Segmentation and Enhancement of Latent Fingerprints: A Coarse to Fine Ridge Structure Dictionary

Kai Cao, Eryun Liu, *Member, IEEE* and Anil K. Jain, *Fellow, IEEE*

Abstract—Latent fingerprint matching has played a critical role in identifying suspects and criminals. However, compared to rolled and plain fingerprint matching, latent identification accuracy is significantly lower due to complex background noise, poor ridge quality and overlapping structured noise in latent images. Accordingly, manual markup of various features (e.g., region of interest, singular points and minutiae) is typically necessary to extract reliable features from latents. To reduce this markup cost and to improve the consistency in feature markup, fully automatic and highly accurate ("lights-out" capability) latent matching algorithms are needed. In this paper, a dictionary-based approach is proposed for automatic latent segmentation and enhancement towards the goal of achieving "lights-out" latent identification systems. Given a latent fingerprint image, a total variation (TV) decomposition model with L_1 fidelity regularization is used to remove piecewise-smooth background noise. The texture component image obtained from the decomposition of latent image is divided into overlapping patches. Ridge structure dictionary, which is learnt from a set of high quality ridge patches, is then used to restore ridge structure in these latent patches. The ridge quality of a patch, which is used for latent segmentation, is defined as the structural similarity between the patch and its reconstruction. Orientation and frequency fields, which are used for latent enhancement, are then extracted from the reconstructed patch. To balance robustness and accuracy, a coarse to fine strategy is proposed. Experimental results on two latent fingerprint databases (i.e., NIST SD27 and WVU DB) show that the proposed algorithm outperforms the state-of-the-art segmentation and enhancement algorithms and boosts the performance of a state-of-the-art commercial latent matcher.

Index Terms—Latent fingerprint, image decomposition, segmentation, ridge enhancement, sparse coding, dictionary learning.

1 INTRODUCTION

LATENT fingerprints (or simply latents or finger marks) refer to fingerprints lifted from the surfaces of objects inadvertently touched or handled by a person typically at crime scenes. Compared to rolled and plain fingerprints (or exemplar fingerprints), which are acquired in an attended mode, latents are typically of poor quality in terms of ridge structure, containing background noise and non-linear distortion (see Fig. 1). Due to these factors, the latent identification (i.e., latent to exemplar matching) accuracy is much lower than that of exemplar fingerprints (exemplar to exemplar matching). As an example, in NIST evaluations, while the best performing Automated Fingerprint Identification System (AFIS) achieved a rank-1 identification rate of 99.4% on a background database of 10,000 plain fingerprints [1], the best performing commercial latent matcher could only achieve a rank-1 identification rate of 63.4% in searching 1,114 latents against a background database containing 100,000 exemplar prints [2]. Fig. 2 shows examples of rolled to rolled fingerprint matching and latent to rolled fingerprint matching. In these examples, features in the rolled fingerprints are extracted automatically by an AFIS, but features in the latent image are manually marked.

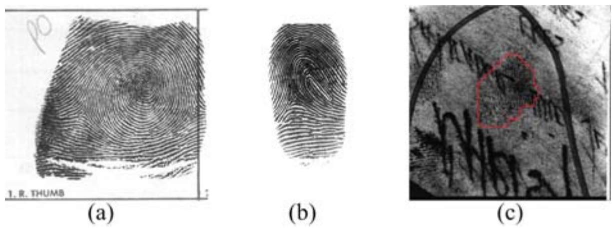


Fig. 1: Three types of fingerprint images. (a) rolled fingerprint, (b) plain fingerprint, (c) latent fingerprint with foreground (friction ridge pattern) highlighted by red outline. Notice the presence of different types of noise and distortion in (c).

One of the challenging problems in latent identification is how to automatically extract reliable features in latents, especially latents with poor quality. Given the difficulty of automatic feature extraction, a manual

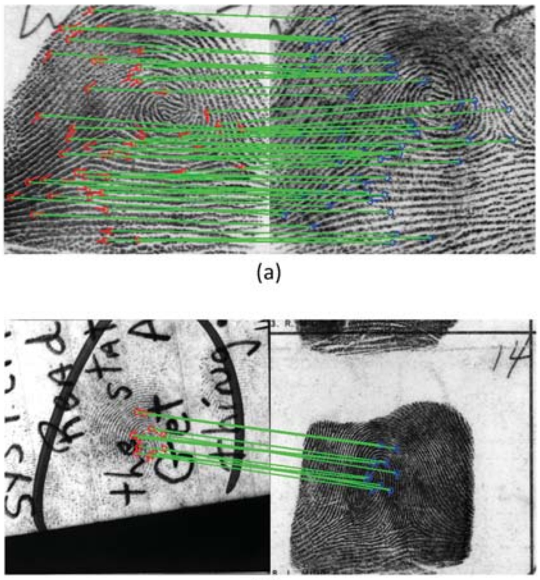
- *Kai Cao and Eryun Liu are with the Dept. of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 U.S.A. They are also affiliated with the School of Life Sciences and Technology, Xidian University, Xi'an, Shaanxi 710126, China.*
  E-mail: {kaicao,liueryun}@cse.msu.edu
- *Anil K. Jain is with the Dept. of Computer Science and Engineering, Michigan State University, 3115 Engineering Building, East Lansing, MI, U.S.A. He is also affiliated with the Department of Brain and Cognitive Engineering, Korea University, Seoul 136-713, Republic of Korea.*
  E-mail: Jain@cse.msu.edu

markup of various features in latents, such as region of interest (ROI), singular points and minutiae, is the current practice. However, this human factor issue in latent examination has raised some concerns related to repeatability and reliability [3],[4]. For example, markup of a specific feature type (e.g., minutiae) by different latent examiners or even by the same examiner at different times may not give the same results. A study conducted by NIST showed that the accuracy of a latent matcher is highly affected by the precision of latent examiner markup, especially when the latent image itself is not available to the matche [2]. One of the factors that affects human markup performance is that latent examiners often work under extreme time pressure due to heavy case work. Studies have shown that when the comparison time is limited, latent examiners are more likely to make an inconclusive matching decision between a latent and its mated rolled print [5]. One of the most infamous cases involving mistaken identity based on latent matching is the Brandon Mayfield case [6]. Other cases of mistaken identifications have been reported by the Innocence project [7]. One of the priorities of FBI's Next Generation Identification (NGI) is to support the development of a lights-out[1] capability for latent identification [9]. An essential component of this lights-out capability is to develop a fully automatic latent feature extraction module. This is highly desirable to (i) increase the throughput of latent matching systems, (ii) improve repeatability of latent feature extraction and, (iii) increase the compatibility between features extracted in the latents and features extracted in the reference prints by an AFIS [10].

An AFIS, whether for rolled/plain print matching or latent matching, typically contains a number of modules, including region of interest (ROI) segmentation (separating friction ridge from background), enhancement, feature extraction, and matching. Segmentation, especially for latents, is critical to avoid extraction of features (e.g., minutiae) in the noisy background [11]. Enhancing ridge and valley structures and removing noise in the foreground region are essential to extract accurate features. The Gabor filter based fingerprint enhancement[12] can adaptively improve the clarity of ridge and valley structures; the filters are tuned based on the local ridge orientation and frequency. Therefore, for latent enhancement it is essential to get good estimates of ridge orientation and frequency fields.

There is a rich body of literature on exemplar fingerprint segmentation [12], [13], [14], [15], orientation field estimation [12], [16], [15], [17], [18] and frequency field estimation [19], [15]. But these approaches do not work well on latent fingerprints. Compared with exemplar fingerprints, there are two types of difficulties in latent segmentation and enhancement: i) presence of structured noise, such as lines, markings, characters and speckles. Some of the structured noise may even overlap with the ridge pattern (see Fig. 1 (c)). This structured noise breaks the ridge flow pattern in the fingerprint and affects the subsequent processing; ii) poor quality of ridge structure in the foreground area. The low clarity of ridge structure makes it difficult to estimate orientation and frequency fields in latents.

1. Lights-out identification refers to an AFIS requiring minimal or no human assistance in which a query fingerprint image is presented as input, and the output consists of a short candidate list [8].



(a)

(b)

Fig. 2: Examples of (a) rolled to rolled fingerprint matching and (b) latent to rolled fingerprint matching. In these examples, features in the rolled fingerprints were extracted automatically by an AFIS, but features (minutiae, region of interest and singular points) in the latent were manually marked.

