To appear in the Proceedings of the 5th International Workshop on Computational Forensics, Tsukuba, Japan, November 11, 2012

On Latent Fingerprint Image Quality

Soweon Yoon, Eryun Liu, and Anil K. Jain

Department of Computer Science and Engineering,
Michigan State University, East Lansing, MI, U.S.A.
{yoonsowo,liueryun,jain}@cse.msu.edu

Abstract. Latent fingerprints which are lifted from surfaces of objects at crime scenes play a very important role in identifying suspects in the crime scene investigations. Due to poor quality of latent fingerprints, automatic processing of latents can be extremely challenging. For this reason, latent examiners need to be involved in latent identification. To expedite the latent identification and alleviate subjectivity and inconsistency in latent examiners' feature markups and decisions, there is a need to develop latent fingerprint identification systems that can operate in the "lights-out" mode. One of the most important steps in "lights-out" systems is to determine the quality of a given latent to predict the probability that the latent can be identified in a fully automatic manner. In this paper, we (i) propose a definition of latent value determination as a way of establishing the quality of latents based on a specific matcher's identification performance, (ii) define a set of features based on ridge clarity and minutiae and evaluate them based on their capability to determine if a latent is of value for individualization or not, and (iii) propose a latent fingerprint image quality (LFIQ) that can be useful to reject the latents which cannot be successfully identified in the "lights-out" mode. Experimental results show that the most salient latent features include the average ridge clarity and the number of minutiae. The proposed latent quality measure improves the rank-100 identification rate from 69% to 86% by rejecting 50% of latents deemed as poor quality. In addition, the rank-100 identification is 80% when rejecting 80% of the latents in the databases assessed as 'NFIQ = 5'; however, the same identification rate can be achieved by rejecting only 21% of the latents with low LFIQ.

Keywords: Latent fingerprints, "lights-out" latent identification, latent value determination, latent fingerprint image quality (LFIQ)

1 Introduction

Fingerprints have been widely used for reliable human identification in forensics and law enforcement applications for over a century. Law enforcement agencies routinely collect tenprint records of all apprehended criminals in two forms: *rolled* and *plain* (or *slap*). Rolled fingerprints are obtained by rolling a finger from nail to nail to capture the complete ridge details on a finger while plain fingerprints are captured by pressing down a finger on the flat surface of a fingerprint card

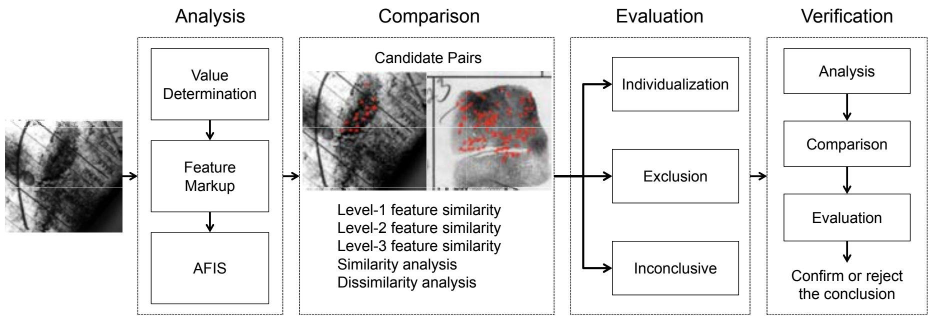


Fig. 1. ACE-V methodology in latent identification.

or a live-scan sensor [1]. Both rolled and plain fingerprints, collectively called reference (exemplar) fingerprints, are believed to contain sufficient ridge details for individualization. To ensure good quality of fingerprint images, reference fingerprints are acquired under the supervision of a human operator; this way, fingerprints can be recaptured in case of poor quality impressions.

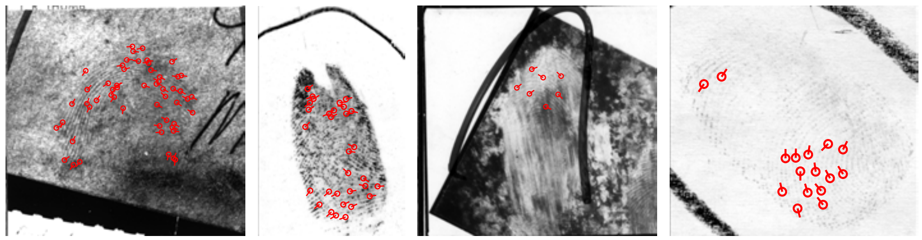
Automated Fingerprint Identification Systems (AFIS) are used by almost every major law enforcement agency worldwide to facilitate tenprint search; rolled or plain fingerprints are submitted as a query to an AFIS to search large-scale reference fingerprint databases. The FBI's Integrated AFIS (IAFIS) receives tens of thousands of requests everyday to search a reference database containing tenprint cards of over 72 million criminals and 34 million civilian job applicants [2]. The matching accuracy of tenprint search has already reached an impressive level; the 2003 Fingerprint Vendor Technology Evaluation (FpVTE) [3] reported that the best performing commercial matcher achieved 99.4% true accept rate at 0.01% false accept rate in searching plain fingerprints against a reference database with 10,000 plain fingerprints.

