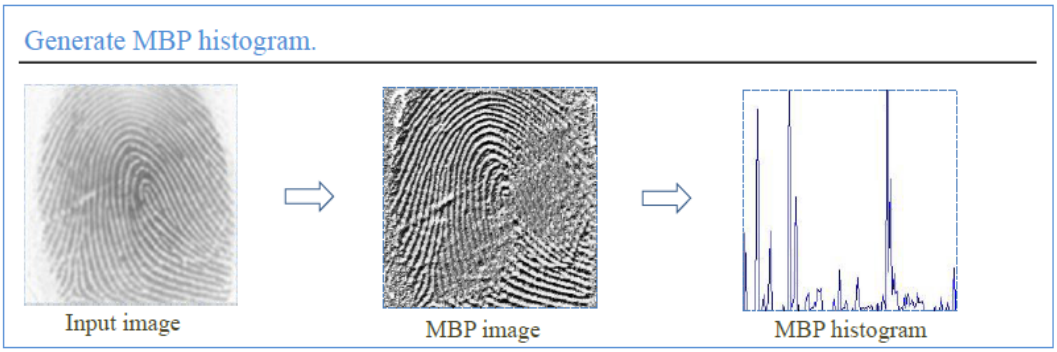


Figure 4.10 – Procedure of calculating MLBP feature.

Figure 4.10 illustrates the calculation of the MLBP feature, in which the MLBP image is firstly obtained by the operator, and then the histogram can be computed. Likewise, the parameter of the operator can be manually specified as well as other extensions of the LBP.

4.3.2 Haralick features (4-C2)

The second class of features in this study is the Haralick feature which is also known as co-occurrence matrix [122]. The co-occurrence matrix of an image reflects the distribution of pixel intensity for either the same intensities or closed intensities. In addition, GLCM also describes the magnitude of the variation of pixel intensities in one or several directions. This kind of texture is not rotation invariant so that the GLCM is calculated in the range of 0 to 180 degrees in general, and directions at 0, 45, 90, 135 degrees are used most. The relationship between pixels of the GLCM matrix is illustrated in Figure 4.11.

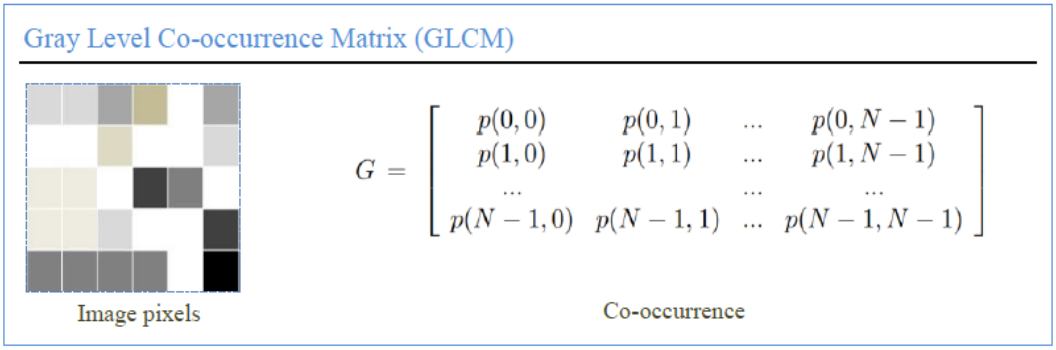


Figure 4.11 – Illustration of Haralick feature.

The element of the gray level co-occurrence matrix in Figure 4.11 is calculated by

$$p\,(i,\ j)\ =\ \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}}, \tag{4.8}$$

where $V_{i,j}$ is the number of occurrence of a pair of pixels, and $p_{i,j}$ is the probability of occurrence of the pixel pair.

Figure 4.12 presents the calculation of several statistic measures generated from the GLCM matrix, which involves in combinations of neighbor pixels in 4 directions that have just been mentioned. In the experiment, we only the default statistic measures given in the reference article.

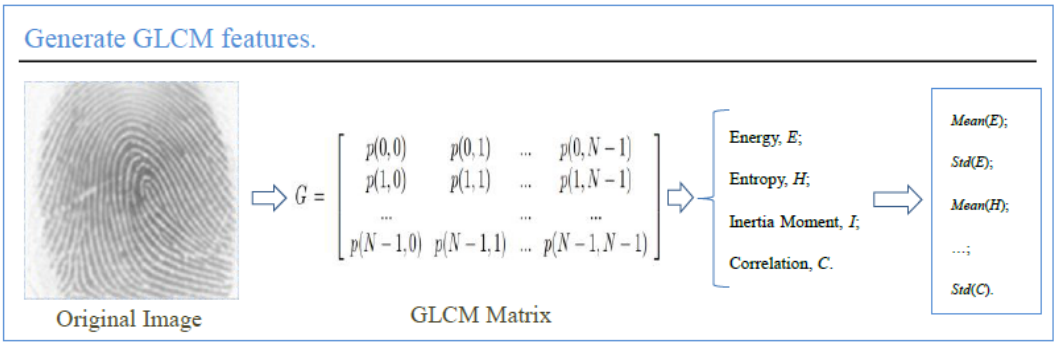


Figure 4.12 – Procedure of calculating Haralick features.

4.3.3 Gabor features (6 ∼ 9-C3)

Likewise, as it has been presented in Chapter 2, Gabor features are also employed in this study. The Gabor filter has distinctive localization properties in both spatial and frequency domain, and is said to be similar to the characteristics of certain cells in the visual cortex of mammals. Because of this, it has been widely used for image processing and analysis such as classification, enhancement and so on. Multi-orientation and multi-scale capability enables Gabor filters to process image information in both local and global level. In this case, 2D Gabor [136] filters perform relatively well in dealing with fingerprint pattern. The 2D Gabor function is a sinusoidal function modulated by a Gaussian window. In the wavelet domain, the basis of Gabor function is hence complete but not orthogonal. A decomposition of a 2D Gabor filter is given in Figure 4.13.

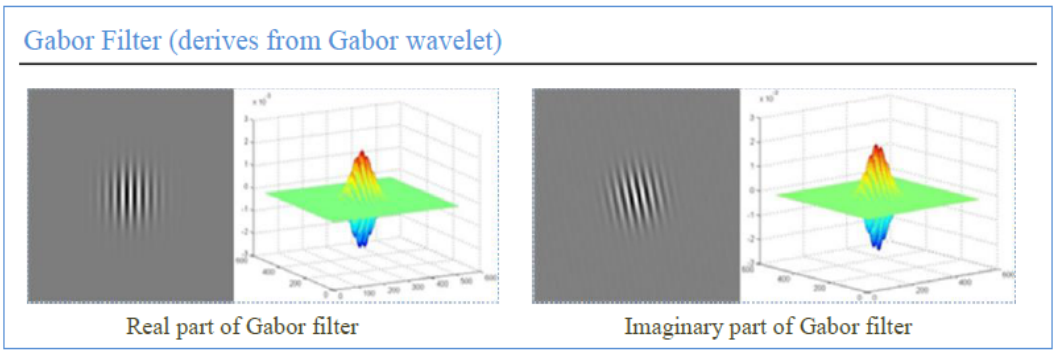


Figure 4.13 – Illustration of Gabor response.

Figure 4.13 contains an example of both imaginary part and real part of the Gabor response of an image generated by one of the Gabor filter bank. A general

form of 2D Gabor function is formulated as

$$h\left(x, y, \theta_k, f, \sigma_x, \sigma_y\right) = exp\left[-\frac{1}{2}\left(\frac{{x_{\theta_k}}^2}{{\sigma_x}^2} + \frac{{y_{\theta_k}}^2}{{\sigma_y}^2}\right)\right] \times exp\left(j2\pi f x_{\theta_k}\right), \tag{4.9}$$

where

$$\begin{bmatrix} x_{\theta_k} \\ y_{\theta_k} \end{bmatrix} = \begin{bmatrix} cos\theta_k & sin\theta_k \\ -sin\theta_k & cos\theta_k \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{4.10}$$

f is the frequency of the sinusoidal plane wave along orientation θ_k, n is the number of orientations, $k = 1 \cdots n$, σ_x and σ_y are the standard deviations of the Gaussian envelope along the x and y axes, respectively.

To generate Gabor features for the proposed metric, we calculate a set of Gabor filters with a Gabor wavelet [136], illustrated by Figure 4.14.

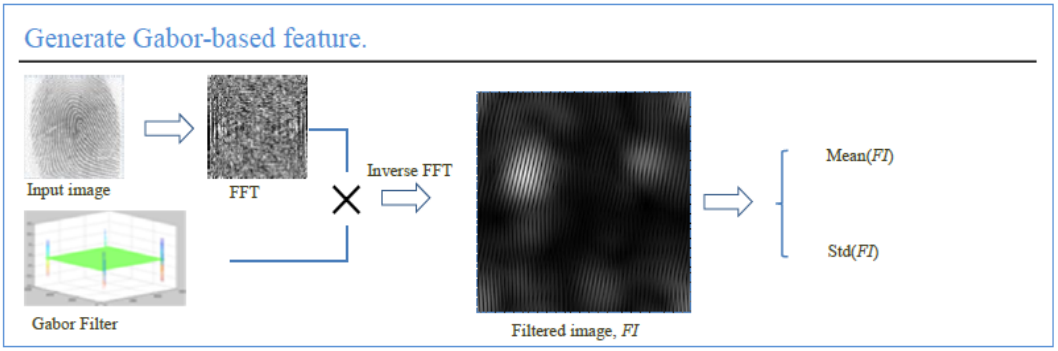


Figure 4.14 – Procedure of calculating Gabor features.

In Figure 4.14, the Gabor filter bank is generated by using the Gabor wavelet in N scales and M orientations. Each Gabor response is obtained by one of the filters in frequency domain multiplying the Fourier transform (FT) of the input image, and then two statistic measures (mean and standard deviation) of the inverse FT of the product are used as the features.

4.3.4 Local Relational String (11-C4)

Another feature applied to the proposed metric is the local relational string (LRS) [124] which is an illumination invariant operator focusing on local variation of the gray level of an image. This operator is based on the local pixels relation in a specified scale, see the left-most graph in Figure 4.15.

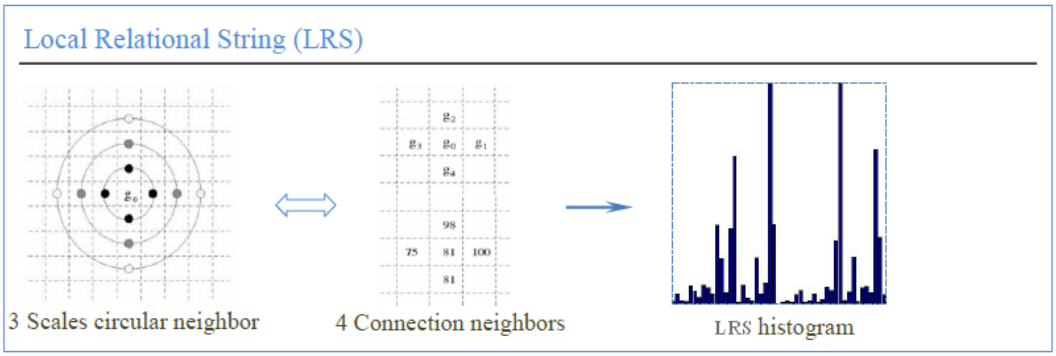


Figure 4.15 – Illustration of local relational string (LRS) feature.

In Figure 4.15, to represent the relational feature, it uses 3 relations to generate local relation string histogram for measuring the local spatial variations. Instead of calculating gray level difference, the LRS operator considers the relation between central pixel and its neighbor pixels, which is represented by an ordered symbolic string (LRS) derived from a linguistic symbol set,

$$LRS : r_1\ r_2\ r_3\ r_4. \tag{4.11}$$

In the string given by expression 4.11, each r_i represents the relation between a central pixel (g_0) and its i^{th} neighbor (g_i). The relation, Z, is defined by

$$S\ =\ \{(g_0, g_i) \in I \mid \exists r_i \in R, r_i\ =\ Z\,(g_0, g_i)\};\ g_i\ \in\ \Omega, \tag{4.12}$$

where $\omega = \{g_1, g_2, g_3, g_4\}\ and\ R = \{<, >, =\}$.

The calculation of the motif histogram is demonstrated in Figure 4.16.

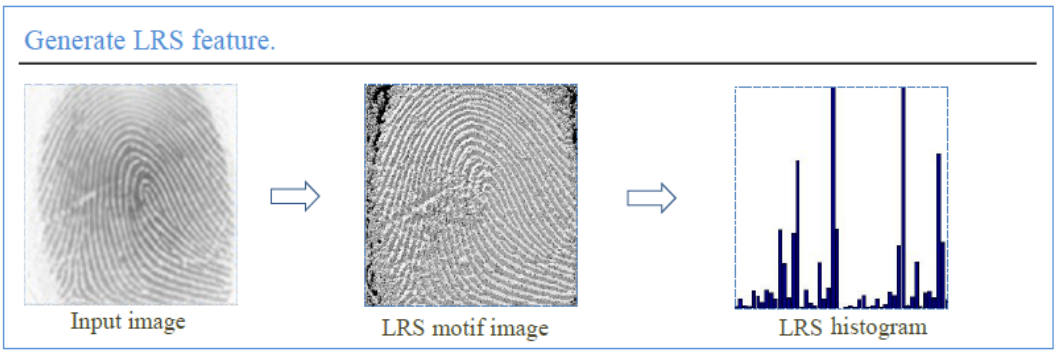


Figure 4.16 – Procedure of calculating LRS feature.

In Figure 4.16, the LRS motif image is composed by the motif value of each pixel of the original image. The feature is represented via a histogram in 1 dimension or multiple dimension which is determined by the spatial scale.

The texture features given above are all extracted from fingerprint image at a global level. Each kind of extracted feature vector was analyzed by using principal component analysis (PCA) [137]. The length of the selected features are different from each other and the redundancy either. However, to combine all the selected features, it is necessary to get a single index from each feature vector. In this case, we choose a statistic measure for each feature vector and the redundancies of them are ignored. The statistic measurement for the features given in Table 4.2 are interquartile range (iqr), standard deviation, the sum of the two maximum values, skewness, entropy, skewness, entropy, iqr, mad, mad, mad and entropy.

Note that we do not assert that these statistic indexes are quality features in this study. In practical, we calculate several statistic measures for each of the employed features (vectors). Next, we calculate the Pearson correlation between each statistic measure and the genuine matching score. At last, these single index are simply determined in terms of their Pearson correlation obtained from different datasets, *i.e.* statistic measure whose Pearson correlation seems commonly stable are selected to represent the feature. The analysis of Pearson correlation for the defined single indexes is given in other sections of this chapter.

4.4 Minutiae-based features

In addition to image-based features, we also make effort in quality assessment with fingerprint minutiae template, for we mainly use minutiae-based matching approach. Hence, the minutiae information somehow reflects the utility [29] of fingerprint. The minutiae is categorized as the level-2 feature of fingerprint pattern, which are represented as a set of spatial points with an orientation property. Basically, a minutia point is denoted as a triplet representation, $m_i = \{x_i, y_i, o_i\}$, where (x_i, y_i) is the spatial position and o_i is the orientation. An illustration of a fingerprint

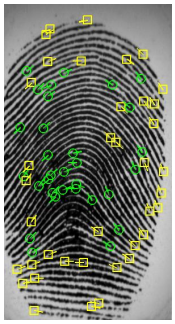


Figure 4.17 – Illustration of detected minutiae of a fingerprint image.

and its minutiae is given in Figure 4.17. In addition to location and orientation of minutiae point, the triplet representation does not carry any other information about fingerprint image. This is sufficient to perform matching task for a recognition system, and several studies have already proved that it is possible to estimate orientation map and reconstruct a fingerprint image [138, 139] by using the triplet representation and it succeeded in performing type-I and type-II attacks against a recognition system.

Table 4.3: Minutiae number-based measures related to fingerprint quality.