Some approaches have been proposed that specifically address the problem of latent fingerprint segmentation [20], [21], [26], [22], [25] and enhancement [23], [24], [10]. In [20], the orientation and frequency in an image block (typically 8×8 pixels) were estimated using a local gray intensity projection method; the distance between center-of-transient points was used for segmentation. However, no performance evaluation was reported. Short *et al*. [21] generated an ideal ridge model template and used the cross-correlation between a local block and the generated template to define the local fingerprint quality. However, this approach is not able to account for minutiae, ridge curvature and variance in ridge frequency. Zhang *et al*. [26] proposed an adaptive total variation (TV) decomposition model for latent fingerprint segmentation. They further proposed an adaptive directional total variation (ADTV) model by incorporating orientation field and local orientation coherence [25]. However, the orientation field and its local coherency were

TABLE 1: A comparison of latent segmentation and enhancement algorithms proposed in the literature

Algorithm	Objective	Approach	Performance evaluation	Database[a]	Comments
Karimi *et al*. [20]	Segmentation & enhancement	Local gray intensity projection	Not reported	Latents: Two latents from NIST SD27	Not robust to noise
Short *et al*. [21]	Segmentation	Correlation with ideal ridge templates	EER of 33.8%	Latents: NIST SD27	Relies on template generation
Choi *et al*. [22]	Segmentation	Orientation tensor and ridge frequency	Rank-1 identification accuracy of 16.28% and 35.19% on NIST SD27 and WVU databases, respectively, by a COTS tenprint matcher	Latents: NIST SD27 & WVU DB Background: 31,997 rolled prints	Orientation and frequency maps are used separately
Yoon *et al*. [23]	Enhancement	Polynomial model and zero-pole model	Rank-1 identification accuracy of 35% by VeriFinger[b] SDK 4.2	Latents: NIST SD27 Background: 258 rolled prints	Requires ROI, singular points and uses fixed ridge frequency
Yoon *et al*. [24]	Enhancement	STFT and R-RANSAC	Rank-1 identification accuracy of 26% by VeriFinger SDK 4.2	Latents: NIST SD27 Background: 27,258 rolled prints	Requires ROI, singular points and uses fixed ridge frequency
Feng *et al*. [10]	Enhancement	Dictionary of reference orientation patches	Rank-1 identification accuracy of 26% by VeriFinger SDK 6.2	Latents: NIST SD27 Background: 27,258 rolled prints	Requires ROI and uses fixed ridge frequency
Zhang *et al*. [25]	Segmentation & enhancement	Adaptive directional total variation model	Rank-1 identification accuracy of less than 2% by VeriFinger SDK 6.6	Latents: NIST SD27 (1,000ppi) Background: 27,258 rolled prints	Relies on orientation field and orientation coherence estimation
Proposed algorithm	Segmentation & enhancement	Image decomposition; ridge structure coarse to fine dictionary	Rank-1 identification accuracy of 61.24% and 70.16% on NIST SD27 and WVU databases, respectively, by a COTS latent matcher	Latents: NIST SD27 & WVU DB Background: 31,997 rolled prints	Fully automatic; no manual markup required

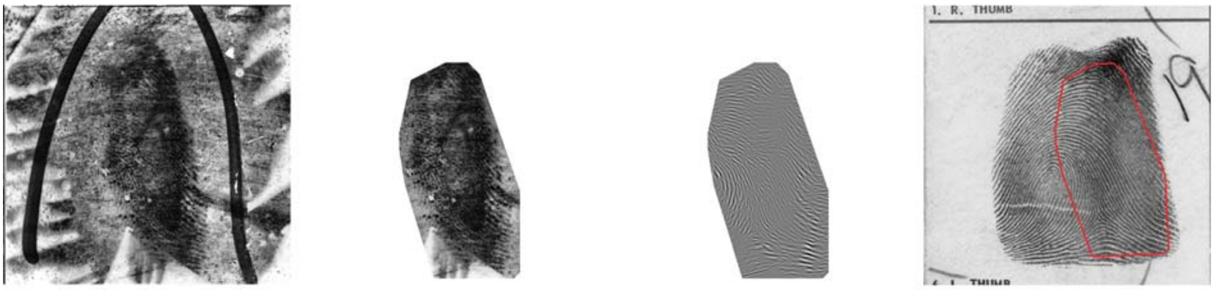
[a] The NIST SD27 contains 258 latent fingerprints and their rolled mates and the WVU DB contains 449 latent fingerprints and their rolled mates (4,290 additional rolled prints are also provided). The background database includes 258 rolled fingerprint from NIST SD27, 4,739 rolled fingerprints from WVU DB and 27,000 rolled fingerprints (including the 449 mated prints) from NIST SD14.
[b] Verifinger is a tenprint matcher.

computed from the gray scale latent images which are not reliable due to the low quality of friction ridge structure. Choi *et al*. [22] used the orientation tensor and Fourier response in valid frequency regions to segment latent fingerprints separately; an intersection of the resulting segmented regions was used to localize the latent fingerprint (foreground) region. However, they did not utilize the correlation between ridge orientation and ridge frequency.

In the latent enhancement approaches reported in the literature, manual markup of ROI and/or singular points are typically needed as input [23], [24], [10]. In [23], the orientation field was obtained by fitting a polynomial model to the coarse orientation field. The authors improved this approach by using short-time Fourier transform (STFT) and randomized RANSAC [24], but it still required manual markup of ROI and singular points. Feng *et al*. proposed a dictionary of orientation patches to estimate the orientation field in the manually marked ROI [10]. The orientation field estimation was posed as an energy minimization problem, which consists of the (i) similarity between orientation patches and orientation dictionary elements, and (ii) compatibility between neighboring orientation dictionary elements. However, all these approaches used a fixed ridge frequency to tune Gabor filters for latent enhancement. In [25], image decomposition using the ADTV model was used to enhance the ridge quality. Table 1 compares various latent segmentation and enhancement algorithms proposed in the literature.

To address the two main difficulties in latent fingerprint matching, namely presence of structured noise and poor quality of ridge structure, we propose a dictionary-based segmentation and enhancement of latent fingerprint. To remove the structured noise, the total-variation model with L_1 fidelity regularization [27] is used to decompose a latent into a texture part and a cartoon part[2]. The relative reduction rate of local total variation (LTV) is used as an index of local oscillatory pattern and a nonlinear decomposition method was proposed for the decomposition. The cartoon part with piecewise-smooth characteristics containing structured noise is discarded. To define ridge quality and estimate orientation and frequency fields, a ridge structure dictionary is proposed. A dictionary is a set of words (or vectors) used to sparsely and linearly represent signals of the same dimension, namely sparse coding. Dictionary and sparse coding have been successfully applied to a number of signal processing tasks, such as image denoising [28], [29], clas-

2. Cartoon image is a piecewise-smooth image whose discontinuity set includes in a union of curves whose overall length is finite [27].



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

Fig. 3: An example of latent segmentation and enhancement by the proposed algorithm. (a) A latent fingerprint image (U286 from NIST SD27); (b) fully automatic segmentation of (a) by the proposed algorithm; (c) enhancement of (b) by the proposed algorithm; (d) the true mate (rolled print) of (a) with the segmentation boundary in (b) outlined on the mate. By feeding the original latent in (a), the segmented latent in (b) and the enhanced latent in (c) into a commercial off-the-shelf (COTS) latent matcher (with a background database of 31,997 reference prints), the mated print is retrieved at ranks 4,152, 26 and 2, respectively. The contrast of images in (a) and (b) has been enhanced for better visual quality.

sification [30], [31] and face recognition [32], [33]. Feng *et al*. [10] proposed a dictionary of orientation patches for orientation field estimation in latent fingerprints. However, their dictionary cannot be successfully used for latent segmentation and frequency field extraction since ridge information is ignored. In this paper, the ridge structure dictionaries are learnt from a set of high quality fingerprint patches from rolled fingerprints and then used to recover the ridge structure from noisy latent patches. The ridge quality of a patch is defined as the structural similarity between the original patch and its reconstruction. Orientation and frequency fields of a patch are estimated from its reconstruction and used to tune the Gabor filters for latent enhancement. Fig. 3 shows an example latent fingerprint where the proposed segmentation and enhancement improves the latent identification performance (retrieval rank) of a COTS latent matcher. The main contributions of this paper are as follows:

1) Ridge structure dictionary is proposed for latent segmentation and enhancement. The dictionary is learnt from hight quality fingerprint patches.
2) The ridge quality of a patch is defined as the structural similarity between the patch and its reconstruction from the learnt ridge structure dictionary. Orientation and frequency fields of a patch are estimated from its reconstruction.
3) To balance robustness and accuracy, a coarse to fine strategy is proposed which uses dictionaries at two levels of resolution (64×64 and 32×32).
4) The proposed segmentation and enhancement algorithms outperform published algorithms and can also significantly boost the performance of a state-of-the-art commercial latent matcher on two latent database (NIST SD27 and WVU DB).

The rest of the paper is organized as follows. The details of the proposed algorithm are presented in section 2. Experimental results are reported in section 3. Finally, conclusions and future research directions are reported in section 4.

2 Proposed Algorithm

2.1 Algorithm Overview

The proposed algorithm consists of an off-line dictionary learning stage and on-line stage for segmentation and enhancement (see Fig. 4). In the off-line stage, two types of dictionaries are learnt: i) a coarse-level dictionary with patch size of 64×64 pixels, and ii) 16 fine-level dictionaries with patch size of 32×32 pixels; these 16 fine level dictionaries correspond to 16 different ridge orientations. The 64×64 patch size dictionary is used for coarse estimation of ridge quality map, orientation and frequency fields, while the 16 orientation specific 32×32 patch size dictionaries are used for fine estimation of ridge quality map, and orientation and frequency fields computation. The patch size in fine-level dictionaries is chosen to be 32×32 pixels since it covers about two ridges and valleys in 500 ppi fingerprints and is robust to structured noise. The patch size in the coarse-level dictionary (64×64 pixels) is twice the size of the fine-level dictionary.