Another type of important fingerprint identification involves searching *latent* fingerprints against reference fingerprint databases. Latent fingerprints (or simply *latents*) refer to the fingerprints captured at crime scenes, and are regarded as an extremely important source of evidence in crime scene investigations to identify suspects. Unlike rolled or plain fingerprints, latents are often of poor quality; latent impressions typically contain partial ridge patterns of a finger, incomplete or missing ridge structures, mixture of ridge pattern and complex background noise or friction ridge structures from other fingers. Due to low quality of latent fingerprints, human intervention is inevitable in latent search, especially for feature markup (i.e., manually marking region of interest, minutiae, core, delta, and extended features). In matching latents to reference prints, latent fingerprint examiners are expected to follow a methodology, called Analysis, Comparison, Evaluation and Verification (ACE-V) [4]. Fig. 1 illustrates the ACE-V methodology. In the analysis phase, an examiner evaluates the ridge information contained in latent images. If the latent is determined to contain sufficient information for identification or exclusion (called "of value" latent), the features in the latent are manually marked by the examiner to search for its mate using an AFIS. In the comparison phase, the examiner compares the "of value" latent with the candidate mates retrieved from the reference database side-by-side and ascertains the similarity between the latent and mated reference print pairs using feature markup in the latent. In the evaluation phase, one of the following decisions is made about the latent in question: individualization, exclusion or inconclusive[1]. Finally, in the verification phase, the decision made by the first examiner is confirmed by having a second examiner analyze the results independently.

Although the ACE-V methodology is widely accepted by forensic community for latent print examination, the influence of human factors in the ACE-V procedure has raised concerns about their reliability and consistency. A noteworthy case is the erroneous identification of Brandon Mayfield as a suspect in the Madrid train bombing incident based on an incorrect match between Mayfield's reference fingerprint and the latent print captured at the bombing site [6, 7]. The National Research Council's report on limitations and recommendations of forensic science [8] pointed out two major shortcomings in the current forensic science discipline: (i) "lack of mandatory and enforceable standards" that can be globally referred to in crime labs and (ii) "unacceptable case backlogs in state and local crime labs which likely make it difficult for laboratories to provide strong evidence for prosecutions and avoid errors that could lead to imperfect justice". Along with the efforts to understand the human factors in latent fingerprint examination [9], standards and guidelines for latent examiners' practices have also been set up. As an example, the Science Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST) published standards which define terminologies and establish the sufficiency level for decisions at each step of the ACE-V methodology to alleviate subjectivity involved in feature markups and decision makings among examiners [5].

Based on the guidelines in SWGFAST standard, latent examiners' practices have been evaluated from various aspects (e.g., reliability of decisions, degree of consensus and consistency of decisions) [10, 11, 12], mainly on two critical decisions that the examiners make in ACE-V methodology: (i) latent value determination in the analysis phase and (ii) latent individualization conclusion in the evaluation phase. Latent value determination assigns one of the following labels to each latent: value for individualization (VID), value for exclusion only (VEO), and no value (NV). Only VID and VEO latents are further considered for comparison. However, recent studies on the consensus and consistency of latent examiners' value determination indicate that a significant amount of

[1] Individualization is the decision that a latent examiner makes on a pair of latent and a reference print indicating that the pair originates from the same finger based on a sufficient agreement between the two ridge patterns. Exclusion, on the other hand, is the decision where an examiner concludes that the pair did not originate from the same finger based on a sufficient disagreement between the two ridge patterns. An inconclusive decision is made when an examiner cannot make a decision of either individualization or exclusion due to insufficient ridge details or small corresponding area between latent and reference print [5].



(a) (b) (c) (d)

Fig. 2. Latents that are assessed as either VEO or NV, but the mated rolled prints are retrieved at rank-1 by AFIS. (a) NIST SD27 (U237): VEO (minutiae from AFIS) and (b) WVU (W514F02C_N): VEO (minutiae from AFIS), (c) NIST SD27 (U254): NV (minutiae marked by an examiner), and (d) WVU (W605F01C_I): VEO (minutiae marked by an examiner).

subjectivity and variation exists in latent value determination. In [11], the consensus of the latent examiners' value determinations was evaluated; among the 356 latents which were reviewed by the examiners (each latent was reviewed by 23 examiners, on average), unanimous decisions (either VID or not-VID) were made only for about 43% of the latents. In terms of consistency, Ulery *et al.* [12] found that 85% of NV decisions and 93% of VID decisions were repeated by the same examiner after a time gap while only 55% of VEO decisions were repeated. More importantly, an examiner's value determination is not always reliable; the ELFT-EFS (Evaluation of Latent Fingerprint Technologies: Extended Feature Sets) report [13] emphasizes that a significant portion of the latents assessed as being of VEO or NV can still be successfully identified by AFIS. Fig. 2 shows examples of latents in the NIST SD27 database [14] and the WVU latent database [15] that were assessed as either VEO or NV by latent examiners, but whose mated reference prints were correctly retrieved at the top rank by an AFIS.

One of the most desirable properties of a latent fingerprint identification system is its ability to process latents in a "lights-out" identification mode with high accuracy. A "lights-out" fingerprint identification system refers to a system that requires only fingerprint images as input (query) and returns a short list of reference prints as potential mates [16]. The advantages of "lights-out" latent identification mode include: (i) avoiding subjectivity in latent print examination and (ii) increasing throughput of latent print matching, given the growing workload on latent examiners. While state-of-the-art latent fingerprint identification systems have already shown excellent performance in matching latents to reference databases[2] [13,17], more research efforts are needed to continue to improve this performance.

[2] The best performing matcher for latent search in the ELFT-EFS achieved 63.4% rank-1 identification rate in the "lights-out" identification mode [13].

