To appear in the Proceedings of the IEEE International Conf. on Biometrics (ICB), Halmstaad, Sweden, June 2016.

Automatic Latent Value Determination

Kai Cao[†], Tarang Chugh[†], Jiayu Zhou[†], Elham Tabassi[‡] and Anil K. Jain[†]

Michigan State University[†]
Email: {kaicao, chughtar, zhou, jain}@cse.msu.edu

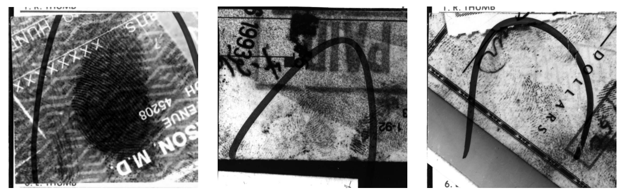
National Institute of Standards and Technology[‡]
Email: elham.tabassi@nist.gov

Abstract

Latent fingerprints are the most frequently encountered and reliable crime scene evidence used in forensics investigations. Automatic methods for quantitative assessment of a latent in terms of (i) value for individualization (VID), (ii) value for exclusion only (VEO), and (iii) no value (NV), are needed to minimize the workload of latent examiners so that they can pay more attention to challenging prints (VID and NV latents). Current value determination is either made by examiners or predicted given manually annotated features. Because both of these approaches depend on human markup, they are subjective and time consuming. We propose a fully automatic method for latent value determination based on the number, reliability, and compactness of the minutiae, ridge quality, ridge flow, and the number of core and delta points. Given the small number of latents with VEO and NV labels in two latent databases available to us (NIST SD27 and WVU), only a two-class value determination is considered, namely VID and $\overline{VID}$, where the $\overline{VID}$ class contains VEO and NV latents. Experimental results show that the value determination by the proposed method (i) obviates the need for examiner markup while maintaining the accuracy of value determination and (ii) can predict the AFIS performance better than examiners.

1. Introduction

Latent fingerprints (latents or marks) refer to friction ridge patterns lifted at crime scenes and are one of the most important evidence used in forensics investigations. Due to the generally low quality of latents, the Analysis, Comparison, Evaluation and Verification (ACE-V) methodology [5] relies on manual feature markup and verification. FBI's Integrated Automated Fingerprint Identification System (IAFIS) conducted 14,653 latent feature searches and 2,682 latent image searches during September 2015 alone [1]. This indicates a significant effort required by examiners for markup and examination of AFIS candidate lists. While human markup is necessary for all the submitted latent queries in the current framework, high quality latents can be handled in a fully automated fashion, or "lights-out" mode. To that end, an accurate and robust method to determine the value of a latent is needed.



(a) (b) (c)

Figure 1: Latents of different values. (a), (b) and (c) are determined by examiners to be VID, VEO and NV, respectively.

In the analysis of ACE-V, each latent is assigned one of the following three values by a latent examiner [3] (see Figure 1 for examples): value for individualization (VID), value for exclusion only (VEO), and no value (NV). VID and VEO latents are typically considered to be "of value" and used for further comparison, while NV latents are typically discarded to avoid unnecessary feature annotation and comparison. Therefore, an erroneous NV determination for a latent may lead to missed opportunity to identify the suspect, while an erroneous "of value" determination takes up examiner time for fruitless comparisons [22]. It should be noted that some forensic agencies only use VID latents for comparison [3]. This prevailing practice of manual value determination highly depends on the examiner skill and experience, therefore it is subjective and lacks transparency. According to the fingerprint examiners' "black box" study, which examined the consensus of latent value determination by examiners [20, 21], the repeatability (intra-examiner variability) of value determination was 84.6% after a gap of several months, while reproducibility (inter-examiner similarity) was only 75.2%; moreover, for some latents, the value determination by examiners even changed from VID to NV when examined at different times. Another serious shortcoming of value determination by examiners is that it may not be indicative of the AFIS performance. The ELFT-EFS report [14] showed that a significant portion of VEO or

NV latents, as determined by examiners, could be successfully identified by an AFIS.

In a typical forensic agency[1], the latent examiners work in a symbiotic manner with the agency's AFIS in "semi-lights-out mode" (semi-automatic search), where the examiners first do the value determination and markup and the AFIS then provides the candidate list to the examiner for verification. In this scenario, the value of a latent determined by examiners is actually coupled to a specific AFIS. This has been previously referred to as "value determination by AFIS" [24].

Given all the above mentioned shortcomings of examiners value determination, it is desirable to develop an objective measure of value determination, which can (1) alleviate subjectivity, (2) reduce workload for examiners and increase throughput, and (3) learn an AFIS-dependent value determination to improve the overall performance (i.e., hit rate and throughput) of the "semi-lights out mode" system.