Given the dictionaries, on-line latent segmentation and enhancement consists of the following steps:

1) Latent decomposition: Input latent is decomposed into cartoon and texture images via local total variations [27]; the cartoon image is discarded.
2) Coarse level estimation of quality map and orientation and frequency fields: The texture image is divided into overlapping patches of size 64×64 pixels (P_L^c). Each patch has 64×48 or 48×64 overlapping pixels with each of its four connected neighboring blocks. For each patch

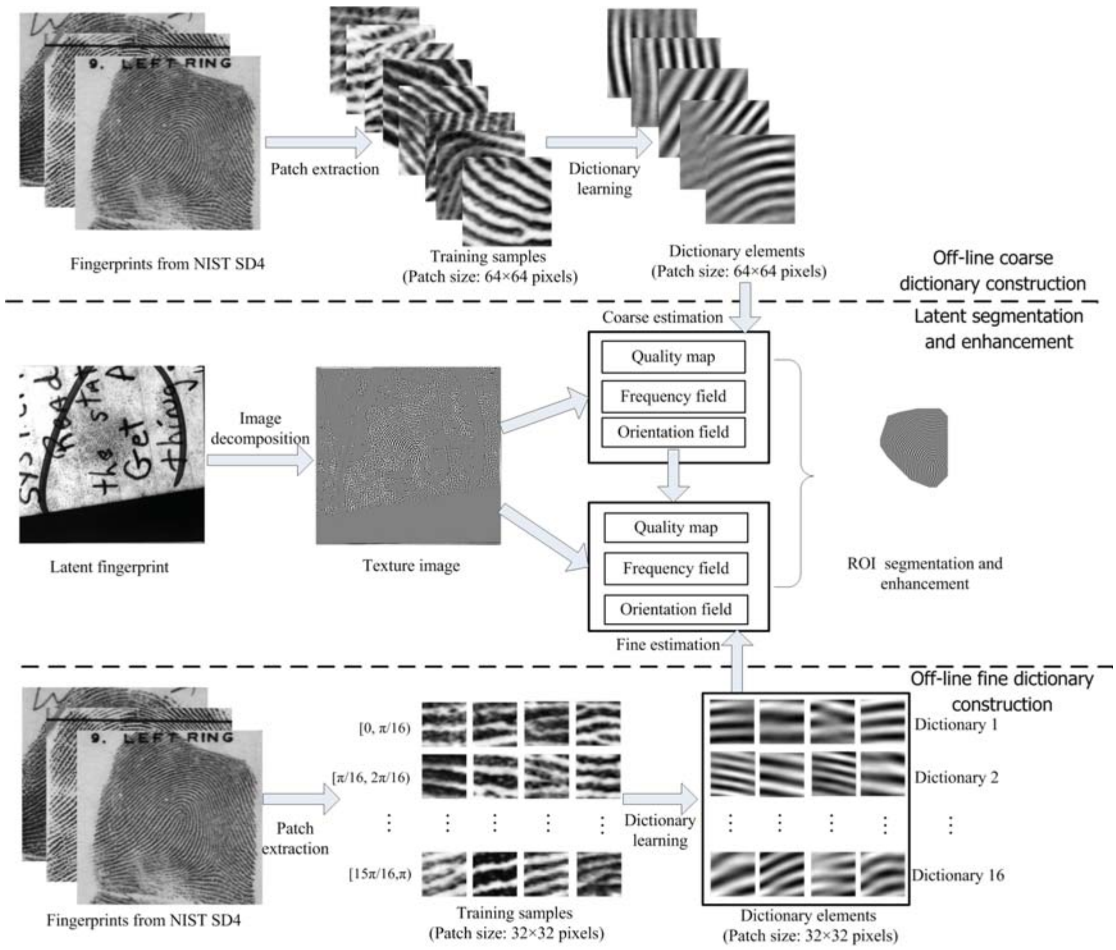


Fig. 4: Overview of the proposed latent segmentation and enhancement algorithm. The off-line dictionary learning and on-line latent segmentation and enhancement stages are separated by dashed lines.

$p \in P_L^c$, its sparse representation and reconstructed patch $\hat{p}$ using the coarse level dictionary are obtained by orthogonal matching pursuit [34]. The structural similarity [35] between p and $\hat{p}$ is defined as the coarse quality of the patch. Since the reconstructed patch $\hat{p}$ generally has a clear ridge pattern, the ridge orientation and frequency in patch p can be replaced with the features in the reconstructed patch $\hat{p}$. In regions where patches overlap, the coarse quality, coarse orientation and frequency fields are obtained by the covering patches.

3) Fine level estimation of quality map and orientation and frequency fields: The texture image is divided into overlapping patches of size 32×32 pixels (P_L^f). Each patch has 32×16 or 16×32 overlapping pixels with each of its four connected neighboring blocks. For each patch $p \in P_L^f$, the coarse ridge orientation value is first used to index the corresponding fine level dictionary. The fine estimation of ridge quality map, orientation and frequency fields are computed as coarse estimation in step 2).

4) Segmentation and enhancement: The coarse quality map and fine quality map are combined for latent segmentation. In the foreground of texture image, Gabor filtering based on the orientation and frequency fields obtained in 2) and 3) is applied for latent enhancement.

2.2 Dictionary Learning

Feng *et al*. [10] were the first to propose the use of dictionary for fingerprint orientation field estimation. However, their dictionary was based on orientation patches and they ignored the ridge structure information. This is the main reason their method in [10] was not very successful for segmentation and frequency field estimation. In this section, we present the proposed ridge structure dictionary learning.

2.2.1 Training Set Selection

In order to construct reliable and robust dictionaries, a large number of high quality fingerprint patches from rolled image in NIST SD4 [36] are selected. The image patches are selected as follows:

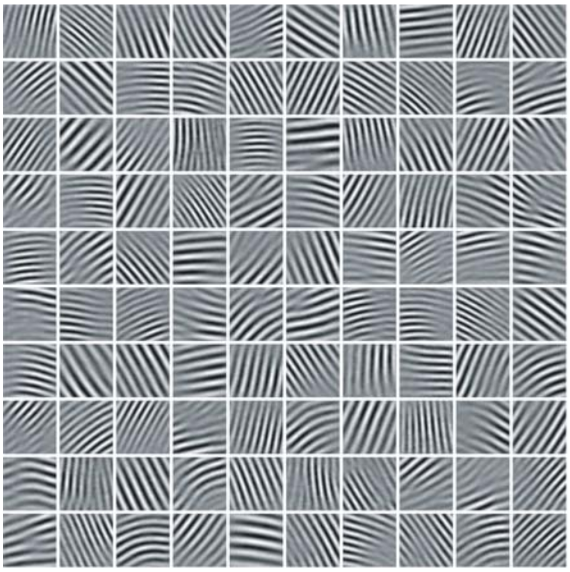


Fig. 5: A subset of elements in the coarse-level dictionary (patch size: 64×64 pixels). The total number of dictionary elements is 1,024.

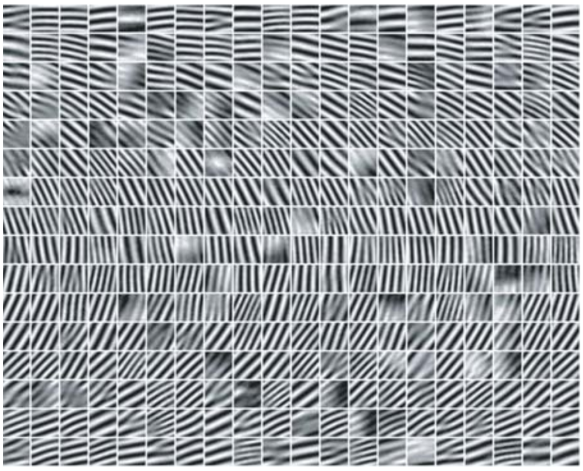


Fig. 6: A subset of elements in the 16 orientation specific dictionaries (patch size: 32×32 pixels). The ith row corresponds to the ith orientation specific dictionary in the orientation range $\left[(i-1) \times \frac{\pi}{16}, i \times \frac{\pi}{16}\right)$, i=1,...,16. The total number of elements in each orientation specific dictionary is 64.

1) High quality fingerprint selection: NIST Fingerprint Image Quality (NFIQ) [37] is used to select 500 fingerprints of high quality[3] (i.e., NFIQ < 3) in NIST SD4.
2) High quality patch selection: MINDTCT [38] (in NIST Fingerprint Image Software (NFIS)) is used to estimate the orientation field and a block-wise ridge quality map of the fingerprints selected in 1). The ridge quality map provides one of the 5 quality levels for each block (with 4 being the highest quality and 0 being the lowest quality). For the coarse-level dictionary, a set of image patches, P^c, of size 64×64 pixels are selected by sliding a window over the fingerprint image with a step size of 8 pixels; if the average quality value of an image patch is larger than a predefined threshold T (T is set to 3.75), the patch is included in the training set. For the orientation specific dictionaries, 16 sets of fingerprint patches $P_i^f, i = 1, \cdots, 16$, are constructed according to ridge orientation. For P_i^f, a 32×32 window is slided over the fingerprint image; a patch is selected if it satisfies the following two conditions: i) average quality value of the patch is larger than T, and ii) average ridge orientation of the patch is within the range $\left[(i-1) \times \frac{\pi}{16}, i \times \frac{\pi}{16}\right)$.
3) Vector normalization: Each patch p in the training set is normalized by Eq. (1) and converted to a vector by concatenating the rows.

$$\tilde{p} = (p - \mu_p)/\sigma_p, \tag{1}$$

where μ_p is the mean intensity and σ_p is the standard deviation of patch p.

Let $P^c = \{p_j^c\}_{j=1}^{N^c}$ be the training set for the coarse-level dictionary, where N^c denotes the number of training patches in P^c, and $P_i^f = \{p_{i,j}^f\}_{j=1}^{N_i^f}$, $i = 1, \cdots, 16$, be the training sets for the 16 fine-level dictionaries, where N_i^f denotes the number of training patches for the ith dictionary specified by ridge orientation. To balance efficiency and accuracy, we randomly select 80,000 patches from P^c and 10,000 patches from each P_i^f, respectively, for dictionary learning.