Measure	Description		NO.
Minutiae number	N_i	N_i, minuitiae number of the i^{th} fingerprint.	$1-M$
Mean based on FT of minutiae	$mean(MFT_i)$	MFT_i, the magnitude of the Fourier transform of minutia point's 3 components	$2-M$
Standard deviation of minutiae	$std(MFT_i)$	Standard deviation of minutiae magnitude	$3-M$
Minutiae number in ROI[1] 1	NR_i	NR_i, minutiae number in a rectangle region.	$4-M$
Minutiae number in ROI 2	NC_i	NC_i, minutiae number in a circle region.	$5-M$
Region-based RMS	$rms = \sqrt{\frac{1}{n}\sum_{i=1}^{n} m_i{}^2}$	Root mean square (RMS) value of minutiae number based on two blocks of the template along its vertical direction.	$6-M$
Region-based median	$med = \frac{1}{2}^{th} sort(m)$	Median value of minutiae number obtained by dividing the template into 4 blocks.	$7-M$
Block-based measure	A block-based quality score is calculated based on the minutiae number in each divided block of the template, the block size is score 64-by-64 here.		$14-M$

1. region of interest.

In this case, to generate the quality metric, one need to find out some features from the minutiae template that are able to reflect quality in one or more aspects. According to the literature [140, 141], fingerprint minutiae cannot be accurately

described by a simple statistical model, which makes difficulty in finding regularities of this kind of feature. However, some studies [142, 143] choose to use minutiae features in a local region, for it is invariant to both the translation and rotation in a small region. In this case, minutiae in a local area or a region of interest (ROI) may have different impacts on matching performance and hence affects quality. Tabassi *et al.* have proposed to use minutiae number in their FQA algorithm. In addition, few of existing studies have considered minutiae information for quality assessment. In this study, we simply define some features of minutiae template for the proposed quality metric. The proposed features are based on minutiae number and the Discrete Fourier Transform (DFT) of the three components of minutiae points. An overall description of these features is given in Table 4.3.

Minutiae-based features given in Table 4.3 are calculated from the template extracted by using NBIS software [68]. This type of template contains a quadruple representation of minutia point which is consisted of the location of detected minutiae, (x, y), the orientation of detected minutiae, θ, and a quality score of detected minutiae. In the experiment, the minutiae positions and orientations are used only for calculating these features. In the following, the definition of these features are presented in detail. In this study, features 2 and 3 are derived from the DFT magnitude of the linear combination of the three minutia components after eliminating DC component, as formulated in Equation 4.13 and 4.14.

$$F_i\left(x, y, \theta\right) = \sum_{n=0}^{N-1} x_n \cdot \mu^{kn} + y_n \cdot \nu^{kn} + \theta \cdot \omega^{kn}, \tag{4.13}$$

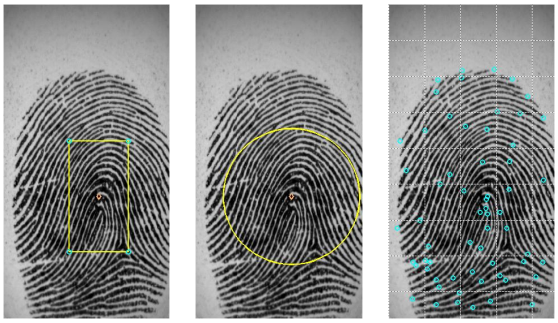
where μ, ν, and ω are frequency samples.

$$F_{2-M} = \overline{\left|F_i\left(x, y, \theta\right)\right|}, \tag{4.14a}$$

$$F_{3-M} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(|F_i| - F_{2-M}\right)}. \tag{4.14b}$$

DC component was eliminated when calculating these two measures because there is no valuable information in this element.

The size of rectangle region of feature 4 is determined by the maximum value of both x and y coordinates of the involved minutiae, for there is no useful information outside the foreground of the fingerprint in this case. This choice also ensures that the region of interest will not go over the effective area of minutiae. An example of rectangle region is shown by Figure 4.18(a).



(a) Rectangle (b) Circular (c) Blocks

Figure 4.18 – Example of region of interest (ROI). Figure 4.18(a) is rectangle region, 4.18(b) is circular region and 4.18(c) is grid division.

The radius of the circle region for feature 5 is also determined by the maximum and minimum location value along the horizontal direction of the fingerprint, for the minutiae located around fingerprint center are said to be more significant to fingerprint matching, *i.e.* they are more informative. As the quadruple representation does not provide information of fingerprint core point, an estimated point was used as the center location of the fingerprint. In the experiments, a comparison is made between the estimated center point and a core point detected by another approach, and it is found that the result does not vary too much. The estimated center position is determined by considering the maximum and minimum minutiae location as well. An example of the circular region is shown by Figure 4.18(b).

For feature 6 and 7, the whole fingerprint region is respectively divided into 2 and 4 blocks, and minutiae number in each block is considered to generate a measure. Another block-based measure is calculated by dividing the whole fingerprint region into 64×64 pixel blocks. An estimated index is assigned to each block in terms of the minutiae number in the block, which involves in a threshold. The block is classified into 3 classes, reasonable block, vague block and unreasonable block. Then, a measure is calculated based on the number of these 3 kinds of blocks. An example of block partition is shown by Figure 4.18(c).

In addition, some rotation and translation invariant features proposed in [139] are calculated in terms of minutiae distribution and orientations. This study adopts the first six features of them to generate the quality metric.

4.5 Proposed quality metric

As it is discussed in previous chapters, the quality of biometric sample is referred to as 3 components [144, 145, 22]: character, fidelity, and utility. In this study, quality

metric is calculated based on the approach proposed in [84] defined in the GREYC research lab. The approach was inspired by the utility of biometric sample quality, i.e. biometric sample quality predicts system matching performance [71, 29, 27]. **Note** that quality predicting matching performance is not an absolute concept throughout of this thesis.

The quality metric in this study is generated by a linear combination of the features that had been presented in Section 4.2, 4.3 and 4.4, as formulated by Equation 4.15.

$$Q = \sum_{i=1}^{N} \alpha_i f_i \tag{4.15}$$

where N is the number of quality features denoted by f_i $(i = 1, \cdots, N)$, and α_i is the weighted coefficient. In our framework, the weighted coefficients are learnt by optimizing a fitness function with a genetic algorithm (GA), which is defined as the Pearson correlation (Equation 4.16) between combined quality results and the corresponding GMS of the training samples. With this definition, the coefficients could be obtained in terms of maximizing the correlation between the genuine matching scores and the quality measures as much as possible [90, 71], and therefore satisfy the utility property. In the experiment, a randomly selected data set of 25% of each database are given to the optimization algorithm calculating the weighted coefficients. By substituting the coefficients of the features into Equation 4.15, the proposed method achieves a continuous quality metric adaptively. The quality scores are normalized on each database.

The fitness function here is defined as a Pearson correlation coefficient between the genuine matching score (GMS) and the quality scores, as in Equation 4.16.

$$\rho_{Q,MS} = \frac{cov\,(Q, MS)}{\sigma_Q \sigma_{MS}}. \tag{4.16}$$

In the experiment, the $N - 1$ GMS of the first sample of each individual is used for this fitness function.

4.6 Experimental results

In this section, the analysis result of each set of features is given by the Pearson correlation coefficients at first, and followed by a validation of the proposed metric with the enrollment selection (ES) approach. Details of both approaches have been mentioned in Chapter 3.

4.6.1 Feature analysis

In Chapter 3, we introduced Pearson correlation between the sample quality and its GMS to observe the validity of a quality metric. In addition, Fernandez *et al.* [71] and Olsen [102] respectively calculated Pearson and Spearman correlation coefficients between different quality metrics to observe their behavior. In this case, as it has been addressed in Chapter 3, this study analyzes the proposed features by computing the Pearson correlation between each of them and several existing quality metrics. Such an analysis is able to observe whether there is a similarity between the proposed features and reference metrics. Quality metrics used for this analysis include OCL [80], orientation flow (OF) [81], standard deviation (STD) [73], Pet Hat's wavelet (PHCWT) [77] and NFIQ. Table 4.4, 4.5, 4.6 presents the correlation results obtained from three of the trial databases where the proposed indexes obtained results as expected. **Note** that we simply give the results of a part of the features due to space limitation.

Table 4.4: Inter-class Pearson correlation for textural features. 02DB2A (top), 04DB1A (middle) and 04DB3A (bottom).

Feature.	LBP	FLBP	CLBP	LBPFT	Gabor128	Gabor256	Gabor512	LRS	MLBP
OCL	-0.6826	0.3002	-0.7037	-0.4462	0.3806	0.5864	0.6832	0.8699	-0.7593
OF	-0.1938	0.1783	0.0098	-0.0452	0.1685	0.2016	0.1590	-0.0012	0.0593
PHC	-0.6926	0.2864	-0.6665	-0.3391	0.3552	0.6329	0.7507	0.8476	-0.7807
STD	-0.6230	0.3958	-0.5590	-0.3016	0.5620	0.8066	0.8940	0.7668	-0.7438
NFIQ	0.3919	0.1240	0.3483	0.1617	0.0401	-0.0676	-0.1307	-0.4569	0.2731
OCL	-0.6899	-0.7979	-0.7798	0.7151	0.4071	0.6708	0.7223	-0.7416	0.7125
OF	-0.2642	-0.3263	-0.3057	0.2073	0.3968	0.4206	0.4539	-0.2281	0.2057
PHC	-0.7060	-0.8206	-0.8416	0.7535	0.4722	0.7548	0.7964	-0.7701	0.7426
STD	-0.5920	-0.7066	-0.7286	0.6471	0.4669	0.6930	0.7264	-0.6646	0.6297
NFIQ	0.1634	0.1607	0.1775	-0.2101	0.0897	-0.0254	-0.0157	0.2295	-0.2143
OCL	-0.5001	-0.6394	-0.7460	0.5144	0.5301	0.6505	0.6948	-0.3814	0.5536
OF	-0.2510	-0.1842	-0.1539	0.0814	-0.1348	-0.1148	-0.0539	-0.1537	0.1566
PHC	-0.1648	-0.2758	-0.4495	0.1439	0.6947	0.7992	0.7450	-0.1928	0.1726
STD	-0.2401	-0.3447	-0.5029	0.2221	0.6398	0.7359	0.7037	-0.2161	0.2550
NFIQ	-0.0532	-0.0886	0.0316	0.0518	-0.3640	-0.4005	-0.2608	-0.0805	0.0907

In Table 4.4, highlighted columns (in yellow) demonstrate a relatively stable correlation for all the three databases, and some others marked with green illustrated their feasibility for certain data sets. According to these observations, one can note that some of the proposed features demonstrate similarity to the reference metrics, and we could make an attempt to reduce some redundant features in next study. Table 4.6 presents only 10 of the minutiae-based features, for the correlation of this type of feature is not very distinctive. Some of them demonstrate good correlated behavior with the reference quality metrics, but greatly vary among the data sets and even not correlated with any of the quality metrics. Likewise, the correlation

Table 4.5: Inter-class Pearson correlation for image-based features. 02DB2A (top), 04DB1A (middle) and 04DB3A (bottom).

Feature	NR-IQA	S.[1] Num	S. DC	S. STD	Block Num	R.[1] Mean	R. Skewness	R. Kurtosis
OCL	0.4816	0.2370	0.2931	0.3659	-0.9137	0.6643	-0.5538	-0.8443
OF	0.0386	-0.0438	0.1038	0.2487	0.0391	0.0875	-0.1092	0.0452
PHCWT	0.4720	0.3650	0.3149	0.4921	-0.7480	0.5860	-0.5316	-0.8469
STD	0.3169	0.2133	0.4788	0.5805	-0.7170	0.6608	-0.5252	-0.8023
NFIQ	0.4434	0.4445	0.1735	0.1164	0.4017	0.2510	0.2598	0.3907
OCL	0.4689	0.0418	0.3839	0.5980	-0.9129	0.8823	-0.3423	0.8753
OF	0.1908	0.0065	0.1216	0.1347	-0.1971	0.1586	0.2351	0.3396
PHC	0.5126	0.2468	0.4225	0.7118	-0.7046	0.6858	-0.2800	0.8687
STD	0.4070	0.2177	0.4946	0.8112	-0.6632	0.6887	-0.2416	0.7591
NFIQ	-0.1890	-0.3808	0.1444	-0.3420	0.0132	-0.0121	0.0069	-0.0719
OCL	0.3414	0.2499	0.2271	0.6927	-0.2067	0.6544	-0.0068	0.7988
OF	-0.0558	-0.0645	-0.1039	0.0883	-0.0079	-0.1368	-0.0017	0.0122
PHC	0.3580	0.4141	0.5300	0.5679	0.2351	0.8933	0.0515	0.6215
STD	0.4175	0.4266	0.4661	0.6575	0.1858	0.9157	0.0575	0.6319
NFIQ	-0.2256	-0.3925	-0.1761	-0.2670	-0.2824	-0.4156	0.0112	-0.1193

1. S. is the abbr. of 'SIFT', R. is the abbr. of 'RMS'.

results of the image-based features are given in Table 4.5. We use all these features to calculate the quality metric, which enables qualifying fingerprint samples with various information and reaches the purpose to answer the questions that had been proposed in Chapter 3.

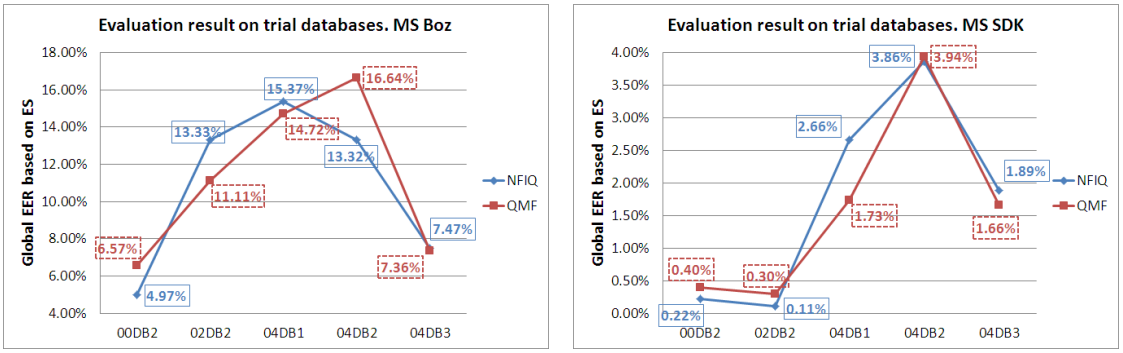
Table 4.6: Inter-class Pearson correlation for minutiae-based features. 02DB2A (top), 04DB1A (middle) and 04DB3A (bottom).

OCL	0.4077	0.3768	0.4040	0.2780	0.0826	0.3166	0.4214	-0.3196	-0.2799	-0.2568
OF	0.0327	0.0391	0.0442	-0.0096	0.0019	-0.0035	0.0491	-0.0040	-0.0987	-0.0521
PHC	0.3717	0.3445	0.3735	0.2306	0.0298	0.2787	0.3829	-0.3230	-0.2704	-0.2791
STD	0.2391	0.2267	0.2376	0.1247	-0.0615	0.1630	0.2490	-0.2389	-0.2027	-0.1832
NFIQ	-0.6052	-0.5393	-0.5949	-0.4783	-0.4639	-0.5807	-0.5554	0.4461	0.3544	0.3198
OCL	0.5576	0.5290	0.5570	0.4649	0.5088	0.4505	0.5536	-0.3599	-0.3677	-0.3986
OF	0.0835	0.0946	0.0859	0.1661	0.0721	0.1334	0.0128	0.0159	0.1975	-0.0372
PHC	0.4036	0.4153	0.4150	0.3462	0.3731	0.3124	0.4184	-0.2908	-0.3245	-0.3121
STD	0.3876	0.4017	0.4003	0.3275	0.3446	0.3149	0.3865	-0.2992	-0.3093	-0.3095
NFIQ	-0.1532	-0.1840	-0.1796	-0.1175	-0.1457	-0.1058	-0.1603	0.1778	0.1040	0.1825
OCL	0.2447	0.2362	0.2521	0.1304	-0.0361	0.2280	0.2630	-0.2231	-0.1557	-0.2140
OF	0.2929	0.2577	0.2724	0.2786	0.3218	0.3043	0.2854	-0.0661	0.1458	-0.1077
PHC	-0.1438	-0.1170	-0.1215	-0.1919	-0.3633	-0.1563	-0.1144	-0.1132	0.0373	0.0406
STD	-0.0421	-0.0243	-0.0220	-0.1007	-0.2618	-0.0561	-0.0130	-0.1491	-0.0271	-0.0281
NFIQ	0.3195	0.2497	0.2741	0.3971	0.4406	0.3391	0.2953	-0.0524	-0.0716	-0.0885

4.6.2 Validation with ES

In this part, the ES is simply performed with NFIQ and the proposed quality metric (denoted as QMF hereinafter) on five datasets. The utility value defined in Chapter

3 is not involved in the experiments. The plots of the global EER values obtained by the ES are given in Figure 4.19.



