Adaptive Directional Total-Variation Model for Latent Fingerprint Segmentation

Jiangyang Zhang, *Student Member, IEEE,* Rongjie Lai, and C.-C. Jay Kuo, *Fellow, IEEE*

***Abstract*—A new image decomposition scheme, called the adaptive directional total variation (ADTV) model, is proposed to achieve effective segmentation and enhancement for latent fingerprint images in this work. The proposed model is inspired by the classical total variation models [1], but it differentiates itself by integrating two unique features of fingerprints; namely, scale and orientation. The proposed ADTV model decomposes a latent fingerprint image into two layers: cartoon and texture. The cartoon layer contains unwanted components (e.g. structured noise) while the texture layer mainly consists of the latent fingerprint. This cartoon-texture decomposition facilitates the process of segmentation, as the region of interest can be easily detected from the texture layer using traditional segmentation methods. The effectiveness of the proposed scheme is validated through experimental results on NIST SD27 latent fingerprint database. The proposed scheme achieves accurate segmentation and enhancement results, leading to improved feature detection and latent matching performance.**

***Index Terms*—Latent fingerprints, total variation, fingerprint recognition, fingerprint segmentation.**

I. Introduction

Latent fingerprint identification plays a critical role for law enforcement agencies in identifying and convicting criminals. An important step in an automated fingerprint identification systems (AFIS) is the process of fingerprint segmentation. While a tremendous amount of efforts has been made on plain and rolled fingerprint segmentation, latent fingerprint segmentation remains to be a challenging problem. Collected from crime scenes, latent fingerprints are often mixed with other components such as structured noise or other fingerprints. Existing fingerprint recognition algorithms fail to work properly on latent fingerprint images, since they are mostly applicable under the assumption that the image is already properly segmented and there is no overlap between the target fingerprint and other components.

Fingerprint segmentation refers to the process of decomposing a fingerprint image into two disjoint regions: foreground and background. The foreground, also called the region of interest (ROI), consists of the desired fingerprints while the background contains noisy and irrelevant contents that will be discarded in the following processing steps. Accurate fingerprint segmentation is critical as it affects the accurate extraction of minutiae and singular points, which are key features for fingerprint matching. When feature extraction algorithms are applied on a fingerprint image without segmentation, lots of false features may be extracted due to the presence of noisy background, and eventually leading to matching errors in the later stage. Therefore, the goal of fingerprint segmentation is to discard the background, reduce the number of false features, and thus improve the matching accuracy.

Based on the collection procedure, fingerprint images can generally be divided into three categories, namely, *rolled*, *plain* and *latent* [2]. Rolled fingerprints are obtained from rolling the finger from one side to the other in order to capture all ridge details of the fingerprint. Plain fingerprints images are acquired by pressing the fingertip onto a flat surface. Both rolled and plain prints are obtained in an attended mode, so they are usually of good visual quality and contain sufficient information for reliable matching. On the contrary, latent fingerprints are usually collected from crime scenes, in which the print is lifted from object surfaces that were inadvertently touched or handled. The matching between latents and rolled/plain fingerprints plays a crucial role in identifying suspects by law enforcement agencies.

Segmentation on rolled and plain fingerprint images has been well-studied in literature. In the early work of [3], segmentation was achieved by partitioning the fingerprint image into blocks, followed by block classification based on gradient and variance information. This method was further extended to a composite method [4] that takes advantage of both the directional and variance approaches. Ratha *et al.* [5] considered the gray-scale variance along the direction orthogonal to the ridge-flow orientation as the key feature for block classification. In [6], fingerprints were segmented using three pixel-level features (coherence, mean and variance). An optimal linear classifier was trained for pixel-based classification and morphology operators were applied to obtain compact segmentation clusters.

While significant effort has been made on developing segmentation algorithms for rolled/plain fingerprints, latent fingerprint segmentation remains to be a challenging problem. Although automated identification has already achieved high accuracy for plain/rolled fingerprints, manual intervention is still necessary for latent prints processing [2]. The difficulty mainly lies in: 1) the poor quality of fingerprint patterns in terms of the clarity of ridge information, and 2) the presence of various structured noise in the background. Traditional segmentation methods fail to work properly on latent fingerprints as they are based on many assumptions that are only valid for rolled/plain fingerprints. In recent works on latent fingerprints [7], [8], [9], the region-of-interest (ROI) is still manually marked and assumed to be known. Undoubtedly,

Jiangyang Zhang and C.-C. Jay Kuo are with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, 90089

Rongjie Lai is with the Department of Mathematics, University of Southern California, Los Angeles, CA, 90089.