2.2.2 *Dictionary Learning*

Without loss of generality, for a training set $P = \{p_j\}_{j=1}^N$, the goal of dictionary learning is to construct a dictionary D of size $N_P \times N_D$ that provides the best sparse representation for each patch in P, where N_P is the dimensionality of the patches in P, and N_D is the number of elements in the dictionary D. A typical objective function for dictionary learning is

$$\min_{D,\Gamma} ||P - D\Gamma||_F^2 \quad s.t. \ \forall j, \quad ||\gamma_j||_0 \leq T_0, \tag{2}$$

where γ_j is the jth column of matrix Γ of size $N_D \times N$, $||\cdot||_0$ is the l^0 norm that counts the number of nonzero entries in the representation, T_0 is a predetermined number of nonzero entries and $||\cdot||_F$ denotes the Frobenius norm. One of the effective algorithms for dictionary learning is K-SVD [39], which minimizes the objective function in Eq. (2) by iterating the following two stages.

- Sparse coding stage: Obtain the representation coefficient vector γ_j for each patch p_j in P by

3. NFIQ ranges from 1 to 5, with 1 indicating the highest quality and 5 indicating the lowest quality fingerprint.

solving the following optimization problem under a fixed dictionary D:

$$\min_{\gamma_j} ||p_j - D\gamma_j||_2^2 \ s.t.\ ||\gamma_j||_0 \le T_0, j = 1, 2, ..., N, \tag{3}$$

- Dictionary update stage: This stage reduces the objective function in Eq. (2) by updating one column of the dictionary D at a time via singular value decomposition (SVD) while fixing all other columns of D.

The initial dictionary D is constructed using the discrete cosine transform (DCT) basis. Each dictionary element is normalized by Eq. (1) after learning.

A total of 17 different dictionaries (the coarse-level dictionary D^c and 16 fine-level dictionaries $D_i^f, i = 1, \cdots, 16$) are constructed by taking P^c and $P_i^f, i = 1, \cdots, 16$ as the training sets. The number of elements N_D^c in the coarse-level dictionary is set to 1,024, and the total number of elements N_D^f in each fine-level dictionary is set to 64. Fig. 5 shows a subset of dictionary elements in the coarse-level dictionary D^c, and Fig. 6 shows a subset of dictionary elements in the 16 fine-level dictionaries.

2.3 Latent Image Decomposition

A latent image, f, is decomposed as a sum of two components: $f = u + v$, where u represents the cartoon (piecewise smooth) part of f and v represents the oscillatory or texture part of f. Based on the characteristic of a cartoon image that its total variation does not decrease by low-pass filtering, Buades *et al.* [27] proposed a fast nonlinear decomposition method based on local total variation (LTV). The local total variation at a pixel $\mathbf{x}$ is defined as

$$LTV_\sigma(f)(\mathbf{x}) = L_\sigma * |\triangledown f|(\mathbf{x}), \tag{4}$$

where L_σ is a σ-sized low-pass filter whose Fourier transform is given by $\hat{L}_\sigma(\xi) = 1/(1 + (2\pi\sigma\xi)^4)$. The relative reduction rate $\lambda_\sigma(\mathbf{x})$ of LTV before and after filtering the image with the low-pass filter measures the local oscillatory behavior which is given by

$$\lambda_\sigma(\mathbf{x}) = \frac{LTV_\sigma(f)(\mathbf{x}) - LTV_\sigma(L_\sigma * f)(\mathbf{x})}{LTV_\sigma(f)(\mathbf{x})}. \tag{5}$$

If $\lambda_\sigma(\mathbf{x})$ is close to 1, it means the LTV decreases a lot after low-pass filtering and pixel $\mathbf{x}$ belongs to the texture part. On the other hand, if $\lambda_\sigma(\mathbf{x})$ is close to 0, there is little relative reduction of the LTV after the low-pass filtering and pixel $\mathbf{x}$ belongs to the cartoon part which is piecewise-smooth. Thus, the cartoon part u and the texture part v can be extracted by a weighted sum of f and $L_\sigma * f$ according to $\lambda_\sigma(\mathbf{x})$:

$$u(\mathbf{x}) = w(\lambda_\sigma(\mathbf{x}))((L_\sigma * f)(\mathbf{x}) - f(\mathbf{x})) + f(\mathbf{x}), \tag{6}$$

$$v(\mathbf{x}) = f(\mathbf{x}) - u(\mathbf{x}), \tag{7}$$

where $w(y)$ is a piecewise linear increasing function, defined as

$$w(y) = \begin{cases} 0 & y < a_1, \\ (y - a_1)/(a_2 - a_1) & a_1 \le y \le a_2, \\ 1 & y > a_2. \end{cases} \tag{8}$$

If $w(\lambda_\sigma(\mathbf{x})) = 0$ (pixel $\mathbf{x}$ belongs to the cartoon part), $u(\mathbf{x})$ is the same as $f(\mathbf{x})$ and $v(\mathbf{x})$ is 0. If $w(\lambda_\sigma(\mathbf{x})) = 1$ (pixel $\mathbf{x}$ belongs to the texture part), $u(\mathbf{x})$ is the same as the pixel value in the low-pass filtered image $L_\sigma * f(\mathbf{x})$ and $v(\mathbf{x})$ is the difference between the original image $f(\mathbf{x})$ and the low-pass filtered image $L_\sigma * f(\mathbf{x})$.

Fig. 9(b) shows the texture component of three different latent images shown in Fig. 9(a); most of the structured noise in latents have been successfully removed while retaining the friction ridge pattern.

2.4 Coarse Estimates of Ridge Quality, Orientation and Frequency

2.4.1 Sparse Coding

Each patch $p \in P_L^c$ is normalized by Eq. (1) and then converted to a vector by row concatenation. The sparse representation of p can be obtained by solving the following optimization problem using orthogonal matching pursuit [34]

$$\min_\alpha ||p - D^c\alpha||_2^2 \quad s.t.\ ||\alpha||_0 \le T_1, \tag{9}$$

where α is a sparse coefficient vector in which at most T_1 entries are non-zeros and D^c is the coarse-level dictionary. According to the non-zero entries in α, a subset of elements D_s^c in the dictionary D^c is selected. Let α' be the corresponding non-zero coefficients of D_s^c. Then p is projected onto the span of the elements of D_s^c. The approximation can be calculated as

$$\hat{p} = D_s^c\alpha', \tag{10}$$

where α' is given by

$$\alpha' = (D_s^{cT} D_s^c)^{-1} D_s^{cT} p. \tag{11}$$

The residual vector can be calculated as

$$r(p) = p - \hat{p} = p - D_s^c\alpha'. \tag{12}$$

2.4.2 Definition of Patch Quality

In general, the reconstruction error $||r(p)||_2$ is small if p is a fingerprint patch. However, this error measure, calculated in a pixel-wise manner, ignores the spatial property of fingerprint images [40], [41]. We have used the structural similarity index (SSIM) [35] between two images I_1 and I_2 defined as

$$\mathrm{SSIM}(I_1, I_2) = \frac{(2\mu_1\mu_2 + C_l)(2\sigma_{12} + C_c)}{(\mu_1^2 + \mu_2^2 + C_l)(\sigma_1^2 + \sigma_2^2 + C_c)}, \tag{13}$$

where μ_1 and σ_1 are the mean intensity and standard deviation of image I_1, μ_2 and σ_2 are the mean intensity and standard deviation of image I_2, σ_{12} is

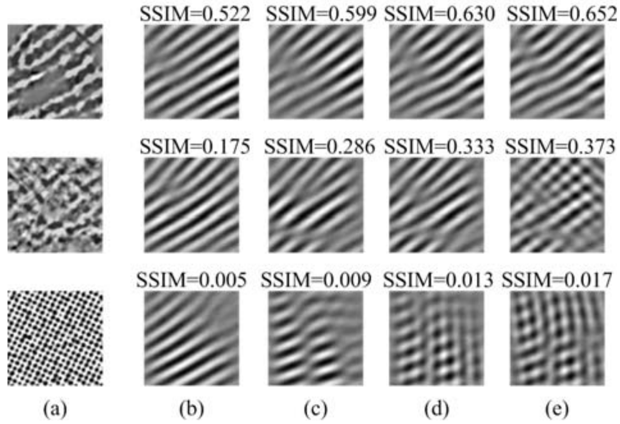


Fig. 7: Patch reconstruction results with different values of parameter T_1. (a) Texture component of high quality fingerprint patch (top), low quality fingerprint patch (middle), and non-fingerprint patch (bottom), (b), (c), (d) and (e) are the reconstruction results when $T_1 = 1$, $T_1 = 2$, $T_1 = 3$ and $T_1 = 4$, respectively. Recall that T_1 determines the number of dictionary elements used to reconstruct the original image. The SSIM indices between the given patch (column (a)) and the reconstructed patch with different value of T_1 are shown.

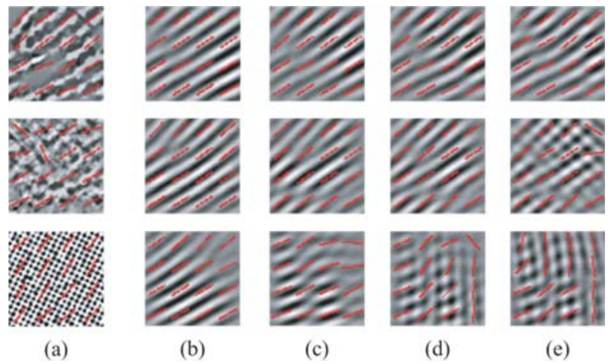


Fig. 8: Orientation fields of the patches corresponding to Fig. 7.

the covariance between I_1 and I_2, and C_l and C_c are parameters to reduce the instability in computation.