To our knowledge, there is no fully automatic method for latent value determination in the published literature. A commonly practiced heuristic for value determination is the number of minutia points in a latent (examiner markup); a minimum of 7 minutiae are required for further comparison as reported in [7]. However, minutiae count alone, without other features such as ridge clarity and minutiae configurations, may not be sufficient for value assessment. The work in [15] utilized similarity score output by a tenprint SDK for value determination without any latent feature extraction. Yoon et al. [24] investigated a set of level-1 and level-2 features to define Latent Fingerprint Image Quality (LFIQ). After feature selection, LFIQ was finally defined in terms of two features: average ridge clarity and the total number of minutiae. The LFIQ was extended in [23] by incorporating additional features, namely connectivity of good ridge structures, minutiae reliability, and finger position. This augmented feature set resulted in good prediction of the identification performance of a state-of-the-art AFIS. However, LFIQ requires manually annotated minutiae, which is both arduous and subjective.

To understand how latent examiners determine the value of latents, Ulery et al. [22] modeled the relationship between value determination and feature annotation by certified latent examiners. The annotated features included level-1 features (fingerprint pattern, orientation and singular points), level-2 features (minutiae), level-3 features (dots, incipient ridges, pores, etc.), and image clarity. Their results showed that while minutiae count was strongly associated with value determination, the correlation between examiner annotation and value determination has limitations due to lack of reproducibility of both manual annotation and value determination.

[1]The workflow for latent processing may vary from one agency to the other.

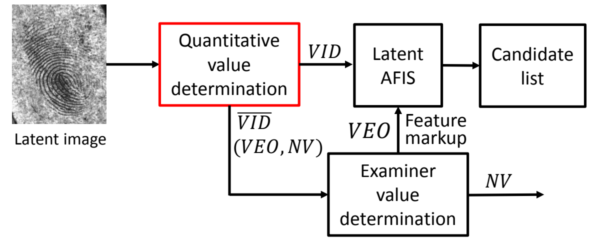


Figure 2: Proposed latent search framework. Quantitative value determination is an important component (shown in red).

Table 1 compares value determination studies reported in the literature which are all based on manually annotated features, specifically minutiae. In this paper, our goal is to develop a fully automatic latent value assessment approach, which can improve the accuracy and efficiency of the latent search process (see Figure 2). We present a fully automatic latent feature extractor where the features are the number of minutiae, ridge clarity, ridge flow, singular points, and minutia reliability. Experimental results on two latent databases, NIST SD27 and WVU, demonstrate that the proposed latent value assessment method is consistent for both examiners and AFIS.

The main contributions of this paper are:

1. A fully automatic feature extraction method that consists of minutiae-based, ridge-based, singularity-based and ridge flow-based features.
2. A proposed method for automatic latent value determination that obviates the need for examiner markup without sacrificing the value determination accuracy.

2. Proposed method for latent value assessment

Figure 3 illustrates the main steps of the proposed value assessment algorithm. We first present details of the proposed latent processing and automatic minutiae extraction, followed by extraction of additional features used for value assessment.

2.1. Latent processing and minutiae extraction

As shown in Figure 4, a latent image often includes multiple impressions. Therefore, marking the region of interest (ROI) is regularly the first task of an examiner. Our proposed algorithm assumes that the ROI is marked; the main steps post-ROI markup are shown in Figure 5 and explained below.

2.1.1 Ridge flow estimation

Ridge flow, or orientation field, is essential for ridge structure enhancement, minutiae extraction, pattern type classification, and identification. For ridge flow estimation of the input latent, we use the ConvNet-based algorithm proposed

Table 1: Value determination studies published in the literature.

Study	Features	Latent database	Comments
Yoon et al. [24]	Ridge quality and number of minutiae	NIST SD27 (258 latents) and WVU (449 latents)	Manually marked minutiae required
Yoon et al. [23]	Ridge quality, minutiae reliability and finger position	NIST SD27 and WVU	Manually marked minutiae required
Ulery et al. [22]	Features at all 3 levels and image clarity	1,850 latents from 6 sources	All features were manually annotated
Sankaran et al. [19]	Ridge clarity and ridge quality	NIST SD27	A tenprint AFIS rather than a latent AFIS was used
Olsen et al. [17]	A subset of features specified in NFIQ 2.0	No latent prints were used, only 749 live-scan prints were used	Tenprints rather than latents were used; manually marked minutiae required
Proposed approach	Number of minutiae, ridge clarity, core and delta, and ridge flow features	NIST SD27 and WVU	Automatic feature extractor developed by the authors; state-of-the-art latent AFIS[2] used