The first and the most significant step in achieving "lights-out" latent identification capability is the quality assessment of latents to (i) determine whether or not they have sufficient fingerprint ridge structure for either automatic feature extraction or manual feature markup and (ii) predict the reliability of latent to reference print matching with the given feature set. Although, in principle, the objectives of latent quality assessment are the same as those of tenprint or reference print quality assessment, defining a latent fingerprint image quality (LFIQ) measure is lot more difficult. Further, while the tenprint quality measures are mainly used to determine if they need to be recaptured, the main purpose of defining a latent quality measure is to expedite latent print examination by identifying 'good' quality latents which can be processed in "lights-out" mode. An appropriate latent quality measure will also help avoid potential erroneous subjective decisions regarding latent value.

Latent fingerprint quality can be assessed as latent value determination, and the latent value can be determined by two different ways: (i) latent examiners and (ii) AFIS. Latent value determination by examiners simply refers to the current practice of assigning one of the following values (i.e., VID, VEO and NV) to a latent by an examiner. However, as mentioned before, the latent values determined by examiners are not always correlated to the latent identification performance. Considering that one of the major goals of defining latent quality measure is to predict latent matching performance (assuming that the mated print in the reference database is of fairly good quality), we define a value determination by AFIS, following the comparison protocol in the ACE-V methodology: a latent fingerprint is declared to be VID in value determination by an AFIS if its mated print in the reference database is retrieved within the top rank-100; otherwise, the latent is declared as not-VID in value determination by the AFIS.

One of the challenges in defining a value determination by AFIS is to find salient features in latents which are directly related to the latent identification performance. Features that can be used in estimating latent fingerprint quality are generally comprised of (i) quality for value and (ii) quality for identification. Quality for value (often referred to as *qualitative* quality [10, 18]) measures the sufficiency of Level-1 features such as fingerprint ridge clarity, pattern class, size of region of interest, fingerprint position (e.g., center, side, or tip of a finger), and determines the utility of the latent for either exclusion or identification. Quality for identification (often referred to as *quantitative* quality [10, 18]) measures the sufficiency of Level-2 (i.e., minutiae) and Level-3 (e.g., pores) features which are directly used in fingerprint identification, and determines the reliability of matching results by an AFIS in the "lights-out" mode.

Another important consideration in defining a latent quality measure is whether it should depend on a specific AFIS. The issue of interoperability of AFIS in latent search is well known [19]. Given that the matching results of AFIS from different vendors on a given latent print can be significantly different as revealed in ELFT-EFS [13], we believe that a latent fingerprint quality measure should be designed for a specific AFIS. We call this a matcher-dependent latent quality measure.

In this paper, we report results of our preliminary study on defining LFIQ. First, we investigate a set of Level-1 features for qualitative quality and Level-2 features for quantitative quality that can be effectively used to define latent fingerprint image quality. The features are evaluated by constructing a 2-class classifier to determine whether a latent is VID or not-VID. The class labels (VID and not-VID) are determined either by latent examiners or an AFIS; this way, two different classifiers – one for value determination by examiners and the other for value determination by an AFIS – are obtained. Based on the observations from latent value estimation, we define a latent quality measure by combining a qualitative quality feature (i.e., the average ridge clarity) and a quantitative quality feature (i.e., the number of minutiae) to estimate the objective target quality: the probability that the mated reference print of a latent will be retrieved within the top rank-100.

2 Latent Fingerprint Image Quality (LFIQ)

To find the most significant features to represent latent fingerprint quality, a latent quality assessment is viewed as a 2-class classification problem: ω_{VID} versus $\omega_{\overline{VID}}$, where ω_{VID} represents the class of VID latents and $\omega_{\overline{VID}}$ represents the class of not-VID latents. Feature vectors with different composition of ridge clarity features and minutiae features are evaluated in terms of the resulting classification accuracy. Based on the empirical results, a latent quality measure is defined by combining the two most significant features: average ridge clarity in the convex hull enclosing all the minutiae and the number of minutiae.

2.1 Ridge Clarity Feature

Given a latent fingerprint image I, the quality of its local ridge structure is measured by the ridge strength in a local block and the ridge continuity in the block's neighborhood. The computation of local ridge clarity map, **RC**, involves the following steps:

1. Preprocessing: The contrast-enhanced latent image, I^*, is obtained by [20]:

$$I^* = \mathrm{sign}(I - \bar{I}) \times \log(1 + |I - \bar{I}|), \tag{1}$$

   where $\bar{I}$ is the smoothed version of I by applying a 15×15 averaging filter, and sign(x)= 1 if $x > 0$, otherwise sign(x)= -1.
2. Fourier analysis: The contrast-enhanced image I^* is divided into 16×16 pixel blocks. The 64×64 subimage, $I^*_{mn}(x, y)$, is constructed by taking a 32×32 subimage of I^* around the center of the block at $[m, n]$ and padding with 0's to get high frequency resolution in the Fourier domain. The subimage is transformed into the Fourier domain, $F_{mn}(u, v)$. The top two local amplitude maxima within the frequency range of $\left[\frac{1}{16}, \frac{1}{5}\right]$ are selected [21]. Let (u_1, v_1) and (u_2, v_2) be the locations of the first and the second amplitude maxima in

$F_{mn}(u,v)$. Then, a 2-dimensional sine wave corresponding to the i-th local maximum of the block at $[m,n]$ can be written as:

$$w_{mn}^{(i)}(x,y) = a_{mn}^{(i)} \sin(2\pi f_{mn}^{(i)}(\cos(\theta_{mn}^{(i)})x + \sin(\theta_{mn}^{(i)})y) + \phi_{mn}^{(i)}),\ i = 1,2, \quad (2)$$

where

$$a_{mn}^{(i)} = |F(u_i,v_i)|,\ f_{mn}^{(i)} = \frac{\sqrt{u_i^2+v_i^2}}{64}$$

$$\theta_{mn}^{(i)} = \arctan\left(\frac{u_i}{v_i}\right),\ \text{and}\ \phi_{mn}^{(i)} = \arctan\left[\frac{Im(F(u_i,v_i))}{Re(F(u_i,v_i))}\right],$$

$a_{mn}^{(i)}$, $f_{mn}^{(i)}$, $\theta_{mn}^{(i)}$, and $\phi_{mn}^{(i)}$ represent the amplitude, frequency, direction, and phase, respectively.