Fig. 7 compares the reconstruction results and SSIM indices for two fingerprint patches (top and middle rows) and one non-fingerprint patch (bottom row) when T_1 is varied from 1 to 4. We observe that the high quality fingerprint patch has large SSIM values between patch p and its reconstruction $\hat{p}$, while the non-fingerprint patch results in very small SSIM values. The quality Q_p of the patch p is defined as

$$Q_p = \text{SSIM}(p, \hat{p}). \tag{14}$$

2.4.3 Ridge Quality Map, Orientation field and Frequency Field Estimation

The reconstructed patch $\hat{p}$ is divided into non-overlapping 16×16 blocks, as used in [12] and [10]. For each block, its orientation and frequency are obtained based on [12], resulting in reconstructed patch orientation field $\theta_{\hat{p}}$ (see Fig. 8) and frequency field $f_{\hat{p}}$. The reconstructed orientation field $\theta_{\hat{p}}$ and frequency $f_{\hat{p}}$ of patch $\hat{p}$ are used as the estimates of orientation field θ_p and frequency field f_p of the patch p in the latent image. For a region covered by multiple patches, its quality, orientation and frequency are defined as follows. The latent image is divided into non-overlapping 16×16 blocks. For each block b in the latent, let $\{q_i, \theta_i, f_i\}$ be the ridge quality, orientation and frequency of the ith patch covering the block b. The coarse estimates of ridge quality Q_b^c, orientation θ_b^c and frequency f_b^c are given as:

$$Q_b^c = \frac{1}{n_b} \sum_{i=1}^{n_b} q_i, \tag{15}$$

$$\theta_b^c = \frac{1}{2} \tan^{-1} \left(\sum_{i=1}^{n_b} q_i \sin 2\theta_i, \sum_{i=1}^{n_b} q_i \cos 2\theta_i \right), \tag{16}$$

$$f_b^c = \frac{1}{\sum_{i=1}^{n_s} q_i} \sum_{i=1}^{n_s} q_i f_i, \tag{17}$$

where n_b is the number of patches covering the block b.

2.4.4 Sparsity Control

The parameter T_1 in Eq. (9) is used to control the degree of sparsity in sparse coding. Figs. 7 and 8 show that while large values of T_1 lead to a better approximation of the input patch, it introduces unwanted noise. In particular, a single dictionary element is enough to recover level 1 features (orientation and frequency). For $T_1 = 1$, the coefficient vector α' becomes a scalar and the computation of α' is simplified to:

$$\alpha' = \frac{1}{N_p - 1} D_s^{cT} p, \tag{18}$$

where N_p is the number of pixels in patch p. Eq. (18) gives an efficient solution since there is no need to perform matrix inversion (see Eq. (11)). As the patch p is normalized to have the mean of 0 and standard deviation of 1, the SSIM index computation in Eq. (13) can be simplified to

$$\text{SSIM}(p, \hat{p}) = \frac{2\alpha'^2 + C_c}{1 + \alpha'^2 + C_c}. \tag{19}$$

It should be noted that when $T_1 = 1$, the reconstructed patch is an element selected from the dictionary D^c. Thus, the orientation and frequency fields of the dictionary elements can be pre-calculated in the off-line stage, which can greatly speed up the orientation and frequency fields estimation. For these reasons, T_1 is set to 1. Fig. 9(c) shows some examples of coarse quality maps under $T_1 = 1$.

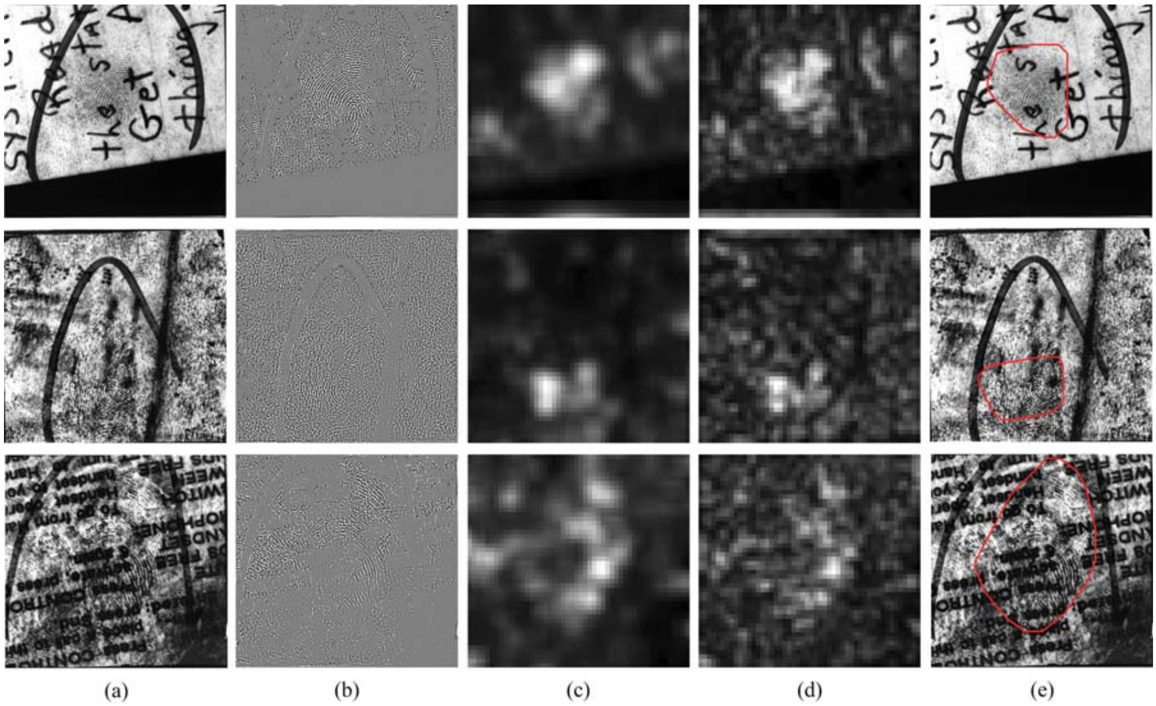


Fig. 9: Illustration of latent fingerprint segmentation. (a) Gray scale latent images, (b) extracted texture component images, (c) coarse quality maps, (d) fine quality maps, (e) segmentation results. The top, middle and bottom latent fingerprints in column (a) are of good, bad and ugly quality as defined in NIST SD27, respectively. The contrast of the middle and bottom latent fingerprints has been enhanced to get better visual quality.

2.5 Fine Estimates of Ridge Quality, Orientation and Frequency

The coarse-level dictionary while robust to local noise, does not enable us to extract detailed ridge information. Small patch size dictionaries enable us to compute the fine quality map and calculate the fine orientation and frequency fields:

1) All patches in P_L^f are normalized by Eq. (1) with mean of zero and standard deviation of one.
2) For each patch $p \in P_L^f$, a dominant orientation θ is obtained by averaging the coarse orientations of the blocks covered by patch p. Then, the corresponding orientation specific dictionary D_k^f is selected for patch p, where $k = \lceil \frac{16\times\theta}{\pi} \rceil$ and $\lceil \cdot \rceil$ is the ceiling operator.
3) The reconstructed patch $\hat{p}$ is obtained by solving the optimization problem in Eq. (9) with dictionary D_k^f.
4) The quality of patch p is determined by the structural similarity between p and $\hat{p}$ (Eq. (19)).
5) The block wise orientation and frequency fields of patch p are computed from $\hat{p}$ by the method in [12].
6) For each 16×16 block b in the latent, the fine quality Q_b^f, orientation θ_b^f and frequency f_b^f are obtained from the covering patches using Eqs. (15), (16) and (17).

2.6 Segmentation and Enhancement

2.6.1 Segmentation

The final quality map Q is computed as the average of coarse-level quality and fine-level quality by

$$Q = \frac{1}{2}(Q^c + Q^f). \tag{20}$$

The quality map Q is then normalized to the range [0,1]. A global threshold T_Q (determined from the normalized Q by Otsu's method [42]) is used to binarize the normalized quality map. The blocks with quality less than T_Q are regarded as background (i.e, 0), otherwise foreground (i.e., 1). To obtain the final segmentation results, morphological operations (dilation and opening) are then applied to remove small foreground blocks as well as to fill holes inside the foreground. Finally, the convex hull of the set of foreground blocks is computed to obtain the final segmentation result. Fig. 9(e) shows some examples of segmentation results of latent images in NIST SD27.

2.6.2 Enhancement

In the foreground region, the latent texture image obtained from the decomposition is enhanced by Gabor filtering [12]. The orientation and frequency parameters of the filter are tuned based on the fine-level orientation field (θ^f) and the average frequency of

Fig. 10: Four latent images (templates) that can be input to a COTS matcher. (a) Original latent image (G084 from NIST SD 27), (b) latent in (a) segmented and enhanced by the proposed algorithm, (c) latent in (a) segmented by the proposed segmentation algorithm, but enhanced by the algorithm in [10] and (d) latent segmented and enhanced by the ADTV model [25], (e) the true mate (rolled print). The retrieved ranks of the mated rolled print by COTS2 matcher for (a), (b), (c) and (d) are 5536, 1, 66 and 12462, respectively. The contrast of latent image in (a) has been enhanced for better visual quality.

coarse-level frequency field and fine-level frequency field ($\frac{f^f+f^c}{2}$); the standard deviation of the Gaussian envelope in the Gabor filter is set to 4.