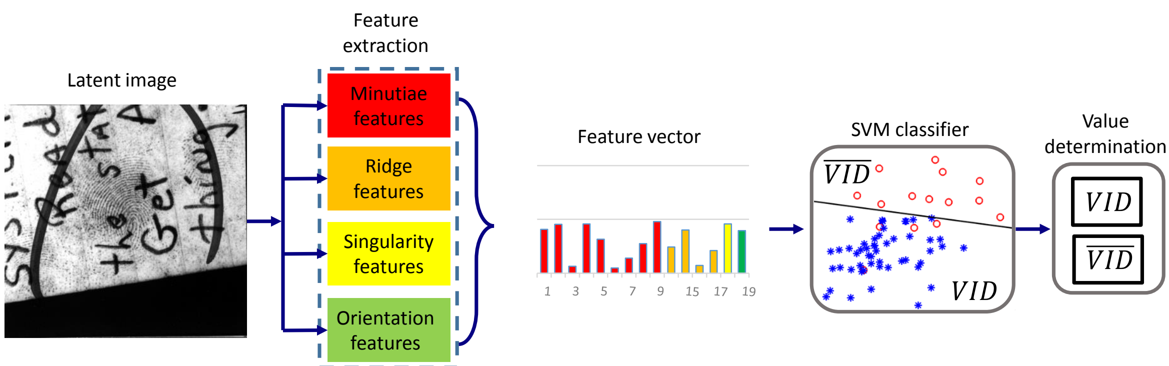


Figure 3: Illustration of the proposed automatic latent value determination algorithm. Target values for the classifier are VID and $\overline{\text{VID}}$ (VEO or NV).

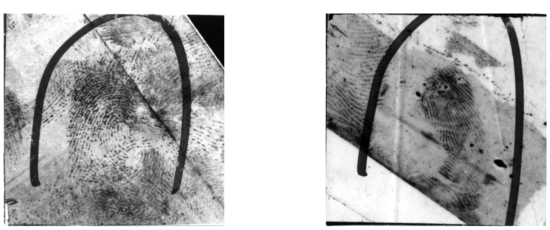


Figure 4: Latents containing multiple friction ridge patterns; region of interest (ROI) is indicated by an examiner.

in [8], which has shown the best performance in terms of the accuracy of the estimated ridge flow compared to manually annotated ridge flow. The ridge flow of the input latent image (Figure 5 (a)) is estimated within each block of 16×16 pixels (illustrated in Figure 5 (b)).

2.1.2 Latent image normalization

Given that the variation of gray values is minimum along the ridge orientation and maximum along the direction normal to the local ridge orientation, the normalized image I_N of the input latent I is obtained as follows:

1. Expand the block-wise ridge flow (section 2.1.1) to pixel-wise ridge flow (θ) by bilinear interpolation.
2. At pixel (i, j), place an oriented window of size $(n \times m)$ defined in the local ridge coordinate system, as shown in Figure 6 (n and m are empirically set to 11 and 7, respectively).
3. Compute the $X - signature$ [13], $\{X_k\}_{k=0}^{n-1}$, of the

[2]Latent SDK does not reveal minutiae or other features.

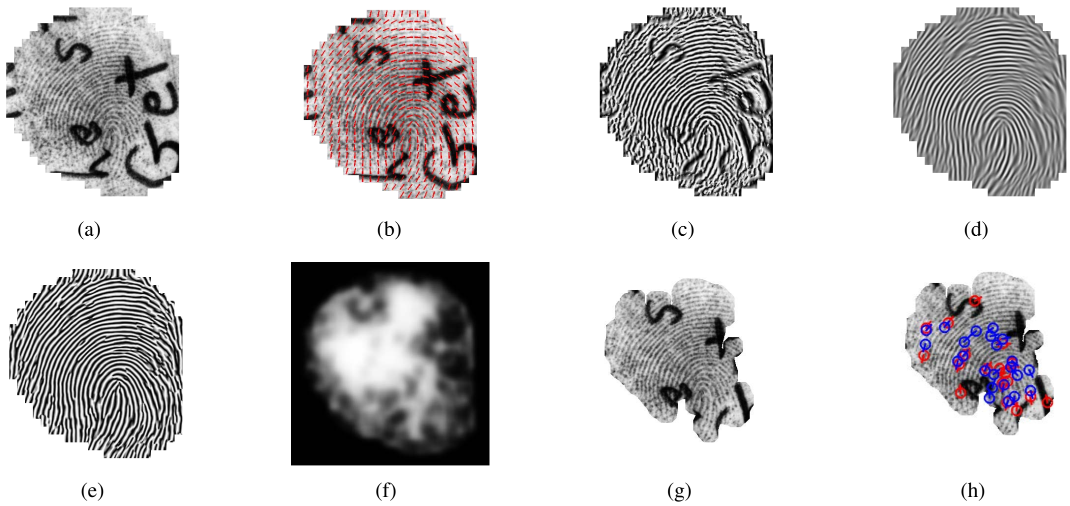


Figure 5: Illustration of the main steps in automatic minutiae extraction. (a) Input latent (I) with ROI, (b) estimated ridge flow, (c) normalization of I (I_N), (d) enhancement of I_N (E), (e) normalization of E (E_N), (f) estimated quality map, (g) cropped latent image based on ridge quality map, and (h) detected minutiae (red) and manually annotated minutiae (blue). In (g) and (h), the automatical cropping explains the irregular boundary.