(a) Result based on MS Boz (b) Result based on MS SDK

Figure 4.19 – Plots of global EERs obtained by using quality-based ES. Figure 4.19(a) is the result based on MS Boz and Figure 4.19(b) corresponds to MS SDK.

In Figure 4.19(a), 'MS Boz' and 'MS SDK' also represent two sets of matching scores as it has been mentioned in Chapter 3. With the graphical result based on 'MS Boz', one can find that QMF performs better than NFIQ in reducing error rate on three databases: 02DB2, 04DB1 and 04DB3. Such results demonstrate that the fusion-based approach achieves a better regression or linearity between the quality and GMS on these 3 datasets, but failed for another two. This is fully dependent on the coefficients obtained by the optimization algorithm. Meanwhile, it relies on the GMS values of the first samples, indicating that coefficients are greatly affected by matching performance, so is the regression result.

This problem happens when the ES is conducted with the MS SDK (Cf. Figure 4.19(b)). The QMF obtained a clearly good result from 04DB1 and a result that is nearly equal to NFIQ from 04DB2. The NFIQ is clearly better than QMF on 02DB2, which is opposite to the result based on 'MS Boz'. One should **note** that both metrics are matching scored-based approaches. Though, the matching algorithm is one point that affects the evaluation result of metric performance, but there would be more uncertainties for such metrics implemented by learning a matching performance. Besides, it also reveals that two metrics do not take the advantages of the employed features for all the datasets, such as 00DB2 for QMF and 04DB1 for NFIQ. These problems could be further revealed via comparisons with metrics rely on a single feature, see later chapters.

The 95% confidence interval (CI) of the global EER values are given in table 4.7.

Table 4.7: The 95% confidence interval of the EERs.

QM et MS / DB	QMF Boz	NFIQ Boz	QMF SDK	NFIQ SDK
00DB2A (CI)	[0.0651 0.0663]	[0.0492 0.0502]	[0.0040 0.0043]	[0.0021 0.0023]
02DB2A (CI)	[0.1104 0.1118]	[0.1326 0.1340]	[0.0029 0.0032]	[0.0011 0.0013]
04DB1A (CI)	[0.1464 0.1480]	[0.1529 0.1545]	[0.0172 0.0178]	[0.0268 0.0276]
04DB2A (CI)	[0.1651 0.1676]	[0.1321 0.1344]	[0.0378 0.0389]	[0.0390 0.0402]
04DB3A (CI)	[0.0730 0.0742]	[0.0741 0.0752]	[0.0162 0.0167]	[0.0190 0.0195]

Furthermore, the AUC ratio are given in table 4.8.

Table 4.8: AUC ratio of each quality metric based on two sets of matching scores.

DB / Metric	00DB2	02DB2	04DB1	04DB2	04DB3
QMF (BOZ)	0.5482	0.8929	0.8096	0.6975	0.7766
NFIQ (BOZ)	0.7506	0.8280	0.7699	0.8242	0.7447
QMF (SDK)	0.8190	0.8152	0.8745	0.7837	0.6490
NFIQ (SDK)	0.9879	0.9910	0.8431	0.7482	0.8218

The AUC ratio given in table 4.8 is basically consistent with the global EER values. However, the AUC and EER are two different global measures, which may sometimes lead to inconsistency between them. This problem happens to 04DB3 when the 'MS SDK' is involved. The AUC (calculated in terms of Section 3.3.2) values of the QMF and NFIQ are 0.0078 and 0.0041, respectively. This result leads to the inconsistency between the AUC ratio and the global EER value. In this case, one can calculate the evaluation result with some auxiliary approaches, such as the 'fraction rejected versus FNMR' defined in Chapter 3.

Figure 4.20 demonstrates the result of the fraction rejected versus FNMR.

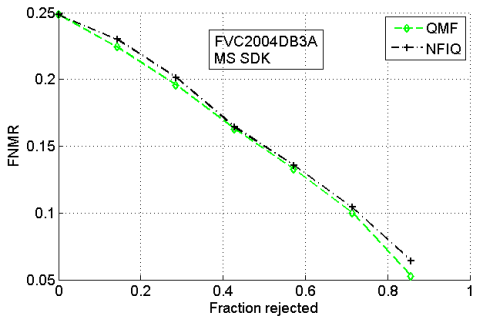


Figure 4.20 – Result of the fraction rejected versus FNMR of the NFIQ and QMF. Database: FVC2004DB3A; Matching score: MS SDK.

This evaluation result is fully dependent on the genuine matchings, while the impostor matching errors are not involved. In this case, this approach is simply used as a complementary evaluation.

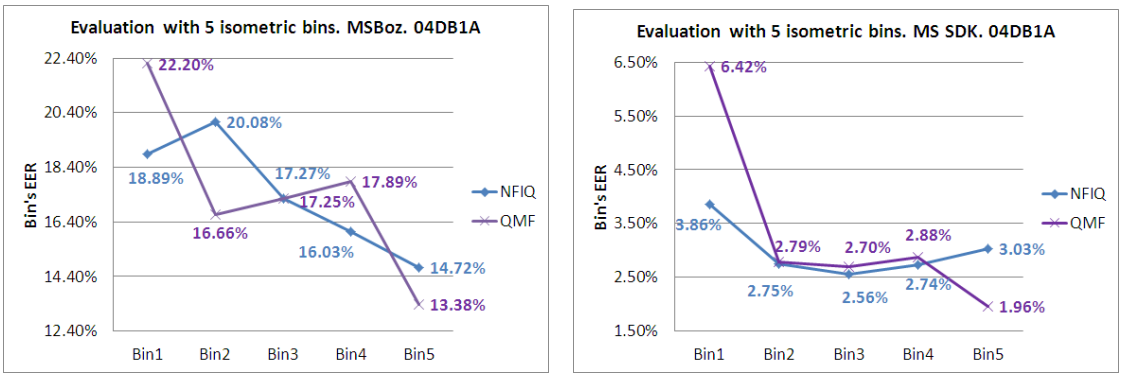
4.6.3 Pearson correlation

According to the analysis given in Section 3.4.2.4, in this part we calculate the Pearson correlation between the maximum GMS and the quality scores. This is simply performed on the datasets where QMF obtain a better result to reveal the linearity between quality metric and genuine matching performance. The Pearson correlation coefficients between each metric and each set of GMS are given in Table 4.9.

Table 4.9: Pearson correlation coefficients between quality score and Max-GMS.

Metric \ DB	00DB2	02DBs	04DB1	04DB2	04DB3
QMF (MS Boz)	-0.0014	0.5217	0.2601	-0.0177	0.5922
NFIQ (MS Boz)	-0.4541	-0.3308	-0.1579	-0.3937	-0.3063
QMF (MS SDK)	-0.0021	0.3254	0.3734	0.0615	0.4142
NFIQ (MS SDK)	-0.4379	-0.2596	-0.1970	-0.5843	-0.4131

The result given in Table 4.9 is basically consistent with the global EERs plotted in Figure 4.19. Both metrics don't show correlation on 04DB1, and we calculate the bin's EER values with the evaluation approach presented in Section 3.4.2.3, see the graphical result in Figure 4.21. In Figure 4.21(a), the square points are results



(a) Result based on MS Boz (b) Result based on MS SDK

Figure 4.21 – Plots of EER values of 5 isometrics obtained from 04DB1 by using Chen's Evaluation with two sets of GMS.

of NFIQ and the cross points are results of QMF. Therefore, one can note in the

figure that bins' EER values are tending to decrease but singularities are found in the first bin for NFIQ and in the second and third bins of QMF, which lead to a very low similarity between the GMS and quality values. Some researchers assert it is not deterministic that the correlation between GMS and quality values does exist. However, in the light of the experiments of this study, one can make an auxiliary validation of the quality metric in terms of this measure.

4.7 Conclusion

This study first proposes a fingerprint quality metric by considering image-based quality features and those derived from minutiae template. Second, the quality metric has been validated by using different validation approaches. In the study, the proposed quality metric was evaluated on different FVC databases, FVC2000 DB2 A, FVC2002 DB2 A and the first 3 datasets of FVC2004 Set A. Among the validation result, it can be observed that the performance of quality metric shows a great variation between different databases, indicating that image specification largely affects the validity of a quality metric which makes difficulties to achieve a commonly effective metric. In addition, the variation of metric performance evaluation also reveals that matching algorithm determines the contribution of quality metric to the overall performance, especially when a more robust matcher is involved. Furthermore, both the trial metric (QMF) and NFIQ are all associated with a prior-knowledge of matching scores which more probably lead to the variation of the evaluation result.

Second, by considering the performance of both the QMF and NFIQ, this study also gives observation that it is not easy to obtain a more robust or more complemental quality metric by fusing a variety of features. For instance, NFIQ performs worse than QMF on some databases even if the NFIQ fuses a group of quality features, yet the features of QMF are not strictly validated (*i.e.* features are not claimed to be quality features). In this case, one can note that quality metric is really not an absolute or a linear predictor of the matching performance, particularly in multi-vendor applications. However, one can also realize that the QMF is more suitable to deal with a specific scenario of FQA. To the end, we believe that it is possible to reach a common good quality metric with a prior-knowledge of matching performance in case the latter is commonly robust.

In the coming chapters, we focus on qualifying fingerprint with pixel-pruning and minutiae template only.

Chapter 5

FQA from Image Segmentation Maps

In this chapter, a general framework for estimating fingerprint image quality is proposed by fusing features after fingerprint segmentation. The quality index is indicated by a ratio of the pixel number of the integrated foreground area to the size (pixel number) of the fingerprint image or simply the pixel number of the remaining foreground area. Experimental results obtained from several trial datasets by using a dual evaluation approach demonstrate the validity of the proposed method in improving the overall performance.

Contents

5.1 Introduction

In previous chapters, two kinds of approaches for generating fingerprint quality metrics have been proposed. These two approaches demonstrate the possibility in assessing fingerprint image from different ways: regression and no-image quality assessment. However, both of them seem more significant in a trial manner than in a practical manner. In addition, according to the literature in Chapter 2, one can note that approaches based on categorizing fingerprint image areas or a single feature are easily affected by the change of image specification. On the other hand, assessing fingerprint quality with weighted or linear fusion is limited by employed coefficients, while the performance of quality metrics based on regression or classification are largely dependent on the involved regression approaches or classifiers, and the accuracy of the employed prior-knowledge. Therefore, it is still a challenge to achieve a common good quality metric for images captured by multi-sensor, even the resolutions of them are quite close to each other. Because of this, one can neither claim that metrics based on multi-feature fusion is able to make the assessment more

robust nor deduce that it is not easy to apply a single feature to images collected with various sensors.

This study proposes a new quality assessment framework based upon trimming foreground pixels of bad quality image as much as possible. Therefore, instead of using solutions presented above, this study generates a quality metric with multiple coarse segmentation. This framework is almost a two-step (or more) work which firstly performs one coarse segmentation to the fingerprint image and followed by another segmentation-like operation for a further pixel-removing. Finally, each of the segmentation results is simply used as a feature, which makes fusing features in segmentation phase possible. The potential advantage of this framework is that it could be improved by integrating other segmentation-based approaches or quality features rather than fusing them in a more complicated manner.

The following of this chapter is organized as follows: Section 5.2 presents the proposed framework in detail. Section 5.3 details the experimental results figured out via the evaluation approach given in Chapter 3, including simple ES and an auxiliary evaluation with NIMS. Section 5.4 concludes the proposed approach.

5.2 Proposed method

As the specialty of the biometric application, fingerprint quality is not only image distortion determination. The purpose of FQA is to guarantee the reliability of the extracted features from the image and hence benefits the matching performance. In this case, segmentation is initially a choice to determine the useful and reliable area of the ridge-valley pattern, which somehow indicates fingerprint's availability in a quantitative manner [74]. Existing studies of fingerprint segmentation also involve in pure feature-based approaches and solutions with learning algorithms [54, 146, 147]. The feature-based approaches are affected by image specification and some learning-based approaches rely on large size training set and may not be appropriate to quality assessment applications. **Note** that segmentation is not equivalent to quality assessment in this case. In addition, prior studies in segmentation-based quality assessment mostly focus on determining foreground block number in terms of one (or more) specific feature(s) or assigning a goodness value to a block [49, 72]. In this case, the quality assessment framework proposed in this study considers to perform segmentation-based operations in multi-task. This is able to ignore the coefficients problem required by fusion-based methods and takes the advantage of the feature(s)

used for segmentation-based approaches. Furthermore, this is not relevant to any prior-knowledge such as training samples. The selected segmentation criteria are common schemes and each of them acts as a module which is possibly to be replaced and improved.

5.2.1 Features given by Morphology Segmentation

The first step of the proposed framework is to obtain a measure of fingerprint foreground area as we have just mentioned before. To do this, a coarse segmentation is adopted in this study, which is achieved via morphological processing of images. Such a processing mainly consists of two tasks: dilation and erosion. Fingerprint image is composed by parallel run ridge-valley pattern with relatively stable frequency. With this property, it is able to connect the edges formed by the ridge-valley pattern (Cf. Figure 5.1).

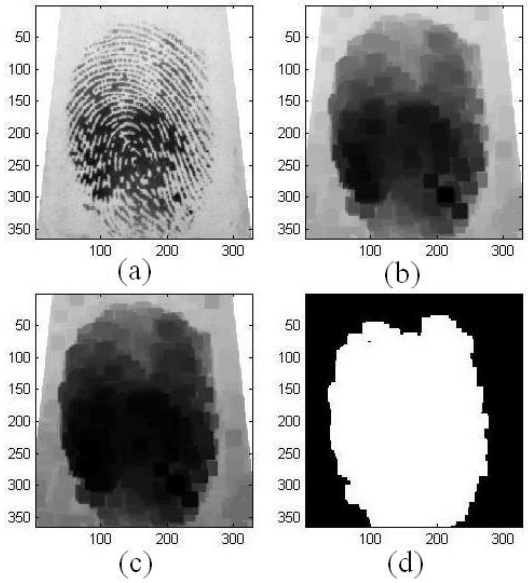


Figure 5.1 – Example of segmentation with morphology operation.

Four images in Figure 5.1 illustrate a morphology processing of a fingerprint image with several iterations, where image 5.1(a) is the original fingerprint pattern, 5.1(b) is the image after erosion processing(s), 5.1(c) is the enhanced version of image 5.1(b), and 5.1(d) is the segmented mask. In this study, we use the approach in [51] to perform the first coarse segmentation. The first feature for indicating fingerprint quality is hence a pixel ratio of the foreground area to the entire image.

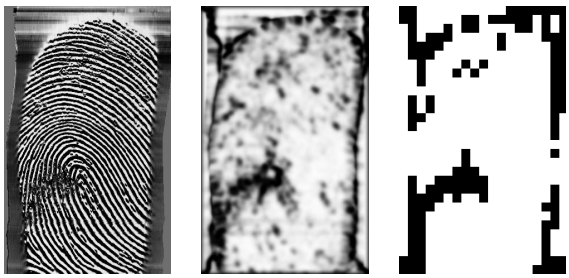
5.2.2 Pixel-pruning based on Coherence

In this task, we propose a pixel-pruning approach by using an existing segmentation criterion namely coherence [148]. The coherence is initially applied onto directional field estimation of fingerprint images and has been used as one of the features [148] for classification-based fingerprint segmentation approaches. The feature is to indicate the uniformity of the foreground gradients. In our experiments, we found that this feature is sensitive to the variation of the ridge-valley direction in a local area. Because of this, we customize an approach by using this feature to extensively remove foreground pixels in a local region where the directional information of the ridge-valley pattern changes abruptly. The definition of the coherence is given by gradient measures of pixel intensity. In a local window W, it is defined by

$$Coh = \frac{|\sum_W (G_{s,x}, G_{s,y})|}{\sum_W |(G_{s,x}, G_{s,y})|} = \frac{\sqrt{(G_{xx} - G_{yy})^2 + 4G_{xy}^2}}{G_{xx} + Gyy} \tag{5.1}$$

where $(G_{s,x}, G_{s,y})$ is the squared gradient, $G_{xx} = \sum_W G_x^2$, $G_{yy} = \sum_W G_y^2$, $G_{xy} = \sum_W G_x G_y$ and (G_x, G_y) represent the local gradient.