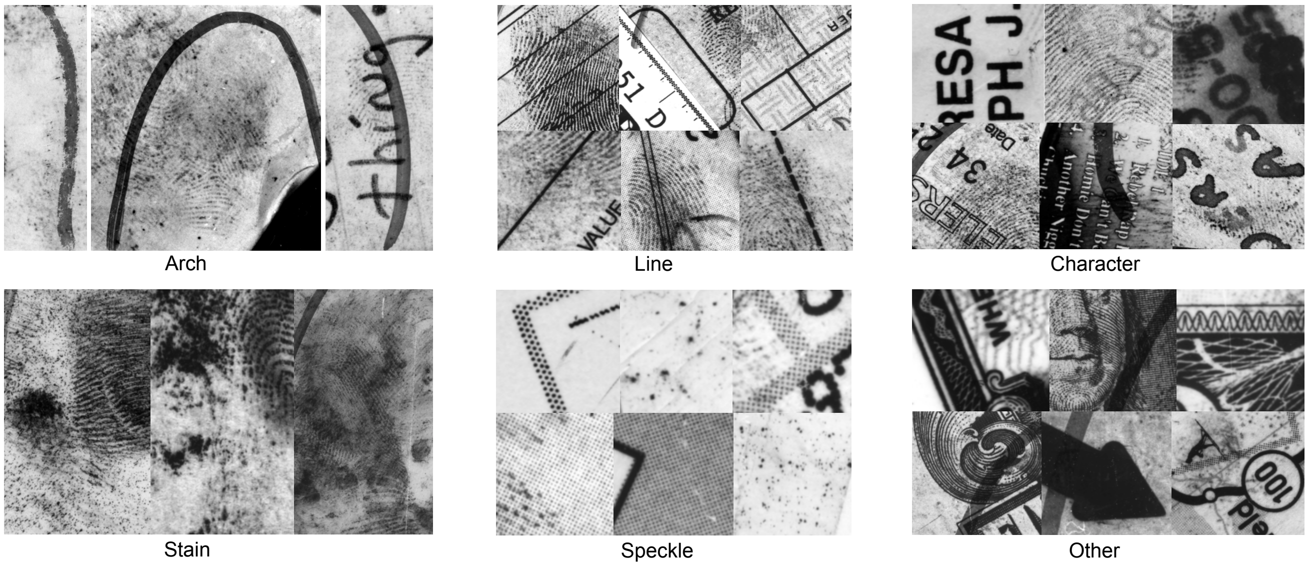


Fig. 1. Illustration of six types of structured noise in latent fingerprint images.

accurate and robust latent segmentation is an essential step towards achieving automatic latent identification, and it is the main focus of our current research.

Most recently, several studies [10], [11] have been conducted to address the problem of latent fingerprint segmentation. Karimi-Ashtiani and Kuo [10] used a projection method to estimate the orientation and frequency of local blocks. After projection, the distance between center-of-transient points measures the amount of data degradation and used for segmentation. Short *et al.* [11] formulated a ridge model template and used the cross-correlation between a local block and the generated template to assign one of six quality scores. Blocks with high quality score are labeled as foreground while the rest are treated as background. While the proposed methods demonstrated improved performance in handling latent fingerprint images, experimental results show that their performances are still limited by the presence of structured noise.

Total-Variation-based (TV-based) image models have been widely used in the context of image decomposition [12], [13]. Among several well known TV-based models, the model using total variation regularization with an L1 fidelity term, denoted by the TV-L1 model, is especially suited for multi-scale image decomposition and feature selection [14], [15]. Besides, a modified TV-L1 model was adopted in [15] to extract small-scale facial features for facial recognition under varying illumination. More recently, the authors proposed an adaptive TV-L1 model for latent fingerprint segmentation in [16], where the fidelity weight coefficient is adaptively adjusted to the background noise level. Furthermore, the Directional Total Variation (DTV) model was formulated in [17] by imposing the directional information on the TV term, which proved to be effective for latent fingerprint detection and segmentation. It appears that the TV-based image model with proper adaptation offers a suitable tool for latent fingerprint segmentation. However, the performance of both models in [16], [17] was evaluated only subjectively, as no objective evaluation was performed to determine whether the proposed scheme improved the matching accuracy, which is the ultimate goal for fingerprint segmentation.