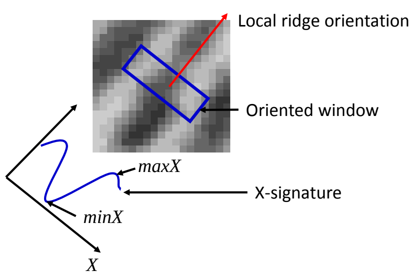


Figure 6: Oriented window and $X-signature$.

ridges and valleys within the oriented window, where

$$X_k = \frac{1}{m}\sum_{d=0}^{m-1} I(u,v), k = 0, 1, ..., n-1, \quad (1)$$
$$u = i + (d - \frac{m}{2})\cos\theta(i,j) + (k - \frac{n}{2})\sin\theta(i,j),$$
$$v = j + (d - \frac{m}{2})\sin\theta(i,j) + (\frac{n}{2} - k)\cos\theta(i,j).$$

4. Select the minimum value, $minX$, and maximum value, $maxX$, of $\{X_k\}_{k=0}^{n-1}$.
5. Convert $I(i,j)$ to the range $[0,1]$ by

$$I_N(i,j) = \begin{cases} 1, & \text{if } I(i,j) > maxX, \\ 0, & \text{if } I(i,j) < minX, \\ \frac{I(i,j)-minX}{maxX-minX}, & \text{otherwise.} \end{cases} \quad (2)$$

Comparison of Figures 5 (a) and (c) shows that the contrast of ridges and valleys in the normalized latent image is significantly improved.

2.1.3 Ridge enhancement

The goal of ridge enhancement is to improve the ridge clarity for minutiae extraction. Fingerprint enhancement methods generally assume that the ridge and valley widths are the same, which does not hold for latent prints. Figure 7 (a) shows an example latent where the ridge width is larger than the valley width which leads to some spurious valleys by Gabor filtering, as shown in Figure 7 (b). To circumvent this, we propose a dictionary based method consisting of the following two steps:

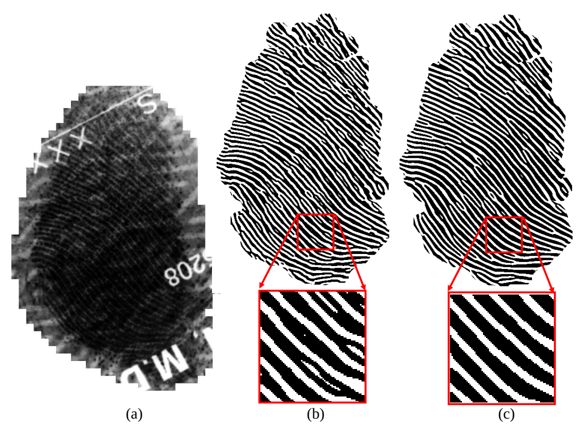
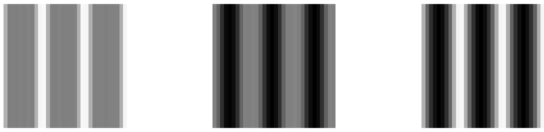


Figure 7: Latent ridge enhancement. (a) Input latent image, (b) binarization of the enhanced latent by Gabor filtering, and (c) binarization of the enhanced latent by the proposed dictionary based approach. Note that some spurious valleys are generated in (b).

1. *Dictionary Construction*: To generate ridge structure



(a) (b) (c)

Figure 8: Dictionary element construction. (a) A valley image, (b) a ridge image, and (c) an element constructed by (a) and (b).

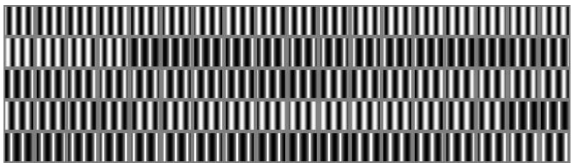


Figure 9: A subset of dictionary elements used for latent ridge enhancement. The ridge width is not always the same as valley width.

dictionary elements with unequal ridge and valley widths, we combine two sine images, a ridge image and a valley image. For a specific orientation α, ridge width w_r and valley width w_v, the ridge image D_r and valley image D_v are constructed as follows:

$$D_v(y,x) = \begin{cases} \sin\frac{\pi l}{w_v}, & \text{if } 0 \leqslant l < w_v, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

$$D_r(y,x) = \begin{cases} -\sin\frac{\pi(l-w_v)}{w_r}, & \text{if } w_v \leqslant l < w, \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $l = mod(d,w)$, $d = x\cos(\frac{\pi}{2}-\alpha) - y\sin(\frac{\pi}{2}-\alpha)+O$; O is the offset in the range $\{0,1,...,w-1\}$ and $w = w_r + w_v$. A dictionary element D is then constructed by the ridge image and valley image ($D = D_r + D_v$). Figure 8 illustrates the dictionary construction process. In order to accommodate various types of ridge structure, α is set in the range [0°,180°) with a step size of 3°, w is in [7,14] and w_v is in [3, w/2+1] with step size of 1 pixel. This results in a ridge structure dictionary with 38,400 elements of size 32×32 pixels. Note that dictionary elements with the same orientation form a subdictionary for the subsequent reconstruction. Example dictionary elements are shown in Figure 9.