Figure 5.2 illustrates an example of the pixel-pruning result of a fingerprint image.



(a) Original (b) Coherence (c) Mask

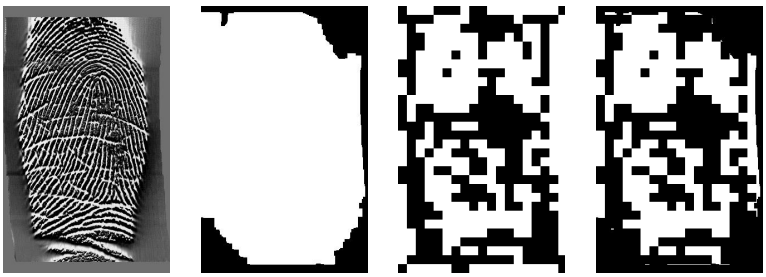
Figure 5.2 – Example of segmentation with Coherence.

In Figure 5.2, image 5.2(b) is the coherence image calculated from the original fingerprint illustrated by 5.2(a), while image 5.2(c) is the segmentation mask obtained by using pixel-pruning method which is carried out via a thresholding operation to the coherence image.

In our study, the coherence image is first normalized into [0,1], and then divided into non-overlapped blocks which is followed by thresholding operations with a baseline value of 0.5. The block size is 16 in this study, and both the block size and the threshold are all empirical values in our study. Finally, the quality feature is also a ratio of the light pixels number to the pixel number of the entire image.

5.2.3 Metric Generation

The proposed framework of fingerprint quality assessment is essentially implemented by fusing two (or more) features in the segmentation phase, i.e. the binary images of mask obtained in the segmentation stage and pixel-pruning session would be combined together. Considering score-based fusion in biometrics [93], one can observe that there are several ways to achieve fusion task such as 'min' and 'max' rules. In the proposed framework, we simply use the logical 'and' rule to fuse two binary mask images, which is actually equivalent to fusing two features (obtained by two steps) in terms of the 'add' rule. An example of such a fusion is given in Figure 5.3.



(a) Original (b) Morphology (c) Coherence (d) Fused

Figure 5.3 – Example of segmentation with Coherence

In Figure 5.3, one can note that the morphology approach coarsely generates a entire foreground area, where coherence is used for removing pixels in terms of the mean value of coherence in block-wise. In our study, the coherence image is normalized into [0,1] and a threshold of 0.5 is involved in pruning pixels as much as possible. The block size is 16 in this study. All of them are empirical values, and there is no experiment data support the block size and the threshold. It could be possible to modify the threshold value in terms of the variation of the global error rate. However, in this study, 0.5 is simply chosen as the coherence image is normalized into [0,1], and this value works for variant datasets in the experiment. The block size is 16 simply because it is a moderate value for a fingerprint. Finally, the quality measure is a ratio of the light pixels number to the pixel number of the image.

In Figure 5.3, one can note that the morphology approach is to coarsely generate an entire foreground area, where the pixel-pruning approach is used for removing pixels in terms of the mean value of coherence in block-wise. The pruning task is particularly effective for bad quality images that contain some abrupt changes of the

direction of the ridge-valley flow. The validation of the proposed approach is given by evaluation experiments, see Section 5.3.

5.3 Experiment results

The experiment given in this section includes two parts based on the Enrollment Selection (ES) given in Chapter 3. At first, the simple ES associated to the proposed quality metric (denoted as MSEG hereinafter) and a reference metric (NFIQ) is performed with the original dataset. Second, the ES with the quality metrics is carried out via the auxiliary NIMS approach. The experiments employ fingerprint databases given in Chapter 3 and two re-organized databases. The experiment is carried out via two different matching algorithms in this study. Details about employed datasets are given below.

5.3.1 Protocol

In experiments, we generate two re-organized datasets from the CASIA FP-Test V1 [1] database. The CASIA database contains fingerprint images of 4 fingers of each hand of 500 subjects, where each finger has 5 samples. In this study, we create the two re-organized databases by using samples of the second finger of each hand, and they are respectively denoted as CASL2 and CASR2. Therefore, each sub-database has 2500 images of 500 individual (5 samples per individual). The samples of the CASIA dataset is captured by an optical sensor, which creates gray-level image of 328×356 and the image resolution is 512-dpi. A glance of the two datasets are given by samples in Figure 5.4.

(a) CASL2 (b) CASR2

Figure 5.4 – Sample illustration of CASIA database. Left and right hand: 5.4(a), 5.4(b).

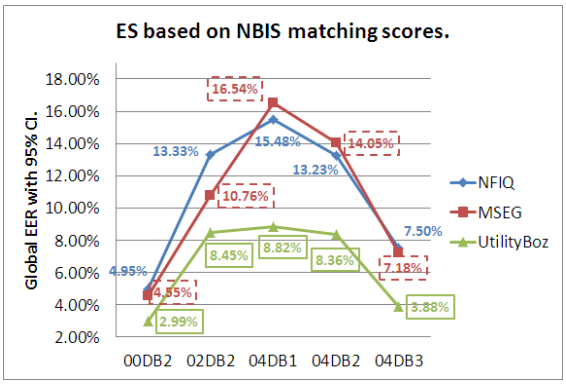
1. http://biometrics.idealtest.org/detailsDatabase.do?id=3

In this case, the intra-class matching scores consisted of 500×4 genuine matching scores (GMS) and the inter-class matching scores contains 500×499×4 impostor matching scores have been computed for these two datasets via two matching algorithms, respectively.

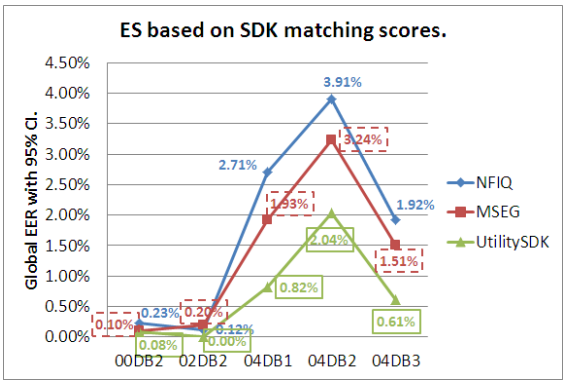
5.3.2 Results

Experiment results are also indicated by a set of global EER values and their 95% confidence interval (CI) obtained from each dataset by substituting the associated **utility** and quality values to the ES, respectively. Results obtained from FVC and CASIA datasets are given separately, for the quality of them are quite different.

Figure 5.5 plots the global EERs of the FVC datasets, where Figure 5.5 (a) is the result calculated from the NBIS matching scores and Figure 5.5 (b) shows the result obtained by using the matching scores of the SDK.



(a) Results based on MS Boz



(b) Result based on MS SDK

Figure 5.5 – Global EER plots. UtilityBoz and UtilitySDK in Figure 5.5(a) and 5.5(b) are plots of the global EER obtained with two sets of sample utilities.

In Figure 5.5 (a), when NBIS matcher is involved, MSEG (red plot) respectively generates 16.54% and 14.05% on 04DB1 and 04DB2 which are relatively bad results in comparing with the reference metric (blue plot), while MSEG shows better results on the other 3 datasets. On the other hand, MSEG (Figure 5.5 (b)) performs relative bad on 02DB2 only and better on the other 4 datasets when a vendor-free matcher (SDK) is used. This is due to the difference of the matching performance between the two algorithms. In addition, the NFIQ is involved in a prior-knowledge of matching performance, which could more probably result in a different evaluation result. The global EERs of MSEG and NFIQ obtained from 02DB2 are 0.2% and 0.12%, respectively. The global EERs obtained by samples **utility** are plotted via green points in each figure. The sample **utility** is simply an approximation of the groundtruth (with respect to the employed matcher) of the original sample as it has been discussed in Chapter 3. The utility-based global EERs are illustrated as a reference, indicating how much the quality metric is close to the best case that one matching algorithm can obtain from a trial dataset.

Table 5.1 gives the CIs of the quality-based global EERs, which has already included the CIs obtained from the two CASIA datasets.

Table 5.1: The 95% CI of the global EER of each metric.

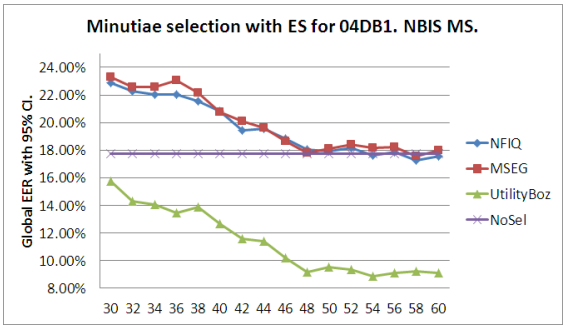
QM / DB	NFIQ	MSEG
00DB2A (NBIS)	[0.0490 0.0500]	[0.0450 0.0461]
02DB2A (NBIS)	[0.1326 0.1340]	[0.1068 0.1084]
04DB1A (NBIS)	[0.1540 0.1557]	[0.1645 0.1662]
04DB2A (NBIS)	[0.1312 0.1334]	[0.1396 0.1413]
04DB3A (NBIS)	[0.0745 0.0756]	[0.0712 0.0723]
00DB2A (SDK)	[0.0022 0.0024]	[0.0009 0.0011]
02DB2A (SDK)	[0.0011 0.0013]	[0.0019 0.0021]
04DB1A (SDK)	[0.0266 0.0275]	[0.0189 0.0196]
04DB2A (SDK)	[0.0384 0.0397]	[0.0319 0.0328]
04DB3A (SDK)	[0.0189 0.0195]	[0.0148 0.0153]
CASL2 (SDK)	[0.4087 0.4097]	[0.3856 0.3866]
CASR2 (SDK)	[0.3815 0.3825]	[0.3592 0.3603]

The global EER values obtained by each quality metrics from the CASIA datasets are presented in Table 5.2.

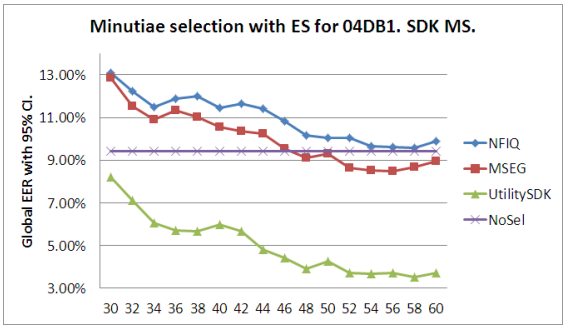
Table 5.2: The global EERs obtained from CASIA datasets.

QM \ DB	CASL2 (NBIS)	CASR2 (NBIS)	CASL2 (SDK)	CASR2 (SDK)
NFIQ	43.09%	43.51%	40.92%	38.20%
MSEG	42.30%	43.20%	38.61%	35.97%

According to the results given in Table 5.2, one can note that the proposed MSEG shows its validity in comparing with the reference quality metric. By using the matching scores of NBIS software, the global EERs obtained by MSEG are 42.30% (CASL2) and 43.20% (CASR2), while NFIQ generates 43.09% (CASL2) and 43.51% (CASR2), respectively. Likewise, the global EERs calculated with the matching scores of the SDK are 38.60% (CASL2), 35.97% (CASR2), 40.92%(CASL2) and 38.20% (CASR2), where the first two values correspond to the MSEG and the last two values belong to the NFIQ. The CIs given in Table 5.1 are also consistent with these global EERs, indicating the validity of the proposed MSEG. Meanwhile, the experimental result also shows that the MSEG is commonly available for multiple image specifications, at least the employed image types.



(a) Results based on MS Boz



(b) Result based on MS SDK

Figure 5.6 – Global EER plot obtained by ES of reduced templates. NoSel means the global EER obtained from the reduced dataset without ES.

Second, we perform validation of the MSEG by using the auxiliary NIMS mentioned in Chapter 3. This kind of operation needs to perform on one dataset for each desired number. Therefore, we simply choose one database as an example to illustrate this auxiliary method. In the experiment, the 04DB1 is used since there is a dissent between two matchers. In addition, the reduced datasets and their matching scores of 04DB1 are already available in [108]. Also, the matching performance with the original template (NoSel) of this database is far from the global EERs obtained by utility-based ES (step 3), which makes a clear illustration. The plots of the global EERs of the reduced datasets are given in Figure 5.6.

In Figure 5.6 (a) and (b), 'UtilityBoz' and 'UtitlitySDK' respectively indicate global EERs' plots obtained via NBIS-based **utility** and SDK-based **utility** for each reduced template set. Obviously, by comparing the plots of the global EER values associated to the metrics, NFIQ (blue) and MSEG (red), the result is basically consistent with the ones given in Figure 5.5. A little bit variation appears as the desired number increased to 48 or so when calculating the global EER with NBIS matching scores. This is reasonable according to the study of NIMS [108]. In this case, with the reference quality metric and the objective measure (**utility**), one can found that the proposed framework is a valid solution for assessing fingerprint quality.

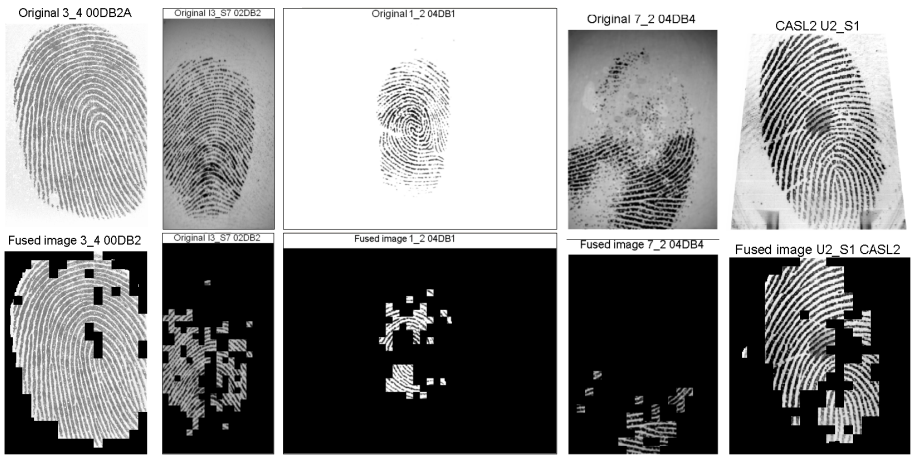


Figure 5.7 – Illustration of the effect of pixel-pruning. The 1^{st} row shows original images and the 2^{nd} row shows the fused images of the originals and the final mask.

At last, several examples are given to illustrate the effect of the proposed pixel-pruning to the performance of the MSEG, see figure 5.7. In Figure 5.7, the first example and the last one are relatively good looking images in FVC2000DB2 and the CASIA dataset, respectively. Others are those relatively bad looking samples randomly picked up from the associated dataset. Apparently, the pixel-pruning

operation achieves the purpose in removing pixels extensively from bad quality images. On the other hand, the singularity areas are mostly removed from the good quality images because the pixel-pruning is dependent on the variation of direction of a local pattern. In this case, it does not affect so much good quality images, and hence gives a valid assessment result.

An implementation of the proposed framework is available online [2].

5.4 Conclusion

In this chapter, a new framework for qualifying fingerprint images is proposed. The proposed solution achieves fusing multi-feature in the segmentation phase, which avoids coefficient problem of the weight-based combination approaches. In addition, the proposed framework shows some generalities to variant fingerprint images. Finally, this approach is not related to any prior-knowledge, which makes the evaluation more independent.