3 Experimental Results

3.1 Databases

We use two latent databases for performance evaluation: NIST SD27 [43] and the West Virginia University latent database[4] (WVU DB) [44]. The NIST SD27 contains 258 latent fingerprints with their mated rolled fingerprints. The WVU DB contains 449 latent fingerprints with their mated rolled fingerprints and an additional 4,290 rolled fingerprints. The algorithm was implemented in MATLAB and C/C++ and run on a machine with Dual-Core 2.66GHz, 4GB RAM and Windows 7 operating system. The average computation time for segmentation and enhancement per latent is about 2.6 seconds for NIST SD27 and 1.6 seconds for WVU DB.

The ultimate goal of segmentation and enhancement of latent images is to improve the latent matching performance. The matching performance is evaluated using three commercial off-the-shelf (COTS) matchers (referred to as COTS1, COTS2 and COTS3); COTS1 and COTS2 are tenprint matchers and COTS3 is a latent matcher. One of the COTS tenprint matcher is VeriFinger SDK 6.3 [45], which has been widely used as a benchmark [10] [25]. We report the matching performance with tenprint matchers and latent matcher separately. Tenprint matchers are used to compare the proposed algorithm with other segmentation and enhancement algorithms reported in the literature (apparently a latent matcher was not available to these researchers). A state-of-the-art latent matcher is used to determine whether the proposed algorithm is able to boost its performance.

4. To request WVU latent fingerprint database, contact Dr. Jeremy Dawson (Email: Jeremy.Dawson@mail.wvu.edu)

3.2 Matching Performance with Tenprint Matcher

In this section, we determine whether the performance of two COTS tenprint matchers can be improved by using the proposed segmentation and enhancement algorithm. We also compared the proposed algorithm with two other algorithms in the literature: i) dictionary of reference orientation patches proposed by Feng *et al.* [10] and ii) adaptive directional total-variation proposed by Zhang *et al.* [25]. Since the algorithm in [10] is only for latent enhancement and requires ROI mask as input, we take the ROI mask obtained from the proposed segmentation as input to [10] for a fair comparison. The following matching scenarios are considered for each latent in NIST SD27 (see Fig. 10):

1) Baseline: Input to the COTS matchers is the original gray scale latent image.
2) Proposed algorithm: Input to the COTS matchers is the segmented and enhanced latent image by the proposed algorithm.
3) Match score fusion: The match scores with and without the proposed segmentation and enhancement are fused by weighted sum. The weights are empirically set as 0.7 and 0.3, respectively.
4) Enhancement in [10]: Input to the COTS matchers is the enhanced latent image by [10] with the proposed algorithm's segmentation mask as input.
5) ADTV model in [25]: Input to the COTS matchers is the segmented and enhanced image by [25].

Since the latents in NIST SD27 used in [25] are 1000 ppi compared to 500 ppi images used in our experiments, we use the segmentation and enhancement results provided by the authors [25] for comparison. Since [25] does not contain results for WVU DB, only four matching results (excluding the ADTV model) are compared for WVU DB. The Cumulative Match Characteristic (CMC) curves of these scenarios

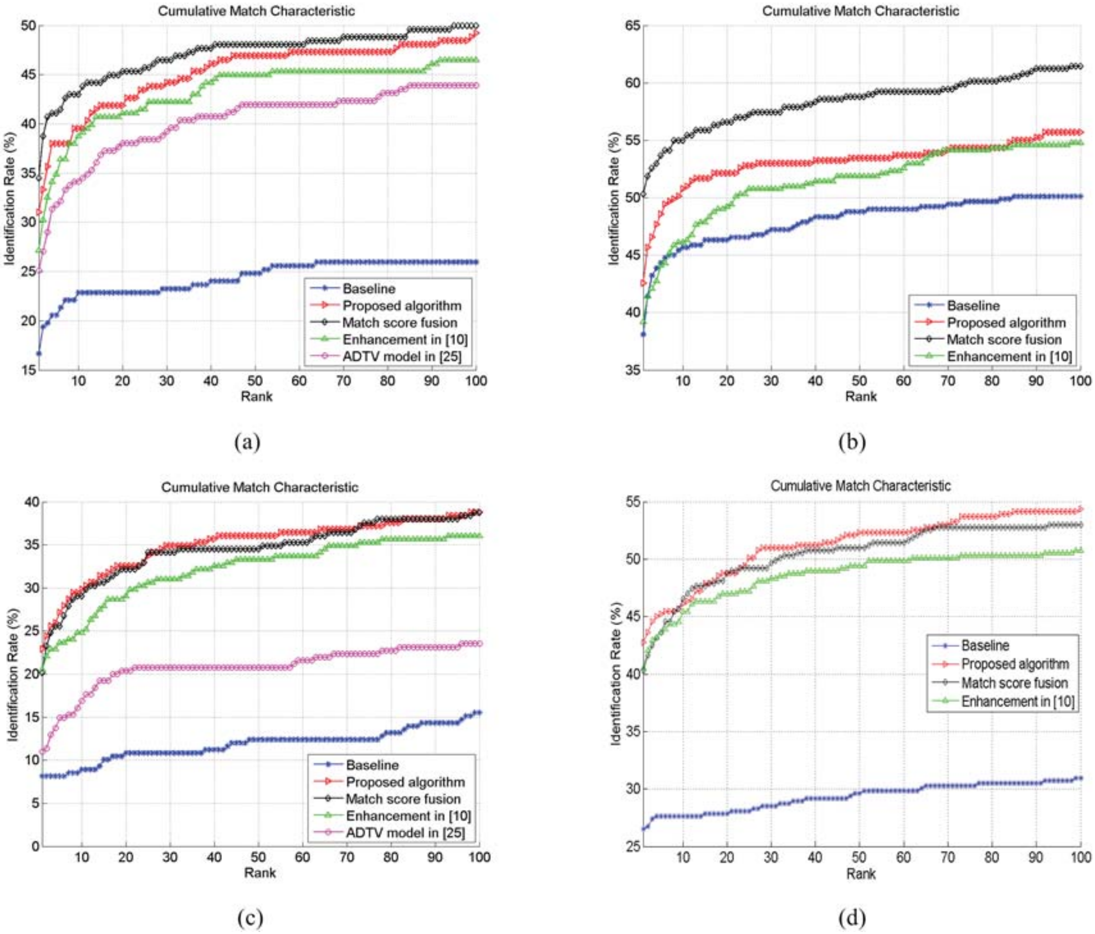
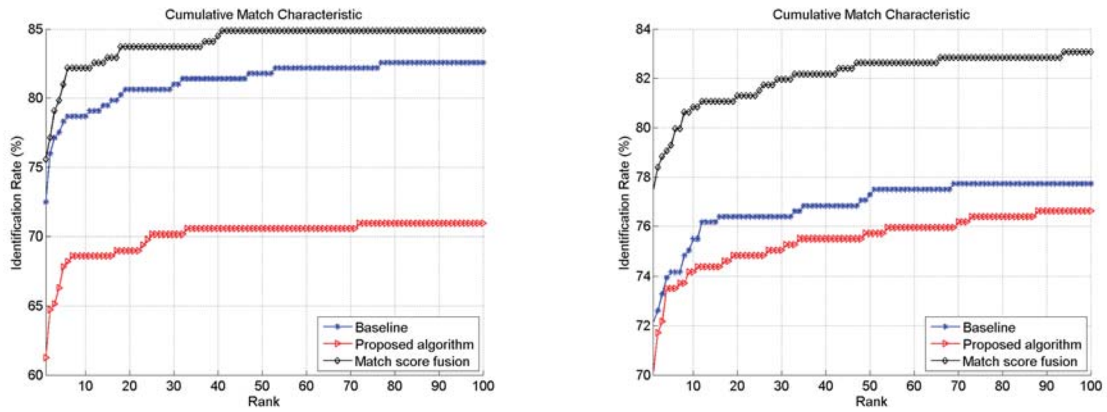


Fig. 11: CMC curves for latent matching under different matching scenarios by COTS1 matcher on (a) NIST SD27 and (b) WVU DB and by COTS2 matcher on (c) NIST SD27 and (d) WVU DB.

on NIST SD27 and WVU DB are shown in Fig. 11. Compared to the baseline, the rank-1 identification rate for COTS1 matcher improved by 14.34% on NIST SD27 and by 4.45% on WVU DB by incorporating latent segmentation and enhancement by the proposed algorithm. The rank-1 identification rate for COTS2 matcher improved by 14.73% on NIST SD27 and by 16.26% on WVU DB. The proposed algorithm also outperforms the algorithms in [10] and [25] for both tenprint matchers. As shown in Fig. 10, the orientation field for the ADTV model [25], extracted by the gradient-based approach, is not robust to structured noise. The algorithm in [10] is not able to account for the variation in ridge frequency in our algorithm that is commonly encountered in poor quality latents. Fig. 12 shows an example where the automatically extracted ridge frequency works better than fixed ridge frequency used in [10]. After fusing the COTS1 match scores using the baseline and the proposed algorithm, the rank-1 identification performance can be further improved (from 31.01% to 34.50% on NIST SD27 and from 42.54% to 50.33% on WVU DB). However, this fusion does not help the COTS2 matcher because there is a large gap in the performance between the baseline and our algorithm.