2. *Enhancement via reconstruction*: The normalized latent I_N is divided into overlapping patches of size 32×32 pixels. For each patch, the dominant orientation is used to find the corresponding subdictionary. The dictionary element most similar to the patch is selected from the subdictionary and used to reconstruct the patch. The enhanced latent is finally obtained by quilting all the reconstructed patches together. Figure 5 (d) shows the enhanced fingerprint image E, and Figure 5 (e) shows the normalized image of E (E_N).

2.1.4 Ridge quality estimation

Ridge quality is crucial to ensure that spurious minutiae are not detected which, in turn, is important for latent value assessment. But, assessing latent ridge quality is challenging due to background noise and poor ridge structure. The dictionary-based quality measure [9] for latent segmentation and enhancement has two limitations: (1) mismatch between the contrast of latent image and dictionary elements and (2) dictionary elements cannot account for all possible ridge patterns. We improve the approach in [9] and combine it with two coherence measures to define an effective ridge quality.

The enhanced and normalized image (E_N) of I_N in Figure 5 (e) is used as a template, and the similarity between patches from I_N and E_N is used to define ridge quality. Since this quality measure is based on patches, it is regarded as a coarse level ridge quality R_c.

Orientation coherence [6] is a measure of the collinearity of gradients in a local region. Let $Coh(I)$ denote the orientation coherence map of fingerprint image I. Based on our observations that (i) values of both $Coh(I_N)$ and $Coh(E)$ are high for foreground ridge structures, (ii) only $Coh(E)$ is high for recoverable corrupted region, and (iii) neither $Coh(I_N)$ nor $Coh(E)$ is high for the noisy background, $Coh(I_N)$ and $Coh(E)$ are used to define fine level ridge quality. R_c, $Coh(I_N)$, and $Coh(E)$ are normalized into the range (0,1) using the sigmoid function. Without ambiguity, we use the same symbols for normalized quality measures. Fine level ridge quality R_f is then defined as $\{Coh(I_N) + Coh(E)\}$. The overall ridge quality R is defined as $[R_f + R_c - T]$, where T is a quality threshold used to separate background and foreground. Figure 5 (f) shows the ridge quality map of the latent in Figure 5 (a).

2.1.5 Minutiae extraction

Minutiae are extracted only in good quality regions ($R > 0$), through the following steps: (1) binarize good quality regions of E_N based on the algorithm in [11]; (2) remove small islands and fill small holes in the binarized image; (3) thin the binarized image to obtain the skeleton; (4) remove short spurs and connect broken ridges; and (5) detect minutiae in the skeleton using the method in [18]. Figure 5 (h) compares our automatically extracted minutiae against minutiae obtained from manual markup by an examiner.

Table 2: Feature vector used for quantitative value assessment.

Feature no.	Description
1	Number of minutiae extracted in the latent
2-8	Sum of reliability of minutiae that have reliability value $\geq t$, $t = 0, 0.1, ..., 0.6$
9	Average area of the triangles in minutiae Delaunay triangulation
10	Area of the convex hull of minutiae set
11-17	Sum of ridge quality of blocks that have quality value $\geq t$, $t = 0, 0.1, ..., 0.6$
18	Number of singular points (core and delta) [16]
19	Standard deviation of the ridge flow (orientation map) in the foreground

2.1.6 Minutiae reliability

Minutiae reliability, which is critical for latent value assessment, is defined in terms of ridge quality (R_M) at minutia M (section 2.1.4), minutiae compactness f, and finger position [23]. Since a minutia with close neighboring minutiae is typically considered to be of poor quality [10], we define compactness f of a minutia (M) as

$$f = \begin{cases} 0 & \text{if } d > d_1 \\ 1 & \text{if } d < d_2 \\ \frac{(d_1 - d)}{d_1 - d_2}, & \text{otherwise,} \end{cases} \quad (5)$$

where d is the average distance between M and its two nearest minutiae, d_1 and d_2 are empirically set as 60 and 20, respectively.

A latent consisting of the central part of a fingerprint is more likely to find a hit with its true mate because it generally contains more information. This suggests that higher weight should be assigned to the minutiae located in the central part of the latent. The reliability of minutia M is then defined as $Q_M = R_M \cdot (1 - f) \cdot w$, where w is a two-dimensional Gaussian function of (x, y), where x (y) is the difference between the x (y)-coordinate of a minutia and the reference point in the latent [23].