3. Ridge continuity map: The 2-dimensional sine waves, w_1 and w_2, in two adjacent blocks are continuous if they satisfy the following conditions [21]:

$$\begin{aligned} &\min\{|\theta_1,\theta_2|, \pi - |\theta_1,\theta_2|\} \le T_\theta, \\ &\left|\frac{1}{f_1} - \frac{1}{f_2}\right| \le T_f,\ \text{and} \\ &\frac{1}{16}\sum_{(x,y)\in\mathcal{L}} \left|\frac{w_1(x,y)}{a_1} - \frac{w_2(x,y)}{a_2}\right| \le T_p, \end{aligned} \quad (3)$$

where $\mathcal{L}$ denotes the 16 pixels on the border of the two adjacent blocks, and the three thresholds, T_θ, T_f, and T_p, are set to $\frac{\pi}{10}$, 3, and 0.6, respectively. Define an indicator function, $I_c(w_1,w_2)$, as follows:

$$I_c(w_1,w_2) = \begin{cases} 1, & \text{if } w_1 \text{ and } w_2 \text{ are continuous,} \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Ridge continuity map is defined by:

$$\widetilde{\mathbf{RC}}[m,n] = \sum_{[m^*,n^*]\in\mathcal{N}} \max\{I_c(w_{mn}^{(1)}, w_{m^*n^*}^{(1)}), I_c(w_{mn}^{(1)}, w_{m^*n^*}^{(2)})\}, \quad (5)$$

where $\mathcal{N}$ is the 8-neighborhood blocks of the block $[m,n]$.

4. Ridge clarity map: The ridge clarity of block at $[m,n]$ is defined by:

$$\mathbf{RC}[m,n] = a_{mn}^{(1)} \cdot \widetilde{\mathbf{RC}}[m,n]. \quad (6)$$

Fig. 3 shows the ridge clarity maps and the intermediate steps for two different latents. Well-defined ridge structures present in a latent correspond to the high ridge clarity regions in the **RC** map.

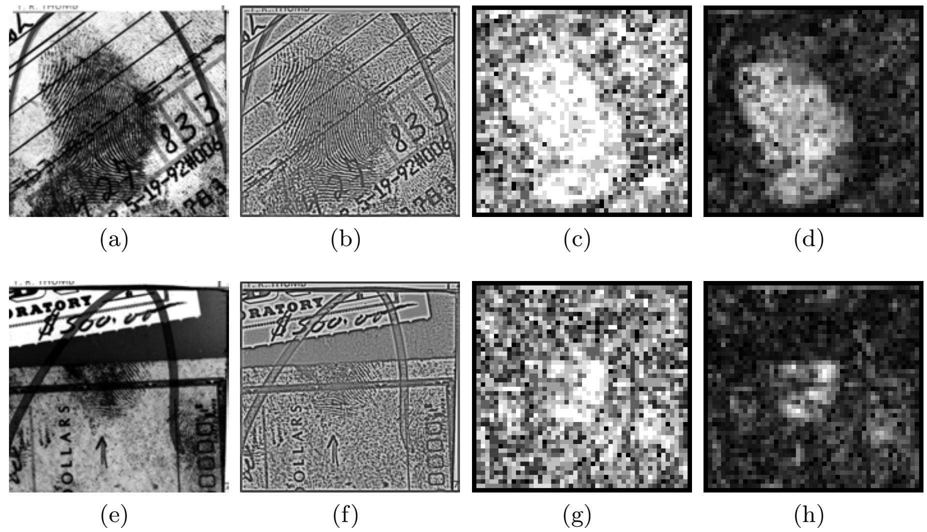


Fig. 3. Ridge clarity maps (**RC**) of two latents in NIST SD27. (a) G051 and (e) G080, (b) and (f) contrast-enhanced latent images, (c) and (g) ridge continuity maps, and (d) and (h) ridge clarity maps (brighter regions indicate higher ridge clarity).

2.2 Minutiae Feature

The most significant features which are directly related to the fingerprint matching performance are the minutiae properties. We investigated a number of different minutiae properties, including the number of minutiae, quality of minutiae based on the **RC** map defined in the Section 2.1, and size of the convex hull enclosing the minutiae in the latent. Let N_M, Q_M, and A_M denote the number of minutiae, average quality of minutiae, and size of the convex hull enclosing all the minutiae in a latent.

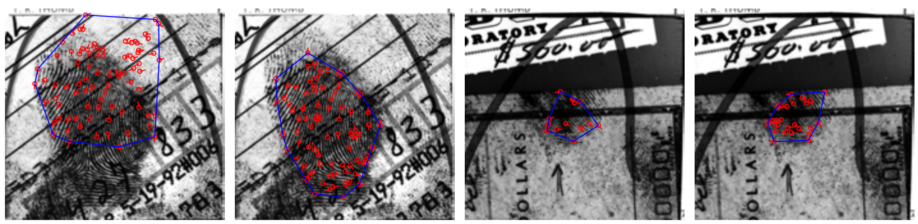
The average quality of minutiae, Q_M, is obtained from the **RC** map as follows:

$$Q_M = \frac{1}{N_M} \sum_{i=1}^{N_M} \mathbf{RC}[bx_i, by_i], \tag{7}$$

where $[bx_i, by_i]$ is the blockwise position of the i-th minutia.