The validity of the proposed framework is proved via a comparison with the reference metric. The proposed quality assessment framework also has a potential advantage in improving itself via any available feature of fingerprint segmentation tasks. Nevertheless, it could be integrated into the other quality assessment framework with multiple features as well.

2. https://www.greyc.fr/users/zyao

Chapter 6

FQA from Minutiae Template

A new quality assessment method of fingerprint derived from only its minutiae points is presented in this chapter. The proposed quality metric is modeled with the convex-hull and Delaunay triangulation of the minutiae points. The validity of this quality metric is verified on several trial datasets by referring to an image-based metric from the state of the art.

Contents

6.1 Introduction

In previous sections, prior studies in estimating fingerprint quality had been introduced, by which one can note that some of the existing methods are carried out by using an unique feature, while others combine multiple features together. To this point, one can found that no matter how many aspects are considered by a fingerprint quality assessment approach, it is all about features employed for generating the quality metric. A valid feature or metric should be able conduct positive contribution to matching performance [27]. This kind of contribution of sample quality has been standardized as a utility property [149], *i.e.* sample quality reflects its impact on the performance of the system. However, system performance fully relies on the matching approach such as minutiae-based system which is employed most in actual deployments. In this case, it is reasonable to consider qualifying fingerprint with minutiae information, particularly when considering quality assessment without its image information.

There are very few quality assessment approaches that take into account minutiae information, such as NFIQ. Moreover, none of the state-of-the-art approaches qualifies a fingerprint from only the minutiae template. The main contribution of the proposed

study is an original algorithm that computes a quality score from a minutiae template associated to the fingerprint. In another word, it could be viewed as a quality metric for assessing the quality of a minutiae template. The benefit of having this type of metric could be related to embedded biometric systems in smart cards or smart objects where only the minutiae template is available due to computational and storage constraints.

This chapter is organized as follows. Section 6.2 presents the proposed quality metric. Details of the experimental results and a discussion are given in Section 6.3. Section 6.4 concludes this chapter and discusses the perspectives.

6.2 Metric definition

Fingerprint matching approaches, according to the literature [43], are broadly classified as minutiae-based, correlation-based or image-based, among which minutiae-based is the most widely studied solution. Minutiae mainly include two types of points known as ridge bifurcation and ridge ending which are the minor details of a fingerprint image, particularly being used for fingerprint alignment and matching [139]. A minutiae template generally provides three kinds of information: 1) minutia location, 2) the orientation and 3) the type of minutia point, see Figure 6.1.

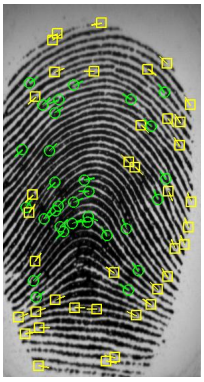


Figure 6.1 – A fingerprint and its minutiae template.

These features are sufficient to reconstruct a synthetic fingerprint from a given template [138] as the orientation field can be estimated. Another kind of information is the amount of detected minutiae points, which has been used as one factor [69] for quality assessment of fingerprint samples. In addition, Prabhakar *et al.* [150] proposed that the matching accuracy can be greatly improved when using minutiae type information as a complementary factor for matching. However, this information is far from being enough for generating an effective quality metric.

In order to do so, the minutiae template is related with the foreground of fingerprint because the area of minutiae indicates an available and useful region for the so-called 'extractability' [29] of features. According to previous chapters, foreground area is known as one index of fingerprint image quality, *i.e.* a good quality fingerprint should have at least enough area for matching. Therefore, when dealing with fusion-based metric in Chapter 3, minutiae account in different regions was considered. However, this kind of measurements failed to reflect any of image characteristics. Because of this, this study pays more attention to the area of minutiae region which is the only element related minutiae template to a fingerprint image. Rectangle and circle regions are able to approximately estimate the foreground area but they have difficulties to measure local characteristics when using minutiae template only. In this case, this chapter proposes to estimate fingerprint quality by modeling the associated minutiae template with the convex-hull and Delaunay triangulation (Cf. Figure 6.2).

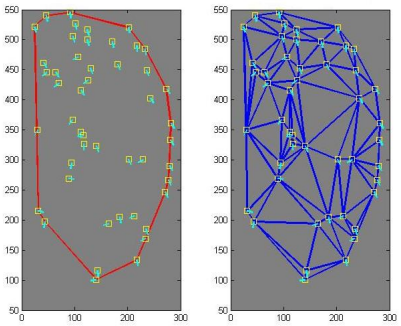


Figure 6.2 – A convex hull (left) and Delaunay Triangulation.

The 2D convex hull [151] $C_i = \{(x_j, y_j)|j = 1, \cdots, N_i\}$ (where N_i is the minutiae number of the i^{th} fingerprint) is defined as the smallest convex set that contains all the points of a given set of minutiae F_i in this case (Cf. Figure 6.2). The Delaunay Triangulation of a set of points P in 2D plane is defined as a triangulation DT(P) [152] that none of the points of the given set are inside the circumcircle of any triangle of DT(P), as illustrated in Figure 6.2.

With the respective properties of these two geometric structures, an area of the smallest informative region of one fingerprint and a set of areas and perimeters of triangles composed by each 3 of its minutiae points are obtained. By doing so, one unavoidable problem is observed that almost all the templates of bad quality fingerprints contain both correctly detected minutiae point and spurious points. As presented in Chapter 2, most of matching algorithms are based on minutiae points,

and geometric structures are also adopted in many studies [153, 154]. Based on these observations, we proposed a quality metric as described below.

Let a minutiae template F_i of a fingerprint containing a set of detected minutiae points represented by $m_j = (x_j, y_j, \theta_j)$, where (x_j, y_j) is the location and θ_j is the orientation of the j^{th} minutia point. There is a convex hull denoted by C_i and a set of triangles (T_k) formed by Delaunay triangulation which is formulated as $DT(F_i) = \{T_k | k = 1, 2, \cdots, L_i\}$, where L_i ($< 2N_i$, is the number of) triangles can be constructed from F_i. Correspondingly, the convex hull encloses all the minutiae of F_i with the smallest area represented by A_i, and a set of areas S_k and perimeters P_k of $DT(F_i)$ are respectively obtained.

In our experiments, we observed that bad quality samples generate tiny and extremely narrow triangles (considered as unreasonable) due to spurious minutiae points. The quantitative values of their area and perimeter are (visually) not proportional to each other, as observed in Figure 6.3.

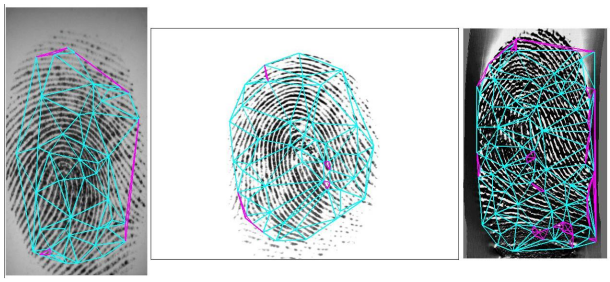


Figure 6.3 – Example of minutiae Delaunay triangulations of 3 different FVC images. NFIQ values are 2, 2 and 1, respectively.

In Figure 6.3, triangles drew by pink color indicate unreasonable minutiae structures. Based on these observations, the quality metric (abbreviated as MQF afterwards) is calculated with several steps, as described by algorithm 1. In the computation of the proposed quality metric, three thresholds are set for triangle perimeter, triangle area and the ratio between the perimeter and area of the triangle, which are represented by Y_p, Y_a and Y_r, respectively. The details for choosing the most appropriate parameters are discussed in Section 6.3. The quality value q is dependent on the area of the geometrical structure of the minutiae region, which could be normalized into the range of [0, 100] on each database.

Algorithm 1 Computation of the quality score.

Input:

Minutiae Template F_i.

Output:

Quality index, q;

1: Calculate the area of the convex hull C_i, denoted as A_i;
2: Calculate perimeter and area for each triangle T_k, denoted as P_k and S_k;
3: $A_{Y_p} = area\left(Search\left(P_k < Y_p\right)\right)$;
4: $A_{Y_a} = area\left(Search\left(S_k < Y_a\right)\right)$;
5: $A_{Y_{pa}} = area\left(Search\left(P_k < Y_p \;\&\; S_k < Y_a\right)\right)$;
6: $A_{Y_r} = area\left(Search\left(P_k / S_k\right) > Y_r\right)$;
7: $S_{area} = \left(A_i - A_{Y_p} - A_{Y_a} - A_{Y_r} - A_{Y_{pa}}\right)$; // No duplicated triangles.
8: **return** $q = S_{area}$;

Obviously, this algorithm relates the minutiae template with the area of a region for matching operations as it has just been mentioned before. However, this factor is not commonly sufficient so that we consider to remove a part of potentially useless area from the informative region due to the lack of image information. Delaunay triangulation gives a relatively ideal solution to this problem thanks to the spurious minutiae of bad quality images. For example, as it is depicted in Figure 6.3, some unreasonable triangles formed by spurious minutiae on the border of the foreground enable us to remove the corresponding area. Another case is the tiny triangle which is mostly happened to the area where the quality of ridge-valley pattern is relatively bad. The study also noted that some fingerprint images have several genuine minutiae clustering in a very small area. This case is not specially considered in the proposed algorithm just because of the limitation of the template. The area of each triangle is obtained by using Heron's formula [155].

6.3 Experimental results

In order to validate the MQF, we adopt several approaches to illustrate the performance of the quality metric: 1) the Pearson correlation coefficients between the MQF and several others (from the state-of-the-art) are calculated and 2) the MQF is evaluated by using the proposed enrollment selection (ES) given in Chapter 3. The experiments are performed with a laptop driving by an Intel Celeron dual-core CPU of 1.73GHz. The experimental protocol is presented first.

6.3.1 Protocol and databases

In this study, experiments are conducted with the five trial databases given in Section 3.4. The minutiae template of MQF is generated by using the OpenSource software MINDTCT addressed in Section 3.4 as well. This extractor generates a quadruple representation of minutia point, $m_i = \{x, y, o, q\}$, where (x, y) is the location of minutia point, o indicates orientation and q is a quality score of minutia point. In the experiment, the location has been used only for calculating the proposed quality metric. In the evaluation stage, we also use the Bozorth3 and SDK to calculate matching scores.

6.3.2 Parameter settings

Fernandez *et al.* [71] and Olsen [102] respectively calculated Pearson and Spearman correlation coefficients between different quality metrics to observe their behavior or similarity. Similarly, we investigate the behavior of the proposed quality metric through the Pearson correlation coefficients, by which the parameters are appropriately selected as well.

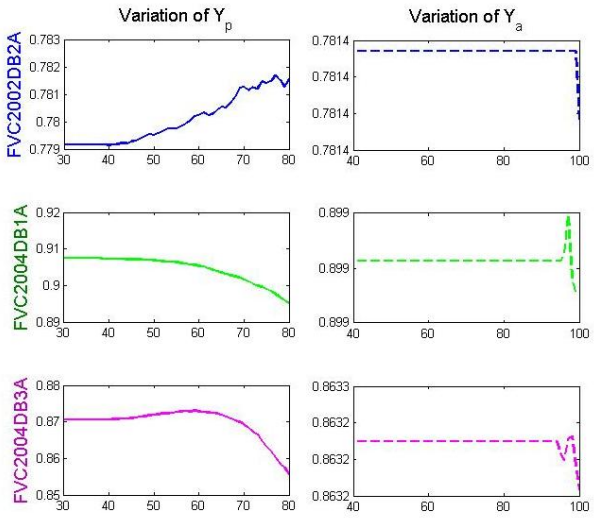


Figure 6.4 – The variation of the Pearson correlation between the proposed metric and OCL as the parameter changes. X-axis: parameter; Y-axis: Pearson correlation.

The three thresholds (Y_p, Y_a and Y_r) are all empirical values observed in the experiments, and they are all dependent on the resolution of the image which is supposed to be over 500 dpi in this study. With the experiments, we noted that the

smallest area of the triangles of each template is generally less than 70 (considering a histogram of 100 bins). This value is only a scalar value without considering the unit of the measurement. In this case, we firstly preferred to choose a value of the area over 70. In addition, it is easy to observe that a threshold larger than 80 may loose the significance of this parameter. The variation of the correlation values (larger than 0.3) demonstrates this problem, see Figure 6.4.

Likewise, the smallest perimeter value of the triangles of each template is mostly within the interval [10, 80], and generally lies around 35 (according to the histogram). In this case, we further consider the relation between the area and the perimeter of those triangles that seemed abnormal. Without considering the unit of the two measurements (one is 1-D and another is 2-D), we observed that the value of the perimeter is generally smaller than the area. In this case, inspiring by the Heron's formula, we choose a threshold to represent the ratio between these two measurements. We observed that the triangle is extremely narrow if the ratio between a perimeter and an area is close to 1. Besides, similarly, it is also not necessary to consider larger values for the perimeter. We first choose a reference quality metric among all the others. The correlation of OCL does not vary so much on variant databases. Therefore, we simply choose the OCL in the experiment. Figure 6.4 provides only the graphical result of the variation of the correlation value between OCL and the proposed metric obtained via a series values of two parameters. In the experiments, the thresholds of the area and perimeter vary in the range of [40, 100] and [30, 80] with an interval of 1, respectively.

According to Figure 6.4, we also observed that the variation of Y_p leads to more impacts on the correlation coefficient than Y_a when other parameters had been set approximately. In order to achieve a generality of the proposed quality metric, the values of Y_p, Y_a and Y_r in this study are 75, 70 and 0.8, respectively. At last, the behavior of these parameters were estimated by using performance validation approach, and the variation of the performance measurement (EER) obtained in a small interval of each of them tends to be stagnant. With this empirical analysis, the parameters are set as what had just been mentioned.

6.3.3 Feature analysis

To validate the MQF, we calculate the correlation coefficients between several quality metrics and reference ones. In addition to the two orientation-based indexes, OCL and the orientation flow (OF) [80], we also employ a wavelet domain feature carried out via the Pet Hat's continuous wavelet (PHCWT) [77]. In the literature, it is said that the Pet Hat wavelet is sensitive to the sharp variations of features such as

fingerprint ridges. Nevertheless, we also use a pixel-based quality metric which is the standard deviation (STD) of fingerprint local block indicating pixel information of the image [73]. The NFIQ is also involved in this section. Table 6.1 presents the Pearson correlation results of the trial quality metrics.

Table 6.1: Inter-class Pearson correlation coefficients. FVC2002DB2A

OCL	OF	PHCWT	STD	MQF	NFIQ
1	0.013	0.932	0.892	0.781	-0.503
0.013	1	0.092	0.122	0.070	0.061
0.932	0.092	1	0.954	0.788	-0.474
0.892	0.122	0.954	1	0.678	-0.374
0.781	0.070	0.788	0.678	1	-0.422
-0.503	0.061	-0.474	-0.374	-0.422	1

Table 6.1 provides only an inter-class Pearson correlation result of the employed quality metrics obtained from FVC2002DB2A. According to the coefficient values indicated with the highlighted cells, one can observe that MQF demonstrates the correlated behavior with the others except the OF. The correlation coefficients between the proposed metric and others calculated from the remaining databases are given in Table 6.2

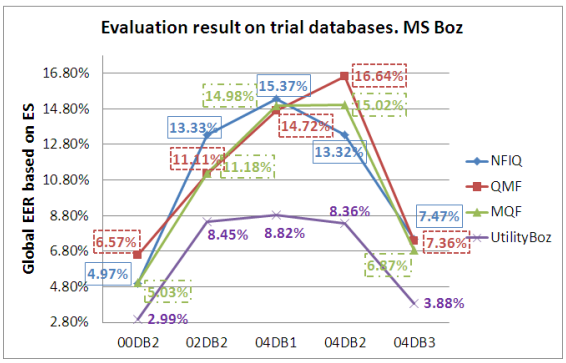
Table 6.2: Pearson correlation coefficients between MQF and others.