2.2. Feature vector for latent value assessment

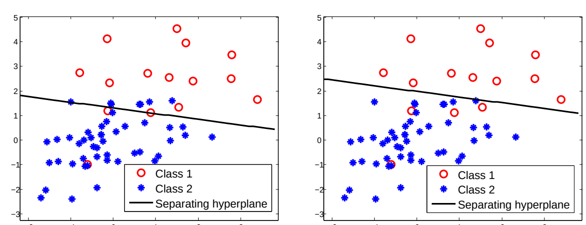
Table 2 summarizes the features used in the proposed automatic latent value assessment. The first 10 features are minutiae-related features, where feature numbers 9 and 10 characterize the spatial distribution of the minutiae set. Features 11-17 are ridge features, feature 18 is the number of singular points using complex filtering [16], and feature 19 is an orientation feature.

2.3. Learning latent value assessment

In order to learn latent value assessment, we need a training database with "true" (ground truth) value labels. True values can be provided either by latent examiners or determined by the hit rate (true mate is retrieved at rank-1 in the candidate list) of a state-of-the-art latent AFIS. Hence we design two different classifiers, one based on hit rate provided by AFIS and the other based on ground truth values provided by examiners. For value determination by examiners, all latents determined as VID by examiners constitute class y_{VID} while latents determined as VEO and NV constitute the second class, namely $y_{\overline{\text{VID}}}$. For value determination by AFIS, a latent belongs to y_{VID} if its mate is the top candidate in the candidate list output by AFIS, otherwise, it belongs to $y_{\overline{\text{VID}}}$. Two binary ($y_{\text{VID}}$ vs. $y_{\overline{\text{VID}}}$) classifiers, *i.e.*, Random Forest (RF) and Support Vector Machine (SVM), are used for value assessment. Due to the imbalanced training set (no. of samples in y_{VID} is significantly larger than the no. of samples in $y_{\overline{\text{VID}}}$), the separating hyperplane of SVM is skewed towards the majority class (Figure 10 (a)). To compensate for this, we use SVM bias refinement (referred to as "SVM+BR") to improve the overall classification accuracy by grid search (Figure 10 (b)).

Table 3: Latent database description. Values determined by examiners were reported in [12] and values determined by AFIS (in parentheses) are obtained using the hit rate of a latent AFIS[3].

	NIST SD27	WVU
No. of latents	258	449
Capture environment	operational casework	laboratory environment
No. of VID latents	210 (176)	370 (317)
No. of $\overline{\text{VID}}$ latents	48 (82)	79 (132)



(a) (b)

Figure 10: SVM bias refinement. (a) Separating hyperplane output by SVM and (b) refined separating hyperplane (SVM+BR).

3. Experimental results

3.1. Databases

Two latent fingerprint databases, NIST SD27 [2] and WVU [4], are used for training and evaluation of the pro-

[3]The latent matcher used here is one of the top performers in NIST evaluations [14].

posed latent value assessment method following 10-fold cross-validation protocol. The characteristics of these two databases are summarized in Table 3. We train two models: one for value determination by examiners and the other for value determination by AFIS. For the AFIS dependent value determination, the hit rate of a state-of-the-art latent matcher is used; the size of the reference database is extended to 100,000, including 707 mates of the latents in the two databases. If the true mate of a latent can be retrieved as the top candidate, it is referred as VID by AFIS (class y_{VID}), otherwise it is referred as $\overline{\text{VID}}$ by AFIS (class $y_{\overline{\text{VID}}}$).

3.2. Performance evaluation

A 10-fold cross validation is adopted where each of the two latent databases is partitioned into 10 subsets. Individual subsets from NIST SD27 and WVU are combined to form 10 folds. The classifier is trained using 9 of the 10 folds and evaluated on the remaining fold; this procedure is repeated 10 times, each time with a different fold for evaluation. We report the mean and variance of the performance metrics. Since learning value assessment is a binary classification problem, the efficacy of the value assessment methods is evaluated using two metrics, i.e., classification accuracy and Area Under the Curve (AUC) of ROC. To compute AUC, we use the signed distance output by SVM (positive for one side of hyperplane and negative on another side) which measures the distance between a test latent and the hyperplane in the feature space. While the two classification methods, RF and SVM+BR, have comparable classification accuracies, SVM+BR has a higher AUC value. Due to space limitation, we only report the performance of SVM+BR. We also compare the proposed method with LFIQ [23], which requires manual minutiae markup.

Table 4: Confusion matrix, classification accuracy (%) and AUC of the proposed method and LFIQ when the ground truth is provided by examiners. Numbers in the brackets indicate s.d. based on 10-fold cross validation. Each fold has ∼70 latents.

Method	Proposed*		LFIQ [23]**	
	$\hat{y}_{\text{VID}}$	$\hat{y}_{\overline{\text{VID}}}$	$\hat{y}_{\text{VID}}$	$\hat{y}_{\overline{\text{VID}}}$
y_{VID}	534	46	546	34
$y_{\overline{\text{VID}}}$	56	71	51	76
Classification accuracy	85.6% (2.4%)		88.0% (3.8%)	
AUC	0.892 (0.033)		0.903 (0.045)	

* Automatically extracted features.
** Manually annotated minutiae.