Two different sets of minutiae are considered: (i) minutiae automatically extracted by an AFIS and (ii) minutiae manually marked by latent examiners. Reliability of the minutiae marked by examiners is significantly higher compared to the minutiae extracted by the AFIS[3] while the consistency of the minutiae markup by the examiners is lower than that of the minutiae extracted by the

[3] The AFIS used in this study is not a state-of-the-art latent-to-reference print matcher, but instead a state-of-the-art AFIS for reference fingerprint matching. Currently, no AFIS for latent matching is available to us.



(a) (b) (c) (d)

Fig. 4. Minutiae in two latents shown in Fig. 3. (a) Minutiae extracted by an AFIS for the latent in Fig. 3a, (b) minutiae marked by an examiner for the latent in Fig. 3a, (c) minutiae extracted by the AFIS for the latent in Fig. 3e, and (b) minutiae marked by an examiner for the latent in Fig. 3e.

AFIS. Fig. 4 shows the two minutiae sets – minutiae extracted by the AFIS and minutiae marked by examiner – of the two latents in Fig. 3 along with the convex hulls enclosing the minutiae.

2.3 Latent Fingerprint Quality Assessment

To establish the most discriminant features to define latent fingerprint quality, the latent quality assessment is formulated as a 2-class classification problem for ω_{VID} and $\omega_{\overline{VID}}$. To build the "best" performing matcher for latents[4], we fuse a state-of-the-art AFIS for tenprint with the latent fingerprint matcher developed in [22] as follows. The match score of a latent i and a reference print j is obtained by [23]:

$$s_{ij} = r\frac{s_{ij}^A}{\frac{1}{N_R}\sum_{k=1}^{N_R} s_{ik}^A} + (1-r)\frac{s_{ij}^B}{\frac{1}{N_R}\sum_{k=1}^{N_R} s_{ik}^B},\ 0 \le r \le 1, \tag{8}$$

where s_{ij}^A and s_{ij}^B are the match scores of a latent i and a reference print j from matcher A and B, respectively, N_R is the size of the reference database, and r is a weight which is empirically chosen to obtain the best rank-100 identification accuracy.

We evaluated feature vectors consisting of different combinations of ridge clarity and minutiae features. The following two feature vectors were observed to be the most discriminative:

$$\mathbf{x_1} = (N_M),$$
$$\mathbf{x_2} = (Q_R, N_M, Q_M, A_M),$$

[4] Based on the latent matching performance evaluation with the fingerprint matchers available to us, the fusion of the two matchers described in this paper showed the best performance to simulate the performance of a state-of-the-art AFIS for latents.

where Q_R is the average ridge clarity in the convex hull of the minutiae in a latent and defined as:

$$Q_R = \frac{1}{|\mathcal{C}|} \sum_{[m,n] \in \mathcal{C}} \mathbf{RC}[m,n], \tag{9}$$

where $\mathcal{C}$ is the set of blocks that belong to the convex hull enclosing the minutiae and $|\mathcal{C}|$ is the number of blocks in the convex hull. Note that each component of the feature vector is normalized by its minimum and maximum values. A decision tree classifier [24] is used to determine whether a latent belongs to ω_{VID} or $\omega_{\overline{VID}}$.

Based on the above observations, a latent fingerprint image quality measure, $LFIQ$, is defined as:

$$LFIQ = Q_R \cdot N_M, \tag{10}$$

where Q_R is the average ridge clarity and N_M is the number of minutiae.

3 Experimental Results

3.1 Databases

The latent quality measure was evaluated on two latent fingerprint databases: NIST SD27 [14] and WVU latent database [15]. NIST SD27, the only public-domain latent database, contains 258 latents obtained from operational case-work. Although all the latents in this database are labeled as belonging to one of three quality levels (‘Good’, ‘Bad’, and ‘Ugly’), we follow a formal definition of value determination by latent examiners (i.e., VID, VEO, and NV) which was reported in [10]. The numbers of VID, VEO, and NV latents in NIST SD27 are 210, 41, and 7, respectively.

The WVU database contains 449 latents collected in a laboratory environment. These latents come with value determination and feature markups. The numbers of VID, VEO, and NV latents in the WVU database are 370, 74, and 5, respectively.

As a reference database, the mated rolled fingerprints of the 707 latents (258 from NIST SD27 and 449 from WVU database) are combined with 4,291 additional rolled prints from the WVU database and 27,000 rolled prints of the first impressions in NIST SD14 [25]. In total, the size of the reference database is 31,998.

3.2 Classification Accuracy of Latent Value Determination

The performance of the classifiers using feature vectors, $\mathbf{x_1}$ and $\mathbf{x_2}$ defined in Section 2.3, to determine whether a given latent is VID or not-VID was evaluated. Two different sources of minutiae set are used to construct feature vectors: (i) minutiae extracted by an AFIS and (ii) minutiae marked by examiners. A 10-fold cross validation is conducted for evaluation: two latent databases are evenly partitioned into 10 sets, and one set from each database is selected at random

Table 1. Confusion matrix of ω_{VID} versus $\omega_{\overline{VID}}$ classification from value determination by examiners. $\hat{\omega}_{VID}$ and $\hat{\omega}_{\overline{VID}}$ are the predicted class labels. Feature vectors are defined based on (a) minutiae extracted by AFIS and (b) minutiae marked by examiners.