In this paper, we take advantage of both TV methods in [16], [17] and combine them into one single model, called the Adaptive Directional Total Variation (ADTV) model. Both the anisotropic directional TV term and the spatially-adaptive fidelity weight are incorporated into the model formulation. The proposed ADTV model decomposes a latent fingerprint image into two layers: cartoon and texture. The cartoon layer contains the unwanted components (e.g. structured noise) while the texture layer mainly consists of the latent fingerprint. This cartoon-texture decomposition facilitates the process of segmentation, as the region of interest can be easily detected from the texture layer using traditional segmentation methods. In addition, the effectiveness of our proposed scheme is validated through experiments on feature detection and latent matching. As compared with our prior work, the materials in Sections III.C and IV are new.

The rest of this paper is organized as follows. In Section 2, we first examine several forms of structured noise that commonly appear in latent fingerprint images. In Section 3, we introduce the proposed ADTV model and explain how it can be utilized for latent fingerprint image decomposition and segmentation. In Section 4, we validate the effectiveness of our proposed scheme through a series of benchmarking experiments. Concluding remarks are given in Section 5.

II. Structured Noise in Latent Fingerprint Images

The difficulty for latent fingerprint segmentation mainly lies in two aspects. On one hand, the fingerprint itself is usually of very poor quality, often with smudged or blurred ridges. It is very common that the image contains only partial area of the finger and large nonlinear distortions exist due to pressure variations. As a result, while a typical rolled fingerprint has around 80 minutiae, a latent fingerprint contains only about 15 usable minutiae with reasonable quality [2].

On the other hand, the presence of various types of structured noise further hinders the proper segmentation for latent prints. As compared with the oscillatory ridge structures of fingerprints, structured noise are usually of much larger scale and can appear in various forms. Based on appearance, structured noise can be classified into six categories: *arch*, *line*, *character*, *speckle*, *stain* and *others*. They are shown in Fig. 1 and elaborated below.

1) *Arch*. The big arch is manually marked by crime-scene investigators to indicate the possible existence of latent fingerprints in the region encircled by the arch. The arch noise is considered to be the simplest type of structured noise.
2) *Line*. The line noise may appear in the format of a single line or multiple parallel lines. Single line is usually detected and removed using methods based on Hough transform [11]. Multiple parallel lines are easily confused with fingerprints since they share many common features.
3) *Character*. The most common type of structured noise that appears in latent fingerprint images. The characters may appear in various font types, sizes, brightness, and can be either handwritten or typed.
4) *Stain*. It is generated when the finger, instead of being properly pressed, was inadvertently smeared on a wet or dirty surface. Stain noise often appears in spongy shape with inhomogeneous brightness.
5) *Speckle*. As compared with lines and characters, the speckle noise tends to consist of tiny-scale structures, which can be either regular (e.g., clusters of small dots) or random (e.g., ink and dust speckles).
6) *Others*. A latent fingerprint image may contain other structured noises such as arrows, signs, etc. Similar to arch and character noise, they usually consist of smooth surfaces with sharp edges.

The line, character and speckle noise often appear when the latent fingerprint is lifted from the surface of a text document (e.g. maps, newspapers, checks, etc).

For latent fingerprint segmentation, the main challenge lies in how to effectively separate latent fingerprints, the relatively weak signal, from all structured noise in the background, which is often the dominant image component. Additional complexity arises when structured noise overlaps with the fingerprint signal. Previous methods proposed for fingerprint segmentation are mostly feature-based, and features commonly used for segmentation include the mean, variance, contrast, coherence as well as their variants [6], [18], [19]. However, these methods may fail to work properly on latent fingerprints as they are based on many assumptions that are only valid for rolled/plain fingerprints. For instance, in [6], the mean feature was used since the background was assumed to be bright and the variance feature was used since the variance of background noise was assumed to be much lower than that of fingerprint regions. However, these assumptions are no longer valid in the context of latent fingerprint images.