MQF of	OCL	OF	PHCWT	STD	NFIQ
00DB2	0.409	-0.131	0.291	0.301	-0.081
04DB1	0.899	0.253	0.905	0.817	-0.201
04DB2	-0.050	-0.489	0.722	0.650	-0.378
04DB3	0.863	0.021	0.818	0.811	-0.363

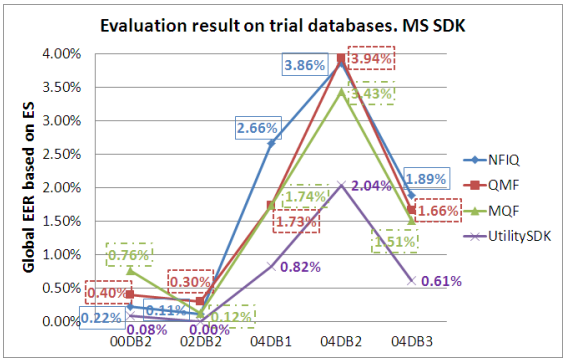
According to the results, in addition to FVC2000DB2A, one can note that MQF shows a relatively stable correlation (similarity) with other metrics, *i.e.* it exhibits an usability to variant databases. The correlation value between the proposed metric and OCL for FVC2004DB2A shows an odd value because a lot of over-inked samples are contained in this database. This problem results in some difficulties for calculating the OCL and leads to a lot of singular values. Note that the proposed metric uses only the set of minutiae location, and the fingerprint image is considered as unavailable.

6.3.4 Evaluation result

In this section, the ES with a reference quality metric (NFIQ), the QMF proposed in Chapter 4, MQF and sample utility (best case) are performed for validating the MQF. The evaluation results obtained from five trial databases are given by the global EER values and their confidence interval. Two graphical plots of the global EER values generated by the ES are illustrated in Figure 6.5.



(a) Result based on MS Boz



(b) Result based on MS SDK

Figure 6.5 – Plots of global EERs obtained by using ES. Figure 6.5(a) is the result based on MS Boz and Figure 6.5(b) corresponds to MS SDK.

In Figure 6.5, the UtilityBoz and UtilitySDK correspond to the global EERs obtained by using ES based on two sets of utility values, respectively. The other three plots are results generated by ES with each quality metric. When the matching score of the Bozorth3 (MS Boz) is used, the proposed MQF shows clearly worse results than NFIQ on 00DB2A and 04DB2A, for which the global EER values obtained by MQF are 5.03% and 15.02%, while the counterparts of the NFIQ are 4.97% and 13.32%, respectively. However, MQF performs worse only on 00DB2A when matching scores of the SDK (MS SDK) are involved in the experiment. The global EERs obtained by two metrics from this dataset are 0.76% (MQF) and 0.22% (NFIQ), respectively.

Apparently, as it has been discussed in Chapter 3, interoperate study shows different evaluation results due to the dissent between matching algorithms. According to this result and the correlation coefficients given in Table 6.2, one can note several problems that are negotiable. For instance, classification with multiple features by learning a prior knowledge of matching score might not be able to take advantages of the employed features. Besides, the Pearson correlation between one measurement and the others could be a solution to estimate the availability of the measurement for being used as a potential feature of a quality metric. However, we do not assert it is sufficient to evaluate the performance of a quality metric.

The statistical index, 95% CI , presented as well in Chapter 3 of the global EER values are listed in Table 6.3.

Table 6.3: The 95% confidence interval of the EERs.

QM et MS / DB	MQF Boz	NFIQ Boz	MQF SDK	NFIQ SDK
00DB2A (CI)	[0.0497 0.0509]	[0.0492 0.0502]	[0.0074 0.0078]	[0.0021 0.0023]
02DB2A (CI)	[0.1109 0.1128]	[0.1326 0.1340]	[0.0011 0.0013]	[0.0011 0.0013]
04DB1A (CI)	[0.1491 0.1506]	[0.1529 0.1545]	[0.0171 0.0177]	[0.0268 0.0276]
04DB2A (CI)	[0.1489 0.1515]	[0.1321 0.1344]	[0.0338 0.0349]	[0.0390 0.0402]
04DB3A (CI)	[0.0681 0.0693]	[0.0741 0.0752]	[0.0148 0.0154]	[0.0190 0.0195]

In Table 6.3, the CI results of the global EER values statistically illustrate the validity of the proposed quality metric. The overlap happened to 02DB2A when the matching score of SDK is employed, for the global EER values obtained from this database is nearly the same. In experiments, the error range of the global EER values has already been discussed in Chapter 3. This result listed in Table 6.3 is fully able to demonstrate the reliability of the global EER in a statistic manner.

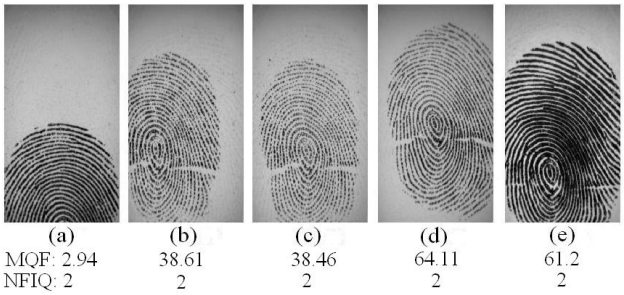


Figure 6.6 – Illustration of the different between MQF and NFIQ.

Moreover, we simply choose several images to illustrate the differences between the two quality metrics and their drawbacks. For instance, the MQF values of the

samples given in Figure 6.6 are (2.94, 38.61, 38.46, 64.11, 61.2), and their NFIQ values are all level 2. According to sample (a), one can find that it contains only a partial of the fingerprint image which is not suitable for matching. In addition, the samples with MQF values under 40 are not as good as a level 2 sample in this example. The samples (b) and (c) would result in spurious minutiae. On the other hand, an obvious shortage of MQF is the area measurement which would generate outliers in many cases, such as the samples illustrated by Figure 6.6 (d) and (e).

At last, by using the CASIA database, a comparison between the state-of-the art quality metric and three proposed solutions is given by the results derived from the vendor-free matching scores (MS SDK), see table 6.4.

Table 6.4: Comparison results of 4 quality metrics.

Metric / DB	CASL2 (MSSDK)	CASR2 (MS SDK)
NFIQ	40.92%	38.20%
QMF	42.72%	41.26%
MSEG	38.61%	35.97%
MQF	42.19%	40.94%

These two databases are used in the experiment because they are relatively difficult and hence more suitable to show the ability of a target quality metric. The approach based on pixel-pruning (MSEG) obtains the best result, while NFIQ comes the second, the MQF given in this chapter obtains better result than the one defined in chapter 4. However, one should note that the MQF uses only minutiae template. In this case, the proposed solutions are validated by using the validation framework defined in 3, especially when relatively normal databases (FVC) are used in the experiment.

The experiments are implemented via Matlab 7.12. The computation takes approximately 0.423 second per sample according to a calculation of 1600 samples when the input is an image, and 0.122 second for template inputs.

6.4 Conclusion

This study mainly focuses on estimating fingerprint quality simply with a minutiae template. By investigating the relation between image-based quality criteria and the structure of minutiae template, we calculate the area of minutiae region via a convex hull and the Delaunay triangulation. The uniqueness of Delaunay triangulation provides a possibility that some unreasonable minutiae could be further eliminated

from template, and the particular rules for calculating convex hull enables estimating a relatively minimum area of the detected minutiae. Accordingly, we define a simple yet efficient quality metric for fingerprint and minutiae template. In another aspect, the quality metric could be affected by some bad samples with large minutiae region. The lack of image information makes this problem inevitable. Correspondingly, we analyzed the proposed quality metric via Pearson correlation coefficients and evaluated its performance with an enrollment phase approach. Therefore, it is reasonable to conclude that the proposed quality metric is an effective measurement for assessing the quality of a various of fingerprint samples when their associated templates are available only. In this thesis, this approach is classified as a segmentation-based solution.

As some of fingerprint quality metrics are generally defined with multiple features, this quality metric therefore could be a candidate for those composite metrics. The future work of this study tends to concentrate on combining this quality metric with some others extracted from minutiae template to obtain better performance.

Conclusions et perspectives

Biometric techniques have been widely deployed in recent years and will be an essential part of the constitution of information security and other social service in the near future. Fingerprint, as it has been detailed in this thesis, has been one of the most important means of biometric applications and will still be a leading role in this domain. Existing studies have noted that biometric sample quality is a significant factor in improving or guarantee the overall performance of biometric systems. Likewise, many attempts in qualifying fingerprint samples have been made so far. This thesis mainly focuses on fingerprint quality assessment (FQA) and the evaluation of quality metrics. During the study of this thesis, we have been making effort in quality assessment with different and new solutions which enable us in observing some potential problems and give us some motivations for the future work, as presented in the following.

Contributions

In Chapter 1 of the thesis, a quick review of biometrics has been given in terms of the general structure of a biometric system and its performance evaluation. Such a technical review gives the prerequisite and a comprehensive understanding for a further study of fingerprint modality.

Chapter 2 introduced the fingerprint modality in detail, including most of the processing operations for this modality. This chapter presents a literature review to most of the existing studies of FQA, which detailed quality assessment approaches in terms of each of the representative solutions that had been proposed so far. Meanwhile, difference between those types of solutions and some potential problems have been discussed. For instance, according to experimental results given in some of the reviewed articles, some metrics that rely on a single feature are better than the benchmark metrics implemented via multiple features. Does this mean that metrics based multi-feature do not really take the advantages of the employed features? This needs to be answered with the support of experiments.

Chapter 3 firstly addresses a quick review of performance evaluation approaches of quality assessment algorithms. Next, a new validation framework for estimating biometric quality metrics has been proposed, which is able to statistically estimate a

quality metric via several global measurements, such as global EER and AUC values. This chapter shows that quality is not absolutely a linear predictor of matching performance.

Chapter 4 proposes the first solution in this thesis for assessing fingerprint quality. The proposed approach is carried out by combining several different types of feature of both fingerprint image and its minutiae template. Meanwhile, this solution relies on a prior-knowledge of matching performance, specifically the genuine matching scores (GMS). Together with the observation of Chapter 3, one can note that the efficacy of quality metrics based on multi-feature fusion really depend on the associated fusion algorithms, and they are constrained by sample types. Furthermore, metrics implemented via a prior-knowledge are easily affected by matching algorithms, *i.e.* the performance of this kind of metrics largely depend on the accuracy of the prior-knowledge.

In Chapter 5, a new solution has been proposed by using multiple image segmentation maps. This approach estimates the fingerprint quality by fusing several features in segmentation phase. One should note that segmentation is not exactly the same as quality assessment. This kind of qualification framework provides possibility to use any kind of segmentation techniques for FQA.

Chapter 6 proposed another new solution of FQA by using only the minutiae template of a fingerprint, which enables qualifying fingerprint when minutiae template is known alone. This is also one of our trials, which none of existing studies has considered to assess quality in such a case, showing the possibility for quality assessment without using full image information.

With all summaries above, this thesis achieves assessing fingerprint quality in different ways: 1) multi-feature fusion with a prior-knowledge of matching performance, 2) geometric structure of minutiae template and 3) fusion in segmentation phase. All these proposed solutions can represent each of the typical solutions of FQA in the existing studies, and able to give us a comparative study to demonstrate whether fusion-based approach really achieves a common good and robust quality index. The assessment approach given in Chapter 5 shows its generality for several different types of fingerprint images. However, according to the experimental results of this thesis, one can notice that it is still a challenge to get a quality metric for every kind of fingerprint images, even the resolution of target gray-level image is limited at about 500-dpi. In addition, some new forms of fingerprint applications have appeared quite recently, such as smart phone. Therefore, quality assessment of fingerprint requires as well appropriate solutions. Nevertheless, several problems still need to be considered in the future studies of the classical fingerprint applications, which are given below.

To this end, FQA is recommended to be independent from the prior-knowledge of matching performance because none of existing matching algorithm is ideal. Similarly, the validation or evaluation of biometric quality assessment approach should consider both the genuine and impostor errors because it is essentially a biometric test.

Perspectives

According to our studies, we believe the FQA is still an open issue for several reasons:

1. Once image specification is fixed, a good quality metric or associated feature should be available for various samples captured via different sensors, while most of the existing studies are not able to satisfy this criterion.

2. Multi-feature fusion for FQA is still challenging, for most of the existing studies do not really make multiple features more robust than a single feature. Typical solution of this kind of scheme concretely includes linear fusion [80], regression [84] and classification. To achieve the fusion adaptively, large scale dataset or a set of specific target sensing devices could be considered.

3. Mobile applications such as camera shot of fingerprint requires qualifying fingerprint image in different aspects such as distortions for image quality assessment (IQA) [110].

4. According to Chapter 3, some quality-related topics could also be improved such as minutiae selection which could combine both image-based selection and NIMS [108] approaches.

5. Meanwhile, it is also possible to consider whether the type of fingerprint pattern impact on assessing fingerprint quality, for existing studies have observed that the type would impact on the matching performance [15].

6. Nevertheless, the quality of fingerprint sample could also be considered as a feature for detecting spoofing fingerprint image.

Publications

International Conference

1. Z. Yao, J-M Le Bars, C. Charrier, C. Rosenberger. Fingerprint Quality Assessment With Multiple Segmentation. In 2015 International Conference on Cyberworlds (CW) IEEE. Scotland, Sweden. Oct. 7, 2015.

2. Z. Yao, B. Vibert, C. Charrier and C. Rosenberger. Blind Minutiae Selection for Standard Minutiae Templates. IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015). HongKong, China. Mar. 2015.

3. Z. Yao, J.M. Le Bars, C. Charrier and C. Rosenberger. Quality Assessment of Fingerprints with Minutiae Delaunay Triangulation. International Conference on Information Systems Security and Privacy (ICISSP). Angers, France, Feb. 2015.

4. Z. Yao, J.M. Le Bars, C. Charrier and C. Rosenberger. Fingerprint Quality Assessment Combining Blind Image Quality, Texture and Minutiae Features. International Conference on Information Systems Security and Privacy (ICISSP). Angers, France. Feb. 2015.

5. B. Vibert, Z. Yao, S. Vernois, et al. EvaBio Platform for the evaluation biometric system: Application to the optimization of the enrollment process for fingerprint device. International Conference on Information Systems Security and Privacy (ICISSP). Angers, France. Feb. 2015

6. Z. Yao, C. Charrier and C. Rosenberger. Utility Validation of a New Fingerprint Quality Metric. International Biometric Performance Testing Conference (IBPC) 2014. National Institute of Standards and Technology (NIST), Washington, USA. Apr. 2014.

Research Reports

1. Z. Yao, C. Charrier and C. Rosenberger. A Preliminary Study of Fingerprint Quality Assessment of Minutiae Template. 11th Summer School for Advanced Studies on Biometrics for Secure Authentication: Biometrics in Forensics, Security and Beyond. Alghero, Italy. Jun. 2014.