(a)

Feature	$\mathbf{x_1}$		$\mathbf{x_2}$	
	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$
ω_{VID}	565	14	496	83
$\omega_{\overline{VID}}$	128	0	91	37

(b)

Feature	$\mathbf{x_1}$		$\mathbf{x_2}$	
	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$
ω_{VID}	564	15	523	56
$\omega_{\overline{VID}}$	69	59	64	64

Table 2. Classification accuracy (standard deviation) for ω_{VID} and $\omega_{\overline{VID}}$ from value determination by examiners. Feature vectors are defined based on (a) minutiae extracted by AFIS and (b) minutiae marked by examiners.

(a)

Feature	$\mathbf{x_1}$	$\mathbf{x_2}$
ω_{VID} classification accuracy	98% (2%)	86% (4%)
$\omega_{\overline{VID}}$ classification accuracy	0% (0%)	29% (19%)
Total classification accuracy	80% (1%)	75% (4%)

(b)

Feature	$\mathbf{x_1}$	$\mathbf{x_2}$
ω_{VID} classification accuracy	97% (3%)	90% (5%)
$\omega_{\overline{VID}}$ classification accuracy	46% (12%)	50% (10%)
Total classification accuracy	88% (2%)	83% (4%)

without replacement to form a fold. Average classification accuracies along with standard deviations are reported in Tables 2 and 4. Target class label for ω_{VID} and $\omega_{\overline{VID}}$ can be assigned by either (i) latent examiners or (ii) AFIS.

Value Determination by Examiners Value determination by latent examiners is one of VID, VEO, and NV. All latents determined as VID by examiners comprise ω_{VID} while $\omega_{\overline{VID}}$ consists of VEO and NV latents determined by examiners. Tables 1 and 2 show the confusion matrices and classification accuracies with standard deviations of predicting value determination by examiners.

Value Determination by AFIS Value determination by AFIS is based on the matching performance of a specific AFIS: a latent belongs to ω_{VID} if its mate in the reference database is retrieved within the top rank-100; otherwise, it belongs to $\omega_{\overline{VID}}$. Tables 3 and 4 show the confusion matrices and classification accuracies with standard deviations of predicting value determination by the AFIS.

Based on these classification results, we make the following observations:

Table 3. Confusion matrix for ω_{VID} versus $\omega_{\overline{VID}}$ classification from value determination by AFIS. $\hat{\omega}_{VID}$ and $\hat{\omega}_{\overline{VID}}$ are the predicted class labels. Feature vectors are defined based on (a) minutiae extracted by AFIS and (b) minutiae marked by examiners.

(a)

Feature	$\mathbf{x_1}$		$\mathbf{x_2}$	
	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$
ω_{VID}	162	157	195	124
$\omega_{\overline{VID}}$	126	262	125	263

(b)

Feature	$\mathbf{x_1}$		$\mathbf{x_2}$	
	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$	$\hat{\omega}_{VID}$	$\hat{\omega}_{\overline{VID}}$
ω_{VID}	456	34	391	99
$\omega_{\overline{VID}}$	117	100	107	110

Table 4. Classification accuracy (standard deviation) for ω_{VID} and $\omega_{\overline{VID}}$ from value determination by AFIS. Feature vectors are defined based on (a) minutiae extracted by AFIS and (b) minutiae marked by examiners.

(a)

Feature	$\mathbf{x_1}$	$\mathbf{x_2}$
ω_{VID} classification accuracy	51% (12%)	61% (7%)
$\omega_{\overline{VID}}$ classification accuracy	67% (11%)	68% (8%)
Total classification accuracy	60% (4%)	65% (4%)

(b)

Feature	$\mathbf{x_1}$	$\mathbf{x_2}$
ω_{VID} classification accuracy	93% (4%)	80% (5%)
$\omega_{\overline{VID}}$ classification accuracy	46% (13%)	51% (9%)
Total classification accuracy	79% (6%)	71% (3%)

- Feature vectors from the minutiae marked by examiners show better performance than those from the minutiae extracted by the AFIS due to the high reliability of the markup features.
- Value determination by examiners shows higher classification accuracy than value determination by AFIS. This is because the value determination by examiners is done by looking at only latent images so that the quality and the sufficiency of features in the latents are closely related to the value determination. On the other hand, the value determination by an AFIS also depends on the quality of mated reference prints. This emphasizes the need for considering the quality of the mates as an independent factor when developing a latent fingerprint quality measure.
- Feature vector $\mathbf{x_1}$ generally shows better performance than feature vector $\mathbf{x_2}$ to predict latent values, except for value determination by AFIS with minutiae extracted by the AFIS. When the reliability of feature extraction is low, the ancillary information on fingerprint ridge quality helps to design a better latent quality measure.