3.3 Performance with Latent Matcher

For each latent, we input two images to the COTS3 latent matcher, i.e. original latent image (baseline) and segmented and enhanced latent image (by the proposed algorithm), as shown in Figs. 3 (a) and (c), respectively. To evaluate if the proposed segmentation and enhancement can boost the performance of a state-of-the-art latent matcher, the match scores from the two input images (original latent and the enhanced and segmented latent by the proposed algorithm) are fused by a weighted sum method (the weights for original latent image and segmented and enhanced are empirically set as 0.7 and 0.3, respectively). The resulting CMC curves for the COTS3 latent matcher on NIST SD27 and WVU DB are shown in Fig. 13. Although the performance with the segmented and enhanced latent is lower than by just that with the original latent image, after fusion, the rank-1 identification rate of COTS3 latent matcher increases from 72.48% to 75.58% for NIST SD27 and from 72.16% to 77.51% for WVU DB. One



(a) (b)

Fig. 13: CMC curves for latent matching with COTS3 matcher, a state of the art latent matcher, under different matching scenarios on (a) NIST SD27 and (b) WVU DB.

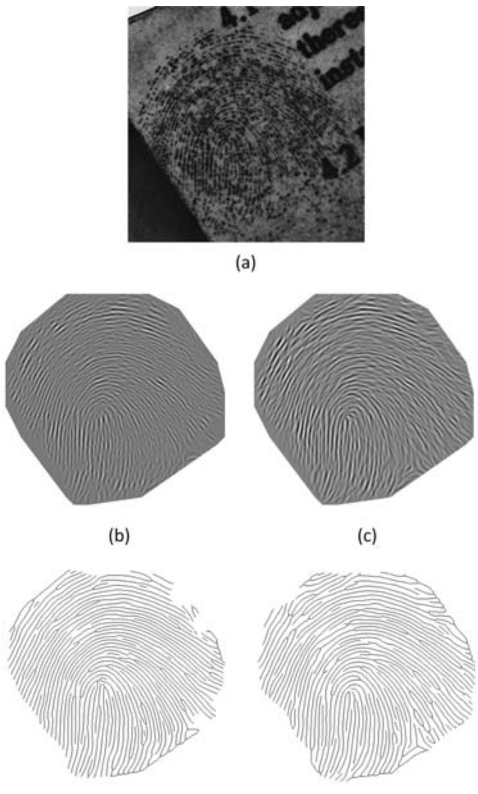


Fig. 12: Latent enhancement results. (a) Original latent image (G032 from NIST SD 27), (b) enhanced latent by the proposed algorithm and (c) enhancement by the algorithm in [10], (d) and (e) are the skeletons of (b) and (c), respectively. The retrieved ranks of the mated rolled print by the COTS2 matcher for (a), (b) and (c) are 13,718, 1 and 10,106, respectively. The contrast of latent image in (a) has been enhanced for better visual quality.

of the main objectives of this paper is to develop a strategy for latent segmentation and enhancement which could lead to a novel template of the latent. This template could be used by any COTS matcher in addition to the propriety templates that the matchers generate internally for a latent. In fact, a fusion of diverse search templates from different segmentation and enhancement algorithms is a common technique to boost the latent matching performance. Hara [46] shows some examples of such a fusion in the NIST latent testing workshop. Fig. 14 shows queries which are correctly retrieved at rank 1 by the COTS3 latent matcher after match score fusion with the proposed algorithm.

3.4 Confidence Value

Some latents that are of extremely poor quality (see Fig. 15) cannot be segmented and enhanced correctly. Therefore, it is necessary to provide a confidence value for latent segmentation and enhancement. If the confidence value is high, it means the segmented and enhanced latent image is suitable for lights-out operation without any human intervention. The confidence value of latent segmentation and enhancement (C) is defined as

$$C = \sum_{(x,y)\in F} Q(x,y)/|F| \tag{21}$$

where F is the segmented foreground and $Q(x,y)$ is the quality at point (x,y) defined in Eq. (20). To validate the usability of the proposed confidence value, we analyzed the latent identification accuracy at various rejection rates (rejected latents will be manually processed) based on the confidence values. At a rejection rate of 20%, the rank-1 identification rate accuracy increases by 7.43% and 7.23%, respectively, for the NIST SD27 and WVU DB using the COTS3 latent matcher. Apparently, both these databases con-

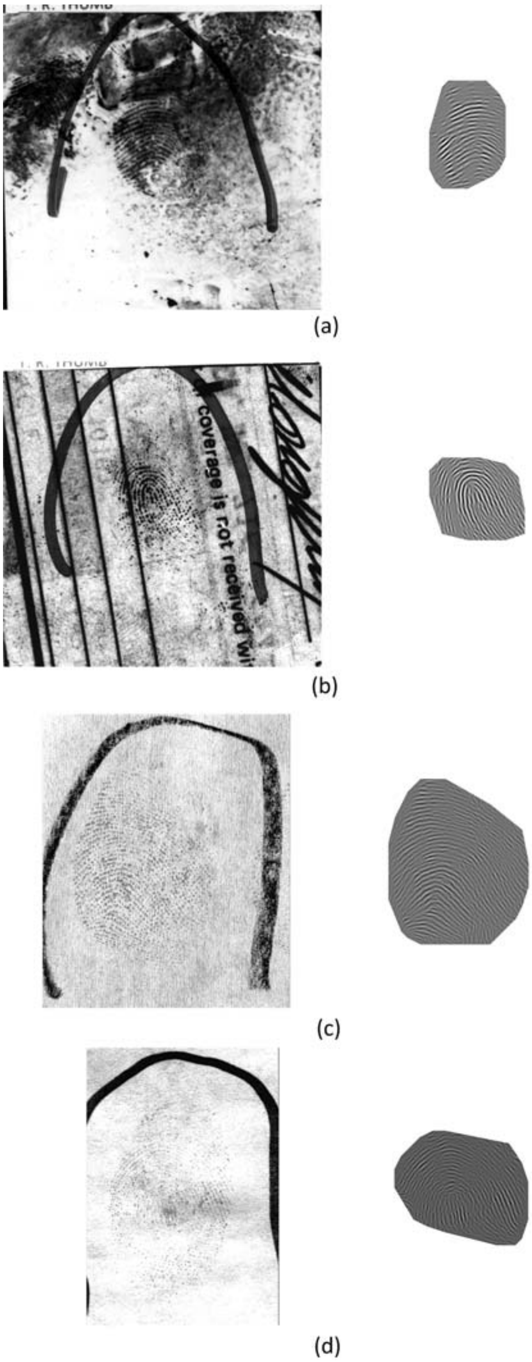


Fig. 14: Examples of latent images which are correctly identified at rank 1 by the COTS3 latent matcher after match score fusion. (a), (b), (c), (d) show original latents (left column) and the segmented and enhanced latents by the proposed algorithm (right column). (a) and (b) are latents from NIST SD27, (c) and (d) are latents from WVU DB. The mated rolled fingerprints of original latents of (a), (b), (c) and (d) are retrieved at ranks 5, 4, 31,997 and 31,997, respectively. The contrast of the latent fingerprints in (c) and (d) has enhanced for better visual quality.

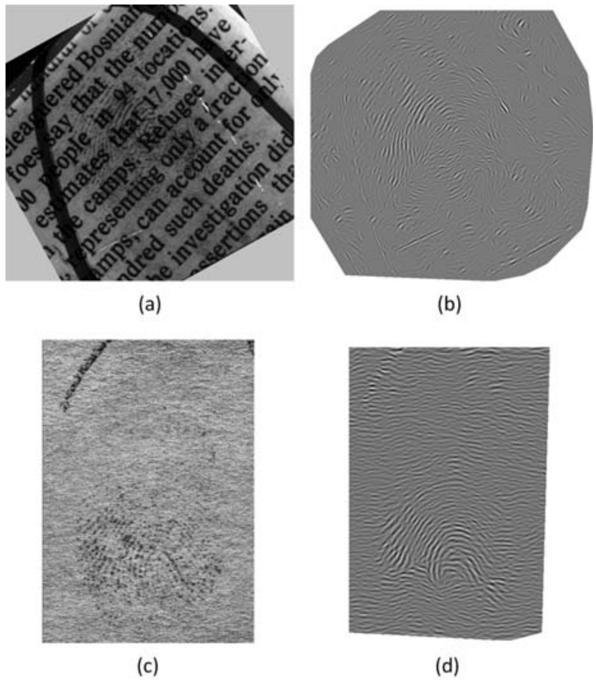


Fig. 15: Examples of poor quality latents. (a) and (c) are original latent images from NIST SD27 and WVU DB, respectively. (b) and (d) are their segmented and enhanced latent images by the proposed algorithm. The contrast of latent image in (c) has been enhanced for better visual quality.

tain some inherently challenging latent images, so the error-reject tradeoff is not as effective.

4 Conclusions and Future Work

Automatic Fingerprint identification Systems (AFIS) have proven to be invaluable not only in identifying suspects and prosecuting criminals, but also in growing applications of national civil registry systems and verifying the identity of travelers at international border crossings. Although state of the art AFIS have already achieved impressive accuracy in tenprint search (rolled prints or slaps), latent matching or search is still a challenging problem due to presence of complex background noise and poor quality of friction ridge structure in many latents. We have proposed an automatic latent segmentation and enhancement algorithm based on image decomposition and coarse to fine ridge structure dictionaries. Experimental results on two different latent fingerprint databases, NIST SD27 and WVU DB, in conjunction with three different COTS matchers show that the proposed algorithm is able to improve the performance of two COTS tenprint matchers and can even boost the performance of a state-of-the-art latent matcher by weighted match score fusion. However, the proposed algorithm still does not work well on very poor quality latent fingerprint images. Our algorithm can be further improved along the following aspects:

1) A robust patch quality definition, especially for dry fingerprint images, where ridges are broken.
2) A better definition of confidence measure for the segmentation and enhancement results.
3) Improve the computational efficiency of the algorithm.