Bibliography

[1] David Zhang. *Automated Biometrics: Technnologies and Systems*. Springer Science & Business Media, 2000. [cité p. 5]

[2] Dengpan Mou. *Machine-based intelligent face recognition*. Springer, 2010. [cité p. 5]

[3] James Wayman, Anil Jain, Davide Maltoni, and Dario Maio. An introduction to biometric authentication systems. In *Biometric Systems*, pages 1–20. Springer, 2005. [cité p. 5, 11, 16]

[4] Ann Cavoukian and Alex Stoianov. Biometric encryption. In *Encyclopedia of Cryptography and Security*, pages 90–98. Springer, 2011. [cité p. 5]

[5] A.K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(1):4–20, 2004. [cité p. 6, 8, 10, 13, 14, 17, 24]

[6] Christophe Rosenberger. Evaluation of biometric systems, 2012. Lecture. [cité p. 6]

[7] Andrea F Abate, Michele Nappi, Daniel Riccio, and Gabriele Sabatino. 2d and 3d face recognition: A survey. *Pattern Recognition Letters*, 28(14):1885–1906, 2007. [cité p. 7]

[8] AnilK. Jain. Biometric recognition: How do i know who you are? In Fabio Roli and Sergio Vitulano, editors, *Image Analysis and Processing ? ICIAP 2005*, volume 3617 of *Lecture Notes in Computer Science*, pages 19–26. Springer Berlin Heidelberg, 2005. [cité p. 7]

[9] Jain Anil K, Nandakumar Karthik, Nagar Abhishek, et al. Biometric template security. *EURASIP Journal on Advances in Signal Processing*, 2008, 2008. [cité p. 9]

[10] YC Feng, Pong C Yuen, and Anil K Jain. A hybrid approach for face template protection. In *SPIE Defense and Security Symposium*, pages 694408–694408. International Society for Optics and Photonics, 2008. [cité p. 9]

[11] Umut Uludag and Anil Jain. Securing fingerprint template: Fuzzy vault with helper data. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pages 163–163. IEEE, 2006. [cité p. 9]

[12] D.L. Woodard, T.C. Faltemier, Ping Yan, K.W. Bowyer, and E. Ortiz. A comparison of 3d biometric modalities. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, pages 57–57, 2006. [cité p. 9]

[13] Arun Ross and Anil Jain. Information fusion in biometrics. *Pattern recognition letters*, 24(13):2115–2125, 2003. [cité p. 10]

[14] Arun Ross and Anil K Jain. Multimodal biometrics: An overview. In *Proceedings of 12th European Signal Processing Conference*, pages 1221–1224, 2004. [cité p. 10, 24]

[15] Salil Prabhakar and Anil K Jain. Decision-level fusion in fingerprint verification. *Pattern Recognition*, 35(4):861–874, 2002. [cité p. 10, 119]

[16] Kevin W Bowyer, Kyong Chang, and Patrick Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition. *Computer vision and image understanding*, 101(1):1–15, 2006. [cité p. 10]

[17] Ping Yan and E. Ortiz. Biometric recognition using 3d ear shape. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(8):1297–1308, 2007. [cité p. 10]

[18] Bao Rong Chang, Hsiu-Fen Tsai, Chi-Ming Chen, and Chien-Feng Huang. Access control of cloud computing using rapid face and fingerprint identification. In *Innovations in Bio-inspired Computing and Applications (IBICA), 2011 Second International Conference on*, pages 179–182, 2011. [cité p. 10]

[19] M. Stojmenovic. Mobile cloud computing for biometric applications. In *Network-Based Information Systems (NBiS), 2012 15th International Conference on*, pages 654–659, 2012. [cité p. 10]

[20] Sharath Pankanti, Ruud M Bolle, and Anil Jain. Biometrics: The future of identification [guest eeditors' introduction]. *Computer*, 33(2):46–49, 2000. [cité p. 10]

[21] Anil K Jain, Arun Ross, and Sharath Pankanti. Biometrics: a tool for information security. *Information Forensics and Security, IEEE Transactions on*, 1(2):125–143, 2006. [cité p. 10, 14]

[22] Christophe Charrier Mohamad El-Abed and Christophe Rosenberger. *Evaluation of Biometric Systems*, chapter 7, page 22. InTech, 2012. [cité p. 11, 15, 85]

[23] Philip C Shinn. System and method of biometric smart card user authentication, December 2 2003. US Patent 6,655,585. [cité p. 15]

[24] EN ISO. 13407: 1999. *Human-centred design processes for interactive systems.* [cité p. 15]

[25] M. El-Abed, R. Giot, B. Hemery, and C. Rosenberger. A study of users' acceptance and satisfaction of biometric systems. In *Security Technology (ICCST), 2010 IEEE International Carnahan Conference on*, pages 170–178, 2010. [cité p. 15]

[26] Simon Liu and M. Silverman. A practical guide to biometric security technology. *IT Professional*, 3(1):27–32, Jan 2001. [cité p. 15]

[27] Patrick Grother and Elham Tabassi. Performance of biometric quality measures. *IEEE Trans. Pattern Anal. Mach. Intell*, pages 531–543, 2007. [cité p. 15, 39, 40, 44, 45, 46, 53, 58, 65, 86, 105]

[28] D Benini et al. Iso/iec 29794-1 biometric quality framework standard, jtc1/sc37/working group 3, jan. 2006. [cité p. 15]

[29] Yi Chen, Sarat Dass, and Anil Jain. Fingerprint quality indices for predicting authentication performance. In *In: Proc. AVBPA, Springer LNCS-3546*, pages 160–170, 2005. [cité p. 15, 32, 38, 44, 58, 82, 86, 107]

[30] P Jonathon Phillips, Alvin Martin, Charles L Wilson, and Mark Przybocki. An introduction evaluating biometric systems. *Computer*, 33(2):56–63, 2000. [cité p. 16, 44, 46]

[31] Raffaele Cappelli, Alessandra Lumini, Dario Maio, and Davide Maltoni. Fingerprint image reconstruction from standard templates. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(9):1489–1503, 2007. [cité p. 16]

[32] Sébastien Marcel. Beat–biometrics evaluation and testing. *Biometric technology today*, 2013(1):5–7, 2013. [cité p. 16]

[33] Weicheng Shen, Marc Surette, and Rajiv Khanna. Evaluation of automated biometrics-based identification and verification systems. *Proceedings of the IEEE*, 85(9):1464–1478, 1997. [cité p. 16]

[34] J A Hanley and B J McNeil. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 143(1):29–36, 1982. PMID: 7063747. [cité p. 18]

[35] Ruud M Bolle, Sharath Pankanti, and Nalini K Ratha. Evaluation techniques for biometrics-based authentication systems (frr). In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 2, pages 831–837. IEEE, 2000. [cité p. 19, 51]

[36] Sinjini Mitra, Marios Savvides, and Anthony Brockwell. Statistical performance evaluation of biometric authentication systems using random effects models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):517–530, 2007. [cité p. 21]

[37] Lawrence O'Gorman. Fingerprint verification. In *Biometrics*, pages 43–64. Springer, 1996. [cité p. 24]

[38] Donald D Rudie. Fingerprint classification by coordinate system, December 31 1968. US Patent 3,419,287. [cité p. 24]

[39] Anil Jain, Lin Hong, and Sharath Pankanti. Biometric identification. *Communications of the ACM*, 43(2):90–98, 2000. [cité p. 24]

[40] Nalini K Ratha, Kalle Karu, Shaoyun Chen, and Anil K Jain. A real-time matching system for large fingerprint databases. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(8):799–813, 1996. [cité p. 24]

[41] Anil K Jain, Lin Hong, Sharath Pankanti, and Ruud Bolle. An identity-authentication system using fingerprints. *Proceedings of the IEEE*, 85(9):1365–1388, 1997. [cité p. 24]

[42] Raffaele Cappelli, Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Fingerprint verification competition 2006. *Biometric Technology Today*, 15(7):7–9, 2007. [cité p. 24]

[43] Davide Maltoni, Dario Maio, Anil K Jain, and Salil Prabhakar. *Handbook of fingerprint recognition*. springer, 2009. [cité p. 25, 26, 27, 28, 106]

[44] Anil Jain, Lin Hong, and Ruud Bolle. On-line fingerprint verification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(4):302–314, 1997. [cité p. 25]

[45] Anil K Jain, Jianjiang Feng, Abhishek Nagar, and Karthik Nandakumar. On matching latent fingerprints. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–8. IEEE, 2008. [cité p. 25]

[46] Arun Ross and Anil Jain. Biometric sensor interoperability: A case study in fingerprints. In *Biometric Authentication*, pages 134–145. Springer, 2004. [cité p. 27]

[47] David Zhang, Feng Liu, Qijun Zhao, Guangming Lu, and Nan Luo. Selecting a reference high resolution for fingerprint recognition using minutiae and pores. *Instrumentation and Measurement, IEEE Transactions on*, 60(3):863–871, 2011. [cité p. 27]

[48] Chris Stein, Claudia Nickel, and Christoph Busch. Fingerphoto recognition with smartphone cameras. In *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the*, pages 1–12. IEEE, 2012. [cité p. 27]

[49] L. Hong, Yifei Wan, and A. Jain. Fingerprint image enhancement: algorithm and performance evaluation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(8):777–789, 1998. [cité p. 27, 35, 94]

[50] S. Greenberg, M. Aladjem, D. Kogan, and I. Dimitrov. Fingerprint image enhancement using filtering techniques. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pages 322–325 vol.3, 2000. [cité p. 27]

[51] Sharat Chikkerur, Alexander N. Cartwright, and Venu Govindaraju. Fingerprint enhancement using stft analysis. *Pattern Recogn.*, 40(1):198–211, January 2007. [cité p. 27, 95]

[52] Xinjian Chen, Jie Tian, Jiangang Cheng, and Xin Yang. Segmentation of fingerprint images using linear classifier. *EURASIP J. Appl. Signal Process.*, 2004:480–494, January 2004. [cité p. 28]

[53] Linlin Shen, Alex Kot, and Waimun Koo. Quality measures of fingerprint images. In *IN: PROC. AVBPA, SPRINGER LNCS-2091*, pages 266–271, 2001. [cité p. 28, 36, 37, 44]

[54] Asker M Bazen and Sabih H Gerez. Segmentation of fingerprint images. In *Proc. Workshop on Circuits Systems and Signal Processing (ProRISC 2001)*, pages 276–280, 2001. [cité p. 28, 37, 94]

[55] Julian Fierrez-Aguilar, Javier Ortega-Garcia, Joaquin Gonzalez-Rodriguez, and Josef Bigun. Kernel-based multimodal biometric verification using quality signals. In *Defense and Security*, pages 544–554. International Society for Optics and Photonics, 2004. [cité p. 28, 32, 39]

[56] Kalle Karu and Anil K Jain. Fingerprint classification. *Pattern recognition*, 29(3):389–404, 1996. [cité p. 29]

[57] HB Kekre and VA Bharadi. Fingerprint core point detection algorithm using orientation field based multiple features. *International Journal of Computer Applications (0975-8887)*, 1(15):106–112, 2010. [cité p. 29]

[58] Xudong Jiang, Manhua Liu, and Alex Chichung Kot. Reference point detection for fingerprint recognition. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 1, pages 540–543. IEEE, 2004. [cité p. 29]

[59] Anil Jain, Yi Chen, and Meltem Demirkus. Pores and ridges: Fingerprint matching using level 3 features. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 4, pages 477–480. IEEE, 2006. [cité p. 29]

[60] Organization for Standardization. Iso/iec 19794-2:2005: Information technology - biometric data interchange formats - part 2: Finger minutiae data, 2005. [cité p. 30, 53, 62]

[61] Rene McIver. Biometric vocabulary, standardization. In StanZ. Li and Anil Jain, editors, *Encyclopedia of Biometrics*, pages 157–160. Springer US, 2009. [cité p. 30]

[62] A.K. Jain, Jianjiang Feng, and K. Nandakumar. Fingerprint matching. *Computer*, 43(2):36–44, 2010. [cité p. 30, 32]

[63] H. Ramoser, B. Wachmann, and H. Bischof. Efficient alignment of fingerprint images. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 3, pages 748–751 vol.3, 2002. [cité p. 30]

[64] Karthik Nandakumar and Anil K. Jain. Local correlation-based fingerprint matching. In *In Indian Conference on Computer Vision, Graphics and Image Processing*, pages 503–508, 2004. [cité p. 31]

[65] Xudong Jiang and Wei-Yun Yau. Fingerprint minutiae matching based on the local and global structures. In *Pattern recognition, 2000. Proceedings. 15th international conference on*, volume 2, pages 1038–1041. IEEE, 2000. [cité p. 31]

[66] Anil K Jain and Jianjiang Feng. Latent fingerprint matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(1):88–100, 2011. [cité p. 32]

[67] Bian Yang, Guoqiang Li, and Christoph Busch. Qualifying fingerprint samples captured by smartphone cameras. In *ICIP*, pages 4161–4165, 2013. [cité p. 32]

[68] Craig I Watson, Michael D Garris, Elham Tabassi, Charles L Wilson, R Michael Mccabe, Stanley Janet, and Kenneth Ko. User's guide to nist biometric image software (nbis). 2007. [cité p. 33, 50, 53, 84]

[69] Elham Tabassi, C Wilson, and C Watson. NIST fingerprint image quality. *NIST Res. Rep. NISTIR7151*, 2004. [cité p. 34, 36, 38, 44, 106]

[70] ISOIEC JTC. 1/sc 37 n 1477: Biometric sample quality standard - part 1: Framework. 2006. [cité p. 34]

[71] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun. A comparative study of fingerprint image-quality estimation methods. *Information Forensics and Security, IEEE Transactions on*, 2(4):734–743, 2007. [cité p. 34, 86, 87, 110]

[72] R. Maarten Bolle, S. Umapatirao Pankanti, and Yi-Sheng Yao. System and method for determining the quality of fingerprint images, Oct 1999. US Patent 5,963,656. [cité p. 36, 37, 94]

[73] B. Lee, J. Moon, and H. Kim. A novel measure of fingerprint image quality using the Fourier spectrum. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. [cité p. 36, 37, 87, 112]

[74] Z Yao, J Le Bars, C Charrier, and C Rosenberger. Quality assessment of fingerprints with minutiae delaunay triangulation. In *International Conference on Information Systems Security and Privacy (ICISSP)*, Feb 2015. [cité p. 36, 94]

[75] Nalini K Ratha and Ruud Bolle. *Fingerprint image quality estimation*. IBM TJ Watson Research Center, 1999. [cité p. 36, 37, 44]

[76] Hartwig Fronthaler, Klaus Kollreider, and Joseph Bigun. Automatic image quality assessment with application in biometrics. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pages 30–30. IEEE, 2006. [cité p. 36, 37]

[77] Loris Nanni and Alessandra Lumini. A hybrid wavelet-based fingerprint matcher. *Pattern Recognition*, 40(11):3146–3151, 2007. [cité p. 36, 61, 87, 111]

[78] Jill D Humphreys, Glenn Porter, and Michael Bell. The quantification of fingerprint quality using a relative contrast index. *Forensic science international*, 178(1):46–53, 2008. [cité p. 36, 37]

[79] Sanghoon Lee, Heeseung Choi, Kyoungtaek Choi, and Jaihie Kim. Fingerprint-quality index using gradient components. *Information Forensics and Security, IEEE Transactions on*, 3(4):792–800, 2008. [cité p. 36, 37]

[80] E. Lim, Xudong Jiang, and WeiYun Yau. Fingerprint quality and validity analysis. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 1, pages I–469–I–472 vol.1, 2002. [cité p. 36, 38, 44, 87, 111, 119]

[81] T.P. Chen, X. Jiang, and W.Y. Yau. Fingerprint image quality analysis. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 2, pages 1253–1256 Vol.2, 2004. [cité p. 36, 38, 44, 87]

[82] Jinqing Qi, Desiree Abdurrachim, Dongju Li, and Hiroaki Kunieda. A hybrid method for fingerprint image quality calculation. In *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*, pages 124–129. IEEE, 2005. [cité p. 36, 38]

[83] Xunqiang Tao, Xin Yang, Yali Zang, Xiaofei Jia, and Jie Tian. A novel measure of fingerprint image quality using principal component analysis(pca). In *Biometrics (ICB), 2012 5th IAPR International Conference on*, pages 170–175, March 2012. [cité p. 36, 38]

[84] M. El Abed, A. Ninassi, C. Charrier, and C. Rosenberger. Fingerprint quality assessment using a no-reference image quality metric. In *European Signal Processing Conference (EUSIPCO)*, page 6, 2013. [cité p. 36, 38, 68, 69, 70, 86, 119, 142]

[85] E. Lim, Kar-Ann Toh, P.N. Suganthan, X. Jiang, and Wei-Yun Yau. Fingerprint image quality analysis. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 2, pages 1241–1244 Vol.2, 2004. [cité p. 36, 38]

[86] Guoqiang Li, Bian Yang, and Christoph Busch. Autocorrelation and dct based quality metrics for fingerprint samples generated by smartphones. In *Digital Signal Processing (DSP), 2013 18th International Conference on*, pages 1–5. IEEE, 2013. [cité p. 36]

[87] M.A. Olsen, E. Tabassi, A. Makarov, and C. Busch. Self-organizing maps for fingerprint image quality assessment. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*, pages 138–145, June 2013. [cité p. 36, 38]