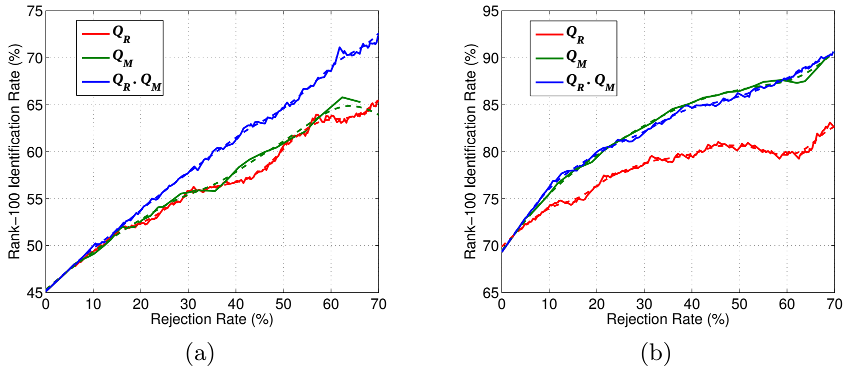


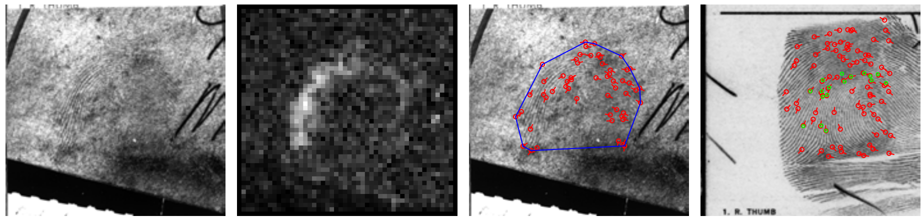
Fig. 5. Rank-100 identification rate versus rejection rate when the minutiae are obtained by (a) an AFIS and (b) latent examiners (solid lines). Dotted lines are the approximation with cubic functions.

3.3 Prediction of Latent Identification Performance by LFIQ

To evaluate the proposed latent quality measure, $LFIQ$, as a predictor of latent identification performance, we consider rank-100 identification rate when the latents with low quality scores are rejected. Fig. 5 shows the rank-100 identification rate as a function of the rejection rate based on $LFIQ$ when two different minutiae sets (one from examiners and the other from AFIS) are used. Based on these plots, we make the following observations:

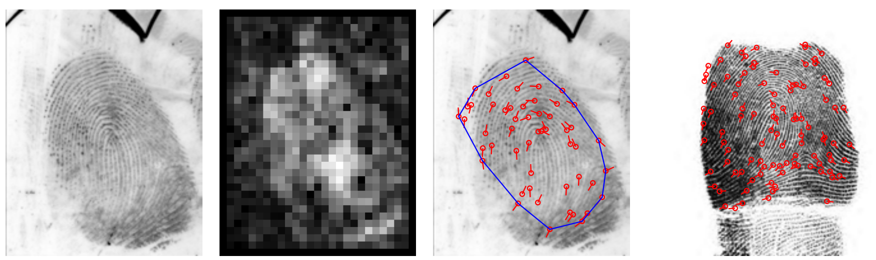
- Quality measure based on the minutiae marked by examiners shows better performance than the quality measure based on the minutiae extracted by the AFIS. This implies that, if highly reliable feature extractors are available, the quality estimation using relatively simple features such as ridge clarity and the number of minutiae will suffice in predicting the latent matching performance.
- For minutiae extracted by an AFIS, the quality measure $LFIQ$ which combines Q_R and N_M shows the best performance compared to either Q_R or N_M alone.
- For minutiae marked by examiners, however, N_M alone shows about the same performance as $LFIQ$, which implies that the amount of information that minutiae contain (i.e., the number of minutiae in the latent) is sufficient to predict the matching performance when the extracted minutiae are highly reliable.

Fig. 6 shows an example where the proposed quality measure $LFIQ$ successfully predicts the latent identification performance. Fig. 7 shows a latent and its mated reference print of good quality, but the latent identification performance is poor due to the large number of missing minutiae and spurious minutiae extracted by the AFIS in the latent even in the high quality ridge region.



(a) (b) (c) (d)

Fig. 6. Example of a latent for which the proposed latent quality $LFIQ$ successfully determines it to be 'high' quality latent. (a) Latent U237 in NIST SD27, (b) ridge clarity map, (c) minutiae set and its convex hull, and (d) mated reference print with minutiae extracted by an AFIS (red) and the mated minutiae with the latent (green). Note that while this latent has a high $LFIQ$ value (corresponding to the rejection rate of 83%) and the mated print is retrieved at rank-1, the value determination by examiners for this latent was not-VID.



(a) (b) (c) (d)

Fig. 7. Example of a latent where the proposed latent quality measure $LFIQ$ determines it to be a 'high' quality latent, but the retrieval rank of the mated reference print is poor (ranked at 4,658). (a) Latent W944F08B_N in WVU database, (b) ridge clarity map, (c) minutiae set and its convex hull, and (d) mated reference print with minutiae extracted by an AFIS. Note that while this latent has a high $LFIQ$ value (corresponding to the rejection rate of 88%) and the value determination by examiners is also VID, the matching performance is poor due to the large number of false minutiae extracted by the AFIS in the latent.

As a comparison to the proposed LFIQ, NIST Fingerprint Image Quality (NFIQ)[5] [26] was used to assess the quality of the latents in the two latent databases. When setting the rejection criterion as 'NFIQ = 5', the rank-100 identification rate of the accepted latents with minutiae extracted by the AFIS was 66% at a rejection rate of 80%. At the same rejection rate, the rank-100 identification rate of the accepted latents by NFIQ with minutiae marked by

[5] NFIQ assigns one of five discrete quality levels ranging from 1 to 5 to a reference print; '1' refers to the highest quality, and '5' indicates the lowest quality. Note that NFIQ was not designed for latent fingerprint quality assessment.

examiners was 80%. The proposed latent quality measure, $LFIQ$, on the other hand, was able to achieve the same rank-100 identification rate at significantly lower rejection rates as follows: (i) rank-100 identification rate of 66% can be achieved, with minutiae extracted by the AFIS, by rejecting 53% of poor quality latents identified by the proposed $LFIQ$; (ii) rank-100 identification rate of 80% can be achieved, with minutiae marked by examiners, by rejecting 21% of poor quality latents identified by the proposed $LFIQ$.