To evaluate the effectiveness of traditional segmentation features, we manually segment a plain and a latent fingerprint image, and plot the distributions of three segmentation features, namely, the mean, variance and coherence, for both foreground and background regions. As shown in the Fig. 2, the distributions of these features in foreground and background regions are well separated for plain fingerprints, while those of latent fingerprints have significant overlaps. These overlaps can be explained by two reasons. First, regions with structured noise often have high contrast and coherent gradient orientations as well, so it's difficult to differentiate them from fingerprints using these features. Second, the qualities of some latent fingerprints are so poor that they cannot be well characterized by traditional fingerprint features. As a result, new features or models need to be considered for more effective separation of latent fingerprint and structured noises.

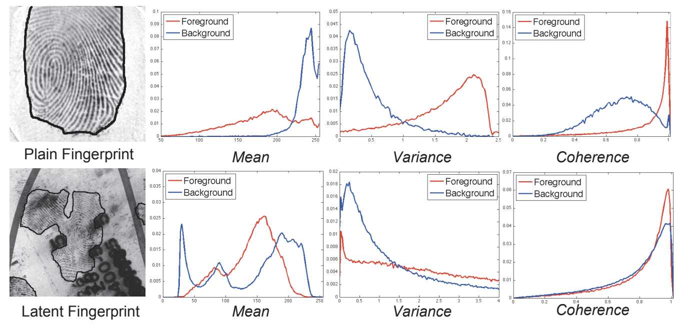


Fig. 2. Comparison of distributions of three features (mean, variance and coherence) in the foreground and background areas of plain and latent fingerprints.

III. Latent Segmentation with ADTV Model

In this section, we introduce the proposed Adaptive Directional Total Variation (ADTV) model and explain how it can be used to effectively separate latent fingerprint with structured noises, and thus facilitate the process of fingerprint segmentation. We begin with introducing the TV-L1 model, which serves as the basis for the proposed ADTV model, and explain its capability in multiscale feature selection. Finally, we propose the ADTV model and discuss the choice of its parameters.

A. The TV-L1 Model

TV-based image models have been widely studied to achieve the task of image decomposition. Among many existing TV models, the total variation regularization model with an L1 fidelity term, denoted by TV-L1, is suitable for multiscale image decomposition and feature selection. In the context of facial recognition under varying illumination, a modified TV-L1 model was proposed in [15] to separate small-scale facial features with nonuniform illumination, and thus leads to improved recognition result.

Similar to other TV-based image models (*e.g.,* the ROF model [1]), the TV-L1 model decomposes an input image, f, into two signal layers:

- Cartoon u, which consists of the piecewise-smooth component in f, and
- Texture v, which contains the oscillatory or textured component in f.

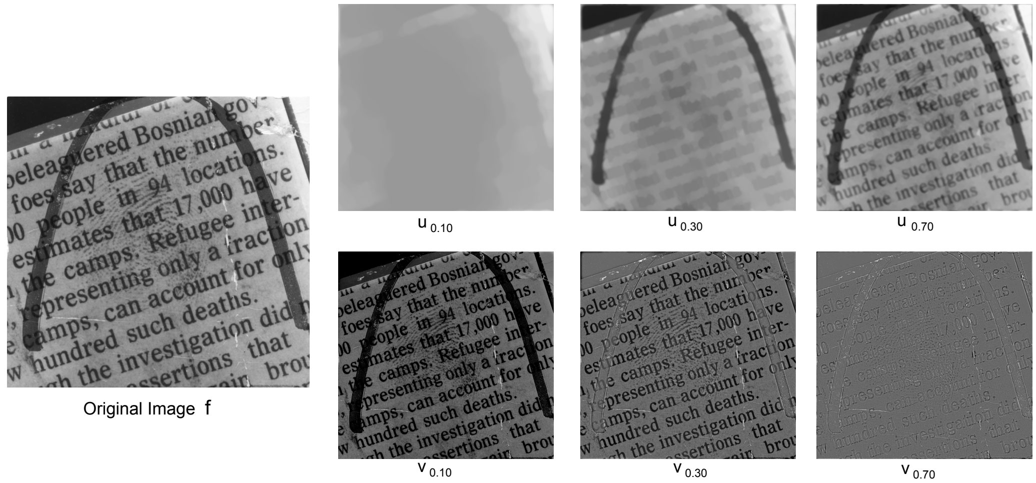


Fig. 3. Feature selection based on the TV-L1 model for latent fingerprint image: input image f (left most) and its TV-L1 decomposed components u and v with its λ value shown in the subscript. As λ increases, only features of smaller scales are extracted to texture output v, while features of larger scales are kept in cartoon u.