Let $\hat{y}_{\text{VID}}$ and $\hat{y}_{\overline{\text{VID}}}$ denote the predicted class labels of y_{VID} and $y_{\overline{\text{VID}}}$, respectively. Tables 4 and 5 compare confusion matrices, classification accuracies and AUC of the proposed method and LFIQ for predicting value determination by examiners and AFIS, respectively. Table 6 shows the confusion matrix of value determinations by examiners and AFIS, illustrating the difficulty of establishing the ground truth for latent value.

Table 5: Confusion matrix, classification accuracy (%) and AUC of the proposed proposed and LFIQ when the ground truth is provided by AFIS (Rank-1 retrieval).

Method	Proposed*		LFIQ [23]**	
	$\hat{y}_{\text{VID}}$	$\hat{y}_{\overline{\text{VID}}}$	$\hat{y}_{\text{VID}}$	$\hat{y}_{\overline{\text{VID}}}$
y_{VID}	447	46	433	60
$y_{\overline{\text{VID}}}$	99	115	103	111
Classification accuracy	79.5% (7.2%)		76.9% (4.1%)	
AUC	0.824 (0.069)		0.835 (0.059)	

* Automatically extracted features.
** Manually annotated minutiae.

Table 6: Confusion matrix illustrating the differences in the value determination by examiners and AFIS on latents in NIST SD 27 and WVU databases (total of 707 latents).

Method	VID by AFIS	$\overline{\text{VID}}$ by AFIS
VID by examiners	458	122
$\overline{\text{VID}}$ by examiners	35	92
Consistency (%)	77.8%	

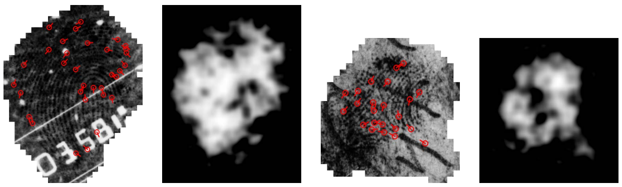
Based on these results, we make the following observations.

- Performance of the proposed automatic value assessment method is comparable to LFIQ, that requires examiner markup.
- Table 5 shows that the proposed value assessment works slightly better (79.5%) than LFIQ (76.9%) in predicting the AFIS hit rate.
- Table 6 shows that only 77.8% of the 707 latents in the two databases have the same value determinations (consistency) by examiners and AFIS. This shows the challenge is establishing the ground truth for value determination of latents.
- Ground truth values provided by examiners lead to better classification accuracy, as well as AUC, than ground truth values provided by AFIS. This was also observed in [24]. The main reason is that value determination by AFIS also depends on the quality of mated reference prints.

Figure 11 shows example latents where the proposed value assessment agrees with value determined by examiners. Figure 12 shows example latents where automatic value assessment differs from value determined by examiners.

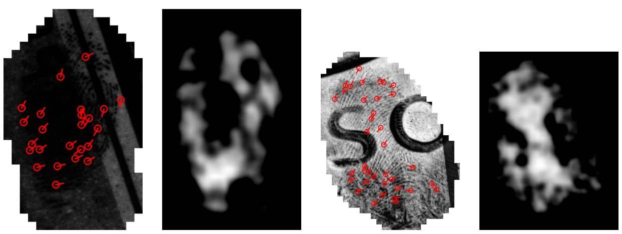
4. Conclusions and future work

An objective method for assessing the value of a latent is critical to avoid unnecessary markup and compari-



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

Figure 11: NIST SD27 latents for which the automatic value determination agrees with the examiner determined value. (a) Cropped latent (G020) and its quality map (b) with ground truth VID; (c) cropped latent (U204) and its quality map (d) with ground truth $\overline{\text{VID}}$.



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

Figure 12: NIST SD27 latents for which the automatic value determination does not agree with examiner determined value. (a) Cropped latent (B120) and its quality map (b) with ground truth VID but classified as VID; (c) cropped latent (G057) and its quality map (d) with ground truth $\overline{\text{VID}}$ but classified as VID.

son efforts by latent examiners. Current latent value determination is made either by examiners or by algorithms that require manual minutiae annotation. Both of these approaches are subjective in nature. In this paper, we present an objective latent value assessment method based on automatically extracted features from latents. Experimental results on NIST SD27 and WVU latent databases demonstrate the efficacy of the proposed method. Our quantitative value assessment can be further improved by (i) extracting multiple minutiae sets (templates) using a multi-resolution approach; (ii) using feature selection algorithms to obtain compact yet salient latent feature representation; and (iii) improving the computational efficiency (current Matlab implementation takes about 15 secs per latent, on average). We are also in the process of acquiring a larger operational latent database for robust training and evaluation of value determination.