ACKOWNLEDGMENTS

The authors would like to thank Prof. Jay Kuo and Jianyang Zhang for providing us their segmentation and enhancement results. Kai Cao and Eryun Liu's research were partially supported by National Natural Science Foundation of China (Grant No. 61101247 and 61100234). Anil Jain's research was partially supported by the Brain Korea 21 PLUS Program through the National Research Foundation of Korea, funded by the Ministry of Education. All correspondence should be addressed to Anil Jain.

REFERENCES

[1] C. Wilson, R. A. Hicklin, H. Korves, B. Ulery, M. Zoepfl, M. Bone, P. Grother, R. Micheals, S. Otto, and C. Watson, "Fingerprint vendor technology evaluation 2003: Summary of results and analysis report," *NISTIR 7123*, 2004.

[2] M. D. Indovina, V. Dvornychenko, R. A. Hicklin, and G. I. Kiebuzinski, "Evaluation of latent fingerprint technologies: Extended feature sets (evaluation 2)," *Technical Report NISTIR 7859, NIST*, 2012.

[3] B. T. Ulery, R. A. Hicklin, J. Buscaglia, and M. A. Roberts, "Repeatability and reproducibility of decisions by latent fingerprint examiners," *PloS one*, vol. 7, no. 3, p. e32800, 2012.

[4] ——, "Accuracy and reliability of forensic latent fingerprint decisions," *Proceedings of the National Academy of Sciences*, vol. 108, no. 19, pp. 7733–7738, 2011.

[5] I. E. Dror, K. Wertheim, P. Fraser-Mackenzie, and J. Walajtys, "The impact of human-technology cooperation and distributed cognition in forensic science: Biasing effects of afis contextual information on human experts," *Journal of Forensic Sciences*, vol. 57, no. 2, pp. 343–352, 2011.

[6] Department of Justice, "A review of the fbi's handling of the Brandon Mayfield case." [Online]. Available: http://www.justice.gov/oig/special/s0601/exec.pdf

[7] "Innocence project," http://www.innocenceproject.org.

[8] M. D. Indovina, V. N. Dvornychenko, E. Tabassi, G. W. Quinn, P. J. Grother, S. Meagher, and M. D. Garris, "ELFT phase II - an evaluation of automated latent fingerprint identification technologies," *NISTIR 7577*, 2009.

[9] "FBI- Next Generation Identification (NGI)," http://www.fbi.gov/about-us/cjis/fingerprints biometrics/ngi.

[10] J. Feng, J. Zhou, and A. K. Jain, "Orientation field estimation for latent fingerprint enhancement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 54, no. 4, pp. 925–940, 2013.

[11] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. Springer, 2009.

[12] L. Hong, Y. Wan, and A. Jain, "Fingerprint image enhancement: Algorithm and performance evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 777–789, 1998.

[13] X. J. Chen, J. Tian, J. G. Cheng, and X. Yang, "Segmentation of fingerprint images using linear classifier," *EURASIP*, vol. 4, pp. 480–494, 2004.

[14] E. Zhu, J. Yin, C. Hu, and G. Zhang, "A systematic method for fingerprint ridge orientation estimation and image segmentation," *Pattern Recognition*, vol. 39, no. 8, pp. 1452–1472, 2006.

[15] S. Chikkerur, A. N. Cartwright, and V. Govindaraju, "Fingerprint enhancement using STFT analysis," *Pattern Recognition*, vol. 40, no. 1, pp. 198–211, 2007.

[16] A. Bazen and S. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 905–919, 2002.

[17] Y. Wang, J. Hu, and D. Phillips, "A fingerprint orientation model based on 2d fourier expansion (FOMFE) and its application to singular-point detection and fingerprint indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 573–585, 2007.

[18] F. Turroni, D. Maltoni, R. Cappelli, and D. Maio, "Improving fingerprint orientation extraction," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1002–1013, 2011.

[19] X. Jiang, "Fingerprint image ridge frequency estimation by higher order spectrum," in *IEEE International Conference on Image Processing*, vol. 1, 2000, pp. 462–465.

[20] S. Karimi-Ashtiani and C.-C. Kuo, "A robust technique for latent fingerprint image segmentation and enhancement," in *IEEE International Conference on Image Processing*, 2008, pp. 1492 –1495.

[21] N. J. Short, M. S. Hsiao, A. L. Abbott, and E. A. Fox, "Latent fingerprint segmentation using ridge template correlation," in *4th International Conference on Imaging for Crime Detection and Prevention*, 2011, pp. 1–6.

[22] H.-S. Choi, M. Boaventura, I. A. G. Boaventura, and A. K. Jain, "Automatic segmentation of latent fingerprints," in *BTAS*, 2012.

[23] S. Yoon, J. Feng, and A. K. Jain, "On latent fingerprint enhancement," in *Proceedings of the SPIE Biometric Technology for Human Identification VII*, 2010, pp. 766 707–766 707–10.

[24] S. Yoon, J. Feng, and A. Jain, "Latent fingerprint enhancement via robust orientation field estimation," in *2011 International Joint Conference on Biometrics (IJCB)*, 2011, pp. 1–8.

[25] J. Zhang, R. Lai, and C.-C. Kuo, "Adaptive directional total-variation model for latent fingerprint segmentation," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 8, pp. 1261–1273, 2013.

[26] J. Zhang, R. Zhang, and C.-C. Kuo, "Latent fingerprint segmentation with adaptive total variation model," in *5th IAPR International Conference on Biometrics (ICB)*, 2012, pp. 189–195.

[27] A. Buades, T. Le, J.-M. Morel, and L. Vese, "Fast cartoon + texture image filters," *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 1978 –1986, 2010.

[28] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736 –3745, 2006.

[29] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53 –69, 2008.

[30] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Discriminative learned dictionaries for local image analysis," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.

[31] X.-C. Lian, Z. Li, C. Wang, B.-L. Lu, and L. Zhang, "Probabilistic models for supervised dictionary learning," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2305 –2312.

[32] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[33] S. Liao, A. K. Jain, and S. Z. Li, "Partial face recognition: Alignment-free approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1193–1205, 2013.

[34] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397 –3415, 1993.

[35] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600 –612, 2004.

[36] "NIST Special Database 4," http://www.nist.gov/srd/nistsd4.cfm.

[37] E. Tabassi, C. Wilson, and C. Watson, "Fingerprint image quality," *NISTIR 7151*, 2004.

[38] M. D. Garris, E. Tabassi, C. I. Wilson, R. M. McCabe, S. Janet, and C. I. Watson, "NIST fingerprint image software 2," 2004. [Online]. Available: http://www.nist.gov/itl/iad/ig/nbis.cfm
[39] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311 – 4322, 2006.
[40] B. Girod, "Digital images and human vision," A. B. Watson, Ed., 1993, ch. What's wrong with mean-squared error?, pp. 207–220.
[41] P. Teo and D. Heeger, "Perceptual image distortion," in *IEEE International Conference Image Processing (ICIP)*, vol. 2, 1994, pp. 982 –986 vol.2.
[42] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62 –66, 1979.
[43] "NIST Special Database 27," http://www.nist.gov/srd/nistsd27.cfm.
[44] "Integrated pattern recognition and biometrics lab, West Virginia University," http://www.csee.wvu.edu/ ross/i-probe/.
[45] Neurotechnology Inc, "Verifinger," http://www.neurotechnology.com/verifinger.html.
[46] Masanori Hara, NEC, "Lights-out latent processing today and tomorrow," http://biometrics.nist.gov/cs_links/latent/workshop09/proc /NEC_AutoL _LatentWS2009MAR_v4.pdf.

Anil K. Jain is a university distinguished professor in the Department of Computer Science and Engineering at Michigan State University. His research interests include pattern recognition and biometric authentication. He served as the editor-in-chief of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (1991-1994). The holder of six patents in the area of fingerprints, he is the author of a number of books, including Introduction to Biometrics (2011), Handbook of Face Recognition (2011), Handbook of Fingerprint Recognition (2009), Handbook of Biometrics (2009), Handbook of Multibiometrics (2006), BIOMETRICS: Personal Identification in Networked Society (1999), and Algorithms for Clustering Data (1988). He served as a member of the Defense Science Board and The National Academies committees on Whither Biometrics and Improvised Explosive Devices. Dr. Jain received the 1996 IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding Paper Award and the Pattern Recognition Society best paper awards in 1987, 1991, and 2005. He is a fellow of the AAAS, ACM, IAPR, and SPIE. He has received Fulbright, Guggenheim, Alexander von Humboldt, IEEE Computer Society Technical Achievement, IEEE Wallace McDowell, ICDM Research Contributions, and IAPR King-Sun Fu awards.

Kai Cao Kai Cao received the Ph.D. degree from the Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2010. He is currently a Post Doctoral Fellow in the Department of Computer Science & Engineering, Michigan State University, East Lansing, USA. He is also affiliated with Xidian University as an Associate Professor. His research interests include biometric recognition, image processing and machine learning.

Eryun Liu received his Bachelor degree in electronic information science and technology from Xidian University, Xi'an, Shaanxi China, in 2006 and Ph.D. degree in Pattern Recognition and Intelligence System from the same university in 2011. He is currently a Post Doctoral Fellow in the Department of Computer Science & Engineering, Michigan State University, East Lansing, USA. He is also affiliated with Xidian University as an Assistant Professor. His research interests include biometric recognition, point pattern matching and information retrieval, with a focus on fingerprint and palmprint recognition.