The decomposition

$$f = u + v,$$

is obtained by solving the following variational problem:

$$\min_u \int |\nabla u| + \lambda \int |u - f| \, dx, \tag{1}$$

where f, u and v are functions of image gray-scale intensity values in $\mathbb{R}^2$, ∇u is the gradient value of u and λ is a constant weighting parameter. We call $\int |\nabla u|$ and $\int |u - f|$ the total variation of u and the fidelity term, respectively.

The TV-L1 model is difficult to compute due to nonlinearity and non-differentiability of the total variation term as well as the fidelity term. A gradient descent approach was proposed in [14], which solves for u as a steady solution of the Euler-Lagrange equation of (1):

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda \frac{f - u}{|f - u|} = 0. \tag{2}$$

Although (2) is easier to implement, the gradient descent approach is slow due to a small time step imposed by the strict stability constraint. That is, the term $\frac{f-u}{|f-u|}$ is non-smooth at $f - u$, which forces the time step to be very small when the solution is approaching the steady state. In addition, $|\nabla u|$ in the term $\frac{\nabla u}{|\nabla u|}$ might be zero, and a small positive constant needs to be added to avoid zero division, which results in inexact solution.

Many numerical methods have been proposed to improve this method. One approach is the split Bregman iteration [20], [21], [22], [23], which uses functional splitting and Bregman iteration for constrained optimization. The equivalence of the split Bregman iterations with the alternating direction method of multipliers (ADMM), the Douglas-Rachford splitting and the augmented Lagrangian method can be found in [23], [24], [25], [26].

The use of TV model is motivated by the analogy between the problem of TV decomposition and latent fingerprint segmentation. As discussed in Section II, the key challenge for latent segmentation is to effectively separate latent fingerprint with different structured noise. Structured noise (e.g. arch, character), with its smooth inner surface and crisp edges, share many similar characteristics with components in the cartoon layer u. On the other hand, fingerprint pattern, which consists of oscillatory ridge structures, matches the characteristics of the texture components in v. This interesting analogy suggests that TV model could be a viable solution to our problem.

B. Multiscale Feature Selection of TV Model

The TV-L1 model distinguishes itself from other TV-based models by its unique capability of intensity-independent multi-scale decomposition. It has been shown both theoretically [14] and experimentally [12] that the fidelity weight coefficient, λ, in (1) is closely related to the scale of features in the texture output v. This relation is supported by the analytic example in [14]. If f is equal to a disk signal, denoted by B_r, which has radius r and unit height, the solution of (1) is given as:

$$u_\lambda(x) = \begin{cases} 0 & \text{if } 0 \le \lambda < \dfrac{2}{r} \\ f(x) & \text{if } \lambda > \dfrac{2}{r} \\ cf(x) & \text{if } \lambda = \dfrac{2}{r}, \text{ for any } c \in [0, 1] \end{cases}$$

In other words, depending on the λ value, the TV-L1 functional is minimized by either 0 or input f. This shows that the TV-L1 model has the ability to select geometric features based on a given scale. Fig. 3 shows an example of feature selection on the latent fingerprint image.

As shown in Fig 3, the numerical results match with the analysis before. The fidelity weight coefficient λ controls the feature selection by manipulating the scale of content captured in each image layer. When λ is very small (*e.g.,* $\lambda = 0.10$), u captures the inhomogeneous illumination in the background while most fine structures are kept in v. When $\lambda = 0.30$, large-scale objects (arch) are captured in u, and separated from structures of smaller scales (characters). As λ continues to increase, only small-scale structures (fingerprint and noise) are left in v while the major content of f is extracted to u.