[88] Nalini K Ratha, Shaoyun Chen, and Anil K Jain. Adaptive flow orientation-based feature extraction in fingerprint images. *Pattern Recognition*, 28(11):1657–1672, 1995. [cité p. 35]

[89] Meltem Ballan, F Ayhan Sakarya, and Brian L Evans. A fingerprint classification technique using directional images. In *Signals, Systems & Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*, volume 1, pages 101–104. IEEE, 1997. [cité p. 37]

[90] R-LV Hsu, Jidnya Shah, and Brian Martin. Quality assessment of facial images. In *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the*, pages 1–6. IEEE, 2006. [cité p. 38, 86]

[91] Anil K Jain, Salil Prabhakar, and Shaoyun Chen. Combining multiple matchers for a high security fingerprint verification system. *Pattern Recognition Letters*, 20(11):1371–1379, 1999. [cité p. 38]

[92] Mayank Vatsa, Richa Singh, Afzel Noore, and Max M Houck. Quality-augmented fusion of level-2 and level-3 fingerprint information using dsm theory. *International Journal of Approximate Reasoning*, 50(1):51–61, 2009. [cité p. 38]

[93] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern recognition*, 38(12):2270–2285, 2005. [cité p. 41, 97]

[94] Norman Poh, Josef Kittler, and Thirimachos Bourlai. Quality-based score normalization with device qualitative information for multimodal biometric fusion. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 40(3):539–554, 2010. [cité p. 41]

[95] Raffaele Cappelli, D Maio, and D Maltoni. Sfinge: an approach to synthetic fingerprint generation. In *International Workshop on Biometric Technologies (BT2004)*, pages 147–154, 2004. [cité p. 44]

[96] Z Yao, Jean Marie Le Bars, C Charrier, and C Rosenberger. Fingerprint quality assessment combining blind image quality, texture and minutiae features. In *International Conference on Information Systems Security and Privacy*. INSTICC, 2015. [cité p. 45]

[97] Dario Maio, Davide Maltoni, Raffaele Cappelli, James L. Wayman, and Anil K. Jain. Fvc2000: fingerprint verification competition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(3):402–412, 2002. [cité p. 47]

[98] Romain Giot, Mohamad El-Abed, and Christophe Rosenberger. Fast computation of the performance evaluation of biometric systems: Application to multibiometrics. *Future Gener. Comput. Syst.*, 29(3):788–799, March 2013. [cité p. 52]

[99] Ruud M Bolle, Nalini K Ratha, and Sharath Pankanti. Error analysis of pattern recognition systems - the subsets bootstrap. *Computer Vision and Image Understanding*, 93(1):1–33, 2004. [cité p. 52]

[100] Ted Dunstone and Neil Yager. *Biometric system and data analysis: Design, evaluation, and data mining*. springer, 2008. [cité p. 52]

[101] Dario Maio, Davide Maltoni, Raffaele Cappelli, Jim L Wayman, and Anil K Jain. Fvc2004: Third fingerprint verification competition. In *Biometric Authentication*, pages 1–7. Springer, 2004. [cité p. 53]

[102] M.A. Olsen, Haiyun Xu, and C. Busch. Gabor filters as candidate quality measure for nfiq 2.0. In *Biometrics (ICB), 2012 5th IAPR International Conference on*, pages 158–163, 2012. [cité p. 58, 87, 110]

[103] Stephen M Stigler. Francis galton's account of the invention of correlation. *Statistical Science*, pages 73–79, 1989. [cité p. 58]

[104] C. Charrier C. Rosenberger. Z. YAO, B. Vibert. Blind minutiae selection for standard minutiae templates. In *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015)*, HongKong, China, Mar. 2015. IEEE Biometrics Council. [cité p. 61, 62, 64]

[105] Claude Barral, Jean-Sébastien Coron, and David Naccache. Externalized fingerprint matching. In *Biometric Authentication*, pages 309–315. Springer, 2004. [cité p. 61]

[106] Chen T., Yau W., and Xudong Jiang. On-card matching. In *Encyclopedia of Biometrics*, pages 1014–1021. Springer, 2009. [cité p. 61]

[107] P Grother, W Salamon, C Watson, et al. Minex ii–performance of match on card algorithms phase ii/iii report. nist interagency report 7477, 2010. [cité p. 62]

[108] B. Vibert, J-M Le Bars, C. Charrier, and C. Rosenberger. Comparative study of minutiae selection algorithms for iso fingerprint templates. In *Media Watermarking, Security, and Forensics XVI*, Proceedings of SPIE, February 2015. [cité p. 62, 64, 102, 119]

[109] R. Nock and F. Nielsen. On weighting clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(8):1223–1235, Aug 2006. [cité p. 62]

[110] Anush Krishna Moorthy and Alan Conrad Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. *Image Processing, IEEE Transactions on*, 20(12):3350–3364, 2011. [cité p. 68, 119]

[111] Zhou Wang, Alan C Bovik, and Ligang Lu. Why is image quality assessment so difficult? In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 4, pages IV–3313. IEEE, 2002. [cité p. 68]

[112] Hamid Rahim Sheikh, Muhammad Farooq Sabir, and Alan Conrad Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *Image Processing, IEEE Transactions on*, 15(11):3440–3451, 2006. [cité p. 68]

[113] M. Saad, A. C. Bovik, and C. Charrier. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *IEEE Transactions on Image Processing*, 21(8):3339–3352, 2012. [cité p. 69]

[114] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, 2004. [cité p. 69]

[115] Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang. Fsim: a feature similarity index for image quality assessment. *Image Processing, IEEE Transactions on*, 20(8):2378–2386, 2011. [cité p. 69]

[116] Zhou Wang and Alan C Bovik. Reduced-and no-reference image quality assessment. *Signal Processing Magazine, IEEE*, 28(6):29–40, 2011. [cité p. 69]

[117] Ravi Soundararajan and Alan C Bovik. Video quality assessment by reduced reference spatio-temporal entropic differencing. *Circuits and Systems for Video Technology, IEEE Transactions on*, 23(4):684–694, 2013. [cité p. 69]

[118] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999. [cité p. 70]

[119] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004. [cité p. 70]

[120] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996. [cité p. 70, 71]

[121] Loris Nanni, Alessandra Lumini, and Sheryl Brahnam. Survey on {LBP} based texture descriptors for image classification. *Expert Systems with Applications*, 39(3):3634 – 3641, 2012. [cité p. 70, 71]

[122] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, (6):610–621, 1973. [cité p. 70, 78]

[123] John G Daugman. Complete discrete 2-d gabor transforms by neural networks for image analysis and compression. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(7):1169–1179, 1988. [cité p. 70]

[124] Adel Hafiane and Bertrand Zavidovique. Local relational string for textures classification. In *Image Processing, 2006 IEEE International Conference on*, pages 2157–2160. IEEE, 2006. [cité p. 71, 80]

[125] Thomas Deselaers, Daniel Keysers, and Hermann Ney. Features for image retrieval: an experimental comparison. *Information Retrieval*, 11(2):77–107, 2008. [cité p. 71]

[126] Loris Nanni and Alessandra Lumini. Descriptors for image-based fingerprint matchers. *Expert Systems with Applications*, 36(10):12414–12422, 2009. [cité p. 71]

[127] Matti Pietikäinen. *Computer vision using local binary patterns*, volume 40. Springer, 2011. [cité p. 71]

[128] Loris Nanni and Alessandra Lumini. Local binary patterns for a hybrid fingerprint matcher. *Pattern recognition*, 41(11):3461–3466, 2008. [cité p. 72]

[129] Chengming Wen, Tiande Guo, and Yuyang Zhou. A novel and efficient algorithm for segmentation of fingerprint image based on lbp operator. In *Information Technology and Computer Science, 2009. ITCS 2009. International Conference on*, volume 2, pages 200–204. IEEE, 2009. [cité p. 72]

[130] Bo Zhang, Jie Zhang, Haixia Jiang, and Xiaojun Jing. A new algorithm of incomplete fingerprint segmentation based on lbp. In *Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on*, pages 610–613, 2011. [cité p. 72]

[131] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002. [cité p. 72]

[132] Lior Wolf, Tal Hassner, and Yaniv Taigman. Descriptor based methods in the wild. In *Workshop on Faces in'Real-Life'Images: Detection, Alignment, and Recognition*, 2008. [cité p. 73]

[133] Zhenhua Guo, D. Zhang, and D. Zhang. A completed modeling of local binary pattern operator for texture classification. *Image Processing, IEEE Transactions on*, 19(6):1657–1663, 2010. [cité p. 74]

[134] Timo Ahonen, Jiří Matas, Chu He, and Matti Pietikäinen. Rotation invariant image description with local binary pattern histogram fourier features. In *Image Analysis*, pages 61–70. Springer, 2009. [cité p. 76, 77, 140]

[135] Adel Hafiane, Guna Seetharaman, and Bertrand Zavidovique. Median binary pattern for textures classification. In *Image Analysis and Recognition*, pages 387–398. Springer, 2007. [cité p. 77]

[136] B.S. Manjunath and W.Y. Ma. Texture features for browsing and retrieval of image data. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(8):837–842, 1996. [cité p. 79, 80]

[137] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1):37–52, 1987. [cité p. 82]

[138] J. Feng and A. K Jain. Fingerprint reconstruction: from minutiae to phase. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(2):209–223, 2011. [cité p. 83, 106]

[139] Arun A Ross, Jidnya Shah, and Anil K Jain. Toward reconstructing fingerprints from minutiae points. In *Defense and Security*, pages 68–80. International Society for Optics and Photonics, 2005. [cité p. 83, 85, 106]

[140] David A Stoney. Distribution of epidermal ridge minutiae. *American Journal of Physical Anthropology*, 77(3):367–376, 1988. [cité p. 83]

[141] Jiansheng Chen and Yiu-Sang Moon. The statistical modelling of fingerprint minutiae distribution with implications for fingerprint individuality studies. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–7. IEEE, 2008. [cité p. 83]

[142] Weiwei Zhang and Yangsheng Wang. Core-based structure matching algorithm of fingerprint verification. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 70–74. IEEE, 2002. [cité p. 84]

[143] Xiaolong Zheng, Yangsheng Wang, Xuying Zhao, and Zheng Wei. A scheme for minutiae scoring and its application to fingerprint matching. In *Intelligent Control*

and Automation, 2008. WCICA 2008. 7th World Congress on, pages 5917–5921, June 2008. [cité p. 84]

[144] ISO/IEC. Information technology - biometric sample quality - part 1: Framework, 01 2009. [cité p. 85]

[145] Fernando Alonso-Fernandez, Julian Fierrez, and Javier Ortega-Garcia. Quality measures in biometric systems. *Security & Privacy, IEEE*, 10(6):52–62, 2012. [cité p. 85]

[146] Sharat Chikkerur, Chaohang Wu, and Venu Govindaraju. A systematic approach for feature extraction in fingerprint images. In *Biometric Authentication*, pages 344–350. Springer, 2004. [cité p. 94]

[147] J. Yin, En Zhu, X. Yang, G. Zhang, and C. Hu. Two steps for fingerprint segmentation. *Image and Vision Computing*, 25(9):1391–1403, 2007. [cité p. 94]

[148] Asker M Bazen and Sabih H Gerez. Systematic methods for the computation of the directional fields and singular points of fingerprints. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):905–919, 2002. [cité p. 96]

[149] British Standards Institute Staff. *Information Technology. Biometric Sample Quality. Framework*. B S I Standards, 2009. [cité p. 105]

[150] Salil Prabhakar, Anil K Jain, and Sharath Pankanti. Learning fingerprint minutiae location and type. *Pattern recognition*, 36(8):1847–1857, 2003. [cité p. 106]

[151] Alex M Andrew. Another efficient algorithm for convex hulls in two dimensions. *Information Processing Letters*, 9(5):216–219, 1979. [cité p. 107]

[152] Boris Delaunay. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7(793-800):1–2, 1934. [cité p. 107]

[153] George Bebis, Taisa Deaconu, and Michael Georgiopoulos. Fingerprint identification using delaunay triangulation. In *Information Intelligence and Systems, 1999. Proceedings. 1999 International Conference on*, pages 452–459. IEEE, 1999. [cité p. 108]

[154] Xuefeng Liang, Arijit Bishnu, and Tetsuo Asano. A robust fingerprint indexing scheme using minutia neighborhood structure and low-order delaunay triangles. *Information Forensics and Security, IEEE Transactions on*, 2(4):721–733, 2007. [cité p. 108]

[155] R. Aufmann, V.C. Barker, and R. Nation. *College Trigonometry*. Cengage Learning, 2007. [cité p. 109]

List of Abbreviations

$AFIS$	Automatic Fingerprint Identification System
AUC	Area Under Curve
CI	Confidence Interval
$CLBP$	Completed Local Binary String
DB	Database
EER	Equal Error Rates
ES	Enrollment Selection
FAR	Failure to Acquire Rate
$FLBP$	Four-patch Local Binary String
FMR	False Match Rate
$FNMR$	False Non-Match Rate
FQA	Fingerprint Quality Assessment
FRR	False Reject Rate
FTA	Failure to Acquire
FTC	Failure to Capture
FTE	Failure to Enroll
$GLCM$	Gray-Level Co-occurrence Matrix
GMS	Genuine Matching Scores
IMS	Impostor Matching Scores
LBP	Local Binary Pattern
$LBPHFT$	Local Binary Pattern Histogram Fourier Transform
LRS	Local Relational String
$MLBP$	Median Local Binary String
$NIMS$	No-Image Minutiae Selection
QM	Quality Metric
ROC	Receiver Operation (or Operating) Characteristic
$SEER$	Sample EER

List of Figures

List of Tables

La qualité d'image est un facteur important de systèmes d'identification automatique par empreintes digitales (AFIS) parce que la performance d'appariement pourrait être affectée de manière significative par des échantillons de mauvaise qualité. La qualité des échantillons d'empreintes digitales, par exemple, pourrait être grandement affectée par un grand nombre de problèmes, tels que les défauts physiques et la performance de l'appareil de détection (facteurs externes). En outre, la différence entre les données capturées telles que la spécification de l'image permet également certaines difficultés pour obtenir une solution commune. Les études existantes ont fait des efforts pour connaître l'approche la plus appropriée pour représenter la qualité de l'empreinte digitale. Cette thèse propose plusieurs méthodes d'évaluation de la qualité, tendant à obtenir des observations de cette question par plusieurs aspects: 1) La mesure de qualité permet-elle de prédire la performance de l'appariement? 2) Dans quelle mesure la fusion multi-attributs permet de réaliser une bonne évaluation de la qualité d'empreintes digitales? 3) Est-il possible de se qualifier les empreintes digitales avec seulement son modèle de minuties? et 4) Est-il possible d'évaluer la qualité de l'empreinte digitale via la fusion de plusieurs éléments dans la phase de segmentation? Les approches proposées dans cette thèse sont des réponses à ces questions.

Digital Fingerprint Quality Assessment

Image quality is an important factor to Automated Fingerprint Identification Systems (AFIS) because the matching performance could be significantly affected by poor quality samples. The quality of fingerprint samples, for instance, could be greatly affected by a wide range of problems, such as physical defect and the performance of sensing device (external factors). In addition, the difference between captured data such as image specification also makes some difficulties to achieve a common solution. Existing studies have made efforts to find out more appropriate approach for representing fingerprint quality. This thesis proposes several quality assessment approaches, tending to get observations of this issue in several aspects: 1) Is quality metric really predictive to the matching performance? 2) Whether multi-feature fusion is able to make quality assessment more complemental? 3) Is it possible to qualify fingerprint with its minutiae template alone? and 4) Is it able to assess fingerprint quality via fusing multiple features in segmentation phase? The proposed approaches in this thesis somhow work out answers to these questions.

Indexation Rameau : MOTS CLÉS À VOIR AVEC LA BU
Indexation libre : Fingerprint, Quality, Assessment, Evaluation, Performance, Biometrics, Minutiae, AFIS, Empreinte Digitale, Qualité, Biométrie, Minutie

Spécialité

Laboratoire GREYC - UMR CNRS 6072 - Université de Caen Basse-Normandie - Ensicaen
6 Boulevard du Maréchal Juin - 14050 CAEN CEDEX