4 Conclusions and Future Work

Towards the goal of designing a "lights-out" latent fingerprint matching system, we proposed a latent fingerprint quality measure. We first defined the value of a latent based on its rank-100 identification rate to directly relate the latent quality to the matching performance. A set of features based on latent fingerprint ridge clarity and minutiae properties was evaluated by posing the latent value determination problem as a 2-class classification problem (i.e., VID versus not-VID). Based on the classification accuracy, the most salient features were selected to define a latent quality measure as a product of the average ridge clarity in the convex hull enclosing the minutiae and the number of minutiae in the latent. The proposed latent quality measure improves the rank-100 identification rate of an AFIS by effectively rejecting poor quality latents.

Based on the preliminary study of the proposed LFIQ, we identify the following topics that are worthy of further consideration:

- Defining a latent quality measure with a more discriminative feature set representing the reliability of minutiae extracted by an AFIS to predict its matching performance.
- Developing a latent quality measure involving Level-3 features.

Acknowledgments. We would like to thank Austin Hicklin of Noblis for providing us the value determination by examiners of latents in NIST SD27. This research was partially supported by a grant from the NSF Center of Identification Technology Research (CITeR).

References

1. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition (Second Edition). Springer-Verlag (2009)
2. The Federal Bureau of Investigation (FBI). Integrated Automated Fingerprint Identification System (IAFIS) `http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/iafis/iafis`.
3. Wilson, C., *et al.*: Fingerprint Vendor Technology Evaluation 2003: Summary of Results and Analysis Report. NISTIR 7123 (2004)
4. Ashbaugh, D.R.: Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology. CRC Press (1999)
5. Science Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST): Standards for Examining Friction Ridge Impressions and Resulting Conclusions (Latent/Tenprint) (2011)

6. Office of the Inspector General, Oversight and Review Division: A Review of the FBI's Handling of the Brandon Mayfield Case (2006)
7. Office of the Inspector General, Oversight and Review Division: A Review of the FBI's Progress in Responding to the Recommendations in the Office of the Inspector General Report on the Fingerprint Misidentification in the Brandon Mayfield Case (2011)
8. National Research Council: Strengthening Forensic Science in the United States: A Path Forward (2009)
9. National Institute of Standards and Technology (NIST) and National Institute of Justice (NIJ): Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach (2012)
10. Hicklin, R.A., *et al.*: Latent Fingerprint Quality: A Survey of Examiners. Journal of Forensic Identification **61**(4) (2011) 385–419
11. Ulery, B.T., Hicklin, R.A., Buscaglia, J., Roberts, M.A.: Accuracy and Reliability of Forensic Latent Fingerprint Decisions. Proceedings of the National Academy of Sciences **108**(19) (2011) 7733–7738
12. Ulery, B.T., Hicklin, R.A., Buscaglia, J., Roberts, M.A.: Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners. PLoS ONE **7**(3) (2012)
13. Indovina, M., Dvornychenko, V., Hicklin, R.A., Kiebuzinski, G.I.: Evaluation of Latent Fingerprint Technologies: Extended Feature Sets [Evaluation #2]. NISTIR 7859 (2012)
14. NIST Special Database 27: Fingerprint Minutiae from Latent and Matching Tenprint Images `www.nist.gov/srd/nistsd27.htm`.
15. West Virginia University Latent Fingerprint Database: `http://www.csee.wvu.edu/~ross/`.
16. Indovina, M., Dvornychenko, V., Tabassi, E., Quinn, G., Grother, P., Meagher, S., Garris, M.: ELFT Phase II - An Evaluation of Automated Latent Fingerprint Identification Technologies. NISTIR 7577 (2009)
17. Jain, A.K., Feng, J.: Latent Fingerprint Matching. IEEE Trans. Pattern Analysis and Machine Intelligence **33**(1) (2011) 88–100
18. Yoshida, A., Hara, M.: Fingerprint Image Quality Metrics That Guarantees Matching Accuracy. Biometric Quality Workshop (2006)
19. NIST: Latent Print AFIS Interoperability Working Group `www.nist.gov/oles/afis_interoperability.cfm`.
20. Fronthaler, H., Kollreider, K., Bigun, J.: Local Features for Enhancement and Minutiae Extraction in Fingerprints. IEEE Trans. on Image Processing **17**(3) (2008) 354–363
21. Jain, A.K., Feng, J.: Latent Palmprint Matching. IEEE Trans. Pattern Analysis and Machine Intelligence **31**(6) (2009) 1032–1047
22. Paulino, A.A., Feng, J., Jain, A.K.: Latent Fingerprint Matching Using Descriptor-Based Hough Transform. IEEE Trans. on Information Forensics and Security (2012) (To appear).
23. Dvornychenko, V.N.: Evaluation of Fusion Methods for Latent Fingerprint Matchers. In: Proc. 5th International Conference on Biometrics (ICB). (2012) 182–188
24. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Chapman and Hall/CRC (1984)
25. NIST Special Database 14: NIST Mated Fingerprint Card Pairs 2 (MFCP2) `http://www.nist.gov/srd/nistsd14.cfm`.
26. Tabassi, E., Wilson, C., Watson, C.: Fingerprint Image Quality. NISTIR 7151 (2004)