We observe that one of the differences between fingerprint patterns and structure noise is their relative scale. By applying

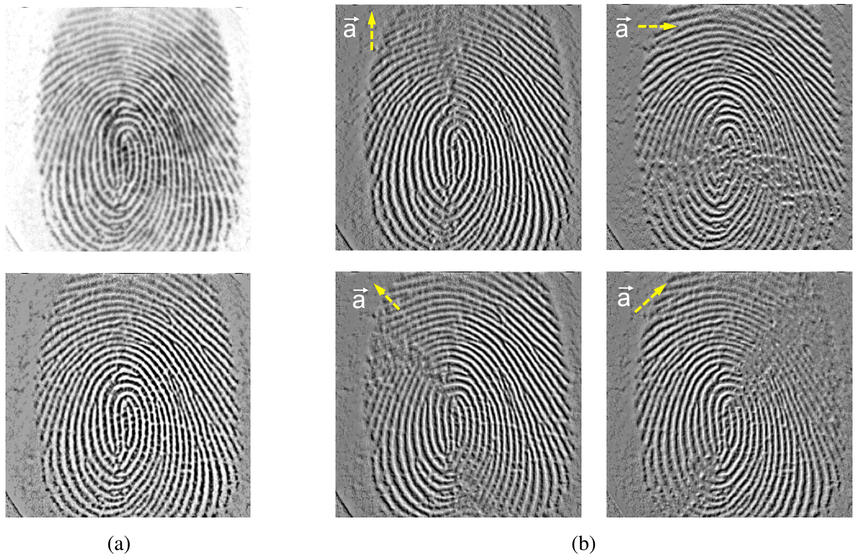


Fig. 4. (a) Top: original image f. Bottom: texture output v after decomposition by the TV-L2 model [1], (b) Texture output v for orientation vector $\vec{a}$ in four different directions. Top: $\vec{a} = (0, 1)$ and $\vec{a} = (1, 0)$. Bottom: $\vec{a} = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and $\vec{a} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$.

the TV-L1 model with an appropriately chosen λ value, it's seemingly possible to extract fingerprints to texture layer v while leaving the unwanted structure noise in cartoon layer u. However, there arise two problems by applying the TV-L1 model directly:

1) The value of λ forces structures that are smaller that or equal to a given scale to appear in v. As a result, structured noises of smaller scales as fingerprints (e.g. speckle, stain), will also be captured by v along with fingerprints.
2) A small amount of boundary signals near non-smooth edges will appear in v (see Fig. 5) due to the non-smoothness of the boundary and the use of finite differencing. This issue was also reported in [15].

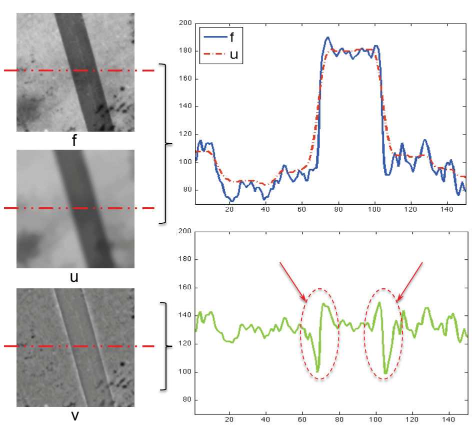


Fig. 5. Illustration of the boundary signal problem in TV-L1 decomposition: a small amount of structure noise edge signal is still kept in texture v (left) and signals along the dash line depicted in f, u and v (right).

To overcome these limitations, we propose the Adaptive Directional Total-Variation (ADTV) model, which will be presented next.

C. The ADTV Model

The TV-L1 model with spatially invariant fidelity (1) does not generate the desired output throughout the entire fingerprint image. In the fingerprint region, when λ is well matched with the scale of fingerprints, all essential contents can be captured in the texture layer v. However, in the noisy region, some unwanted signals will also be extracted to v under the same λ value. In addition, being an isotropic model, the TV model minimizes the total variation of cartoon layer u along all directions. This scheme does not fully exploit the information of orientation coherency, which is one of the most unique characteristics of fingerprints.

These concerns motivate us to consider a more flexible image model that is capable of integrating the special characteristics of fingerprints. We call this new model the Adaptive Directional Total-Variation (ADTV) model:

$$u^* = \operatorname*{argmin}_u \int |\nabla u \cdot \vec{a}(x)|\, dx + \frac{1}{2} \int \lambda(x)\, |u - f|\, dx \quad (3)$$

where $\vec{a}(x)$ is a spatially varying orientation vector adjusted to the local texture orientation, and $\lambda(x)$ is a spatially varying parameters that controls the feature scale.