Acknowledgment

This research is supported by the Forensic Science Program, NIST.

References

[1] Fact sheet - Integrated Automated Fingerprint Identification System (IAFIS). https://www.fbi.gov/about-us/cjis/fingerprints_biometrics/ngi/next-generation-identification-monthly-fact-sheet.

[2] NIST SD27 latent database. http://www.nist.gov/srd/nistsd27.cfm.

[3] Science Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST): Standards for Examining Friction Ridge Impressions and Resulting Conclusions (Latent/Tenprint), version 2.1 (2013).

[4] West Virginia University (WVU) Latent Fingerprint Database. http://www.cse.msu.edu/ rossarun/.

[5] D. Ashbaugh. *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*. CRC press, 1999.

[6] A. M. Bazen and S. H. Gerez. Systematic Methods for the Computation of the Directional Fields and Singular Points of Fingerprints. *IEEE TPAMI*, 24(7):905–919, 2002.

[7] B. Budowle, J. Buscaglia, and R. S. Perlman. Review of the Scientific Basis for Friction Ridge Comparisons as a Means of Identification: Committee Findings and Recommendations. *Forensic Science Communications*, 8(1), 2011.

[8] K. Cao and A. Jain. Latent Orientation Field Estimation via Convolutional Neural Network. In *ICB*, pages 349–356, May 2015.

[9] K. Cao, E. Liu, and A. Jain. Segmentation and Enhancement of Latent Fingerprints: A Coarse to Fine Ridgestructure Dictionary. *IEEE TPAMI*, 36(9):1847–1859, 2014.

[10] K. Cao, E. Liu, L. Pang, J. Liang, and J. Tian. Fingerprint matching by incorporating minutiae discriminability. In *IJCB*, pages 1–6, 2011.

[11] M. D. Garris, E. Tabassi, C. I. Wilson, R. M. McCabe, S. Janet, and C. I. Watson. NIST Fingerprint Image Software 2. 2004.

[12] R. A. Hicklin, J. Buscaglia, M. A. Roberts, S. B. Meagher, W. Fellner, M. J. Burge, M. Monaco, D. Vera, L. R. Pantzer, C. C. Yeung, and T. N. Unnikumaran. Latent Fingerprint Quality: A Survey of Examiners. *Journal of Forensic Identification*, 61(4):385–419, 2011.

[13] L. Hong, Y. Wan, and A. Jain. Fingerprint image enhancement: algorithm and performance evaluation. *IEEE TPAMI*, 20(8):777–789, Aug 1998.

[14] M. Indovina, V. Dvornychenko, R. Hicklin, and G. Kiebuzinski. ELFT-EFS Evaluation of Latent Fingerprint Technologies: Extended Feature Sets [Evaluation# 2]. *National Institute of Standards and Technology, NISTIR*, 7859, 2012.

[15] J. Kotzerke, S. Davis, R. Hayes, L. Spreeuwers, R. Veldhuis, and K. Horadam. Identification performance of evidential value estimation for fingermarks. In *BIOSIG*, pages 1–6, Sept 2015.

[16] K. Nilsson and J. Bigun. Localization of corresponding points in fingerprints by complex filtering. *Pattern Recognition Letters*, 24(13):2135 – 2144, 2003.

[17] M. Olsen, M. Bockeler, and C. Busch. Predicting dactyloscopic examiner fingerprint image quality assessments. In *BIOSIG*, pages 1–12, 2015.

[18] N. K. Ratha, S. Chen, and A. K. Jain. Adaptive flow orientation-based feature extraction in fingerprint images. *Pattern Recognition*, 28(11):1657 – 1672, 1995.

[19] A. Sankaran, M. Vatsa, and R. Singh. Automated clarity and quality assessment for latent fingerprints. In *BTAS*, pages 1–6, 2013.

[20] B. T. Ulery, R. A. Hicklin, J. Buscaglia, and M. A. Roberts. Accuracy and reliability of forensic latent fingerprint decisions. *Proceedings of the National Academy of Sciences*, 108(19):7733–7738, 2011.

[21] B. T. Ulery, R. A. Hicklin, J. Buscaglia, and M. A. Roberts. Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners. *Plos One*, 7(3):e32800, 2012.

[22] B. T. Ulery, R. A. Hicklin, G. I. Kiebuzinski, M. A. Roberts, and J. Buscaglia. Understanding the Sufficiency of Information for Latent Fingerprint Value Determinations. *Forensic Science International*, (1):99–106, 2013.

[23] S. Yoon, K. Cao, E. Liu, and A. Jain. LFIQ: Latent Fingerprint Image Quality. In *BTAS*, pages 1–8, Sept 2013.

[24] S. Yoon, E. Liu, and A. K. Jain. On latent fingerprint image quality. In *Proc. IWCF*, pages 67–82, 2012.