The spatially varying parameter, $\lambda(x)$, can be understood in two ways. First, $\lambda(x)$ is a scalar that controls the scale of features appearing in v at pixel x. A large $\lambda(x)$ value enforces most textures to be kept in u, leaving only tiny-scale structures in v. When $\lambda(x)$ is sufficiently large, $u^*(x) \approx f(x)$, and the original content is almost totally blocked from v. Thus, $v(x) = f(x) - u^*(x) \approx 0$. Second, parameter $\lambda(x)$ can also be interpreted as a weighting coefficient that balances the importance between fidelity and smoothness of u. In the fingerprint region, the $\lambda(x)$ value should be relatively small, since low fidelity ensures the smoothness of u and, thus, more textures could be extracted to v. In regions with structured noise, fidelity becomes important and large $\lambda(x)$ ensures all noise components to be filtered out from texture v.

The orientation vector $\vec{a}(x)$ also controls the content captured in the texture layer v, but in a different manner. By tuning $\vec{a}(x)$ to a specific direction, we are mainly interested in minimizing the total variation of u along that direction, while allowing the variation of u to exist along other directions. As a result, textures along the corresponding direction will be fully captured by v while textures of other directions will be weakened in v. In particular, textures along the orthogonal direction of $\vec{a}(x)$ will be totally blocked from v. In Fig. 4, we illustrate the impact of $\vec{a}(x)$ on the texture output v by applying our ADTV model with fixed orientation vector $\vec{a}$ on a typical fingerprint image. In the proposed ADTV model, $\vec{a}(x)$ is a spatially varying vector that is adaptively chosen according to the image content.

Algorithm 1. Augmented Lagrangian method for our proposed ADTV model.

1) *Initialization:* $u^0 = 0$, $\vec{p}^0 = 0$, $q^0 = 0$, $w^0 = 0$;
2) For $k = 0, 1, 2, ...$, compute:

$$(u^{k+1}, \vec{p}^{k+1}, q^{k+1}, w^{k+1}) = \underset{(u,\vec{p},q,w)}{\operatorname{argmin}}\ \mathfrak{L}(u, \vec{p}, q, w, \vec{\mu_p}^k, \mu_q^k, \mu_w^k) \tag{4}$$

3) *Update:*

$$\begin{aligned} \vec{\mu_p}^{k+1} &= \vec{\mu_p}^k + r_p(\vec{p}^{k+1} - \nabla u^{k+1}) \\ \mu_q^{k+1} &= \mu_q^k + r_q(q^{k+1} - \vec{p}^{k+1} \cdot \vec{a}) \\ \mu_w^{k+1} &= \mu_w^k + r_w(w^{k+1} - u^{k+1}) \end{aligned}$$

We use the augmented Lagrangian method [27], [28], [29], [30] to solve the proposed ADTV model given in (3). The augmented Lagrangian method is both accurate and efficient, as it benefits from the FFT-based fast solver with a closed-form solution. It has been proven that the augmented Lagrangian is equivalent to the split Bregman iteration and its convergence is always guaranteed [23].

In the augmented Lagrangian method, three new variables ($\vec{p}$, q, w) are introduced to reformulate (3) into the following constraint optimization problem:

$$\begin{aligned} &\min_u \quad \int |q| + \frac{1}{2}\int \lambda(x)\,|w - f|\,dx, \\ &\text{s.t.} \quad \vec{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix} = \nabla u, q = \vec{p}\cdot\vec{a}, w = u \end{aligned} \tag{5}$$

To solve (5), the following augmented Lagrangian functional is defined:

$$\begin{aligned} \mathfrak{L}(u, \vec{p}, q, w, \vec{\mu_p}, \mu_q, \mu_w) = &\int |q| + \frac{1}{2}\int \lambda(x)\,|w - f|\,dx \\ &+ \frac{r_p}{2}\int (\vec{p} - \nabla u)^2 + \int \vec{\mu_p}(\vec{p} - \nabla u) \\ &+ \frac{r_q}{2}\int (q - \vec{p}\cdot\vec{a})^2 + \int \mu_q(q - \vec{p}\cdot\vec{a}) \\ &+ \frac{r_w}{2}\int (w - u)^2 + \int \mu_w(w - u), \end{aligned}$$

where $\vec{\mu_p}$, μ_q and μ_w are the Lagrange multipliers and r_p, r_q, r_w are positive constants. The augmented Lagrangian method uses an iterative procedure to solve (5) as shown in Algorithm 1. The iterative scheme runs until some stopping condition is satisfied. Since variables $u, \vec{p}, q, w$ in $\mathfrak{L}(u, \vec{p}, q, w, \vec{\mu_p}, \mu_q, \mu_w)$ are coupled together, it is difficult to solve them simultaneously. Instead, the problem is decomposed into four sub-problems and an alternative minimization process is applied. Instead of solving 4 exactly, we apply the alternating direction method of multipliers (ADMM) [23], [24] and run one iteration for each sub-problem. It should be mentioned that this was also reused in the split Bregman iteration method [25], [26]. This approach of splitting technique is efficient since all sub-problems have closed-form solution, which are given as:

$$\begin{aligned} u^k &= \mathcal{F}^{-1}\left(\frac{\begin{array}{c} -r_p \cdot \mathcal{F}(\mathrm{div}\cdot\vec{p}) - \mathcal{F}(\mathrm{div}\cdot\vec{\mu}_p^k) \\ +r_w\mathcal{F}(w) + \mathcal{F}(\mu_w^k) \end{array}}{r_w - r_p\mathcal{F}(\triangle)} \right), \\ \vec{p}^k(x) &= \nabla u - \frac{1}{r_p}\left(\vec{\mu_p}^k - (r_q q - r_q\rho(x) + \mu_q^k)\cdot\vec{a}(x)\right), \\ q^k(x) &= \mathbf{max}\left\{0, 1 - \frac{1}{r_q\,|\psi(x)|}\right\}\cdot\psi(x), \\ w^k(x) &= \mathbf{max}\left\{0, 1 - \frac{\lambda(x)}{r_w\,|\phi(x)|}\right\}\cdot\phi(x) + f(x) \end{aligned}$$

where $\rho(x) = \frac{r_p(\nabla u\cdot\vec{a}(x)) - \vec{\mu_p}^k\cdot\vec{a}(x) + (\mu_q^k + r_q q)\|\vec{a}(x)\|^2}{r_p + r_q\|\vec{a}(x)\|^2}$, $\psi(x) = \vec{p}\cdot\vec{a}(x) - \frac{\mu_q^k(x)}{r_q}$, and $\phi(x) = u(x) - f(x) - \frac{\mu_w^k(x)}{r_w}$. $\mathcal{F}(u)$ and $\mathcal{F}^{-1}(u)$ denotes the Fourier transform and inverse Fourier transform of u, respectively.

D. Orientation Field Estimation

In order to extract the fingerprint components to texture output v, $\vec{a}(x)$ should be spatially varying and well aligned with the local fingerprint ridge orientation. We use the gradient-based approach [31], [32] for computing the coarse orientation field at each pixel:

$$o(x) = \frac{1}{2}\tan^{-1}\frac{\sum_W 2f_{x_1}f_{x_2}}{\sum_W (f_{x_1}^2 - f_{x_2}^2)} + \frac{\pi}{2}$$

where W is a neighborhood window around x, (f_{x_1}, f_{x_2}) is the gradient vector at $x = (x_1, x_2)$, and $\tan^{-1}$ is a 4-quadrant arctangent function with output range of $(-\pi, \pi)$.

The estimation above is relatively accurate at fingerprint regions, while it becomes less reliable at noisy regions. We evaluate the reliability of the estimated orientation field by its local coherency:

$$c(x) = \frac{(\sum_W (f_{x_1}^2 - f_{x_2}^2))^2 + 4(\sum_W f_{x_1}f_{x_2})^2}{(\sum_W (f_{x_1}^2 + f_{x_2}^2))^2}$$

where $c(x) \in [0, 1]$ (close to 1 for strongly oriented pattern, and 0 for isotropic regions). The value of $c(x)$ provides a reliability measure of the estimated orientation field and will be utilized to generate the final orientation vector $\vec{a}(x)$.

The coarse orientation field $o(x)$ still contains inconsistencies caused by creases and ridge breaks of the fingerprint pattern. We further improve the estimation by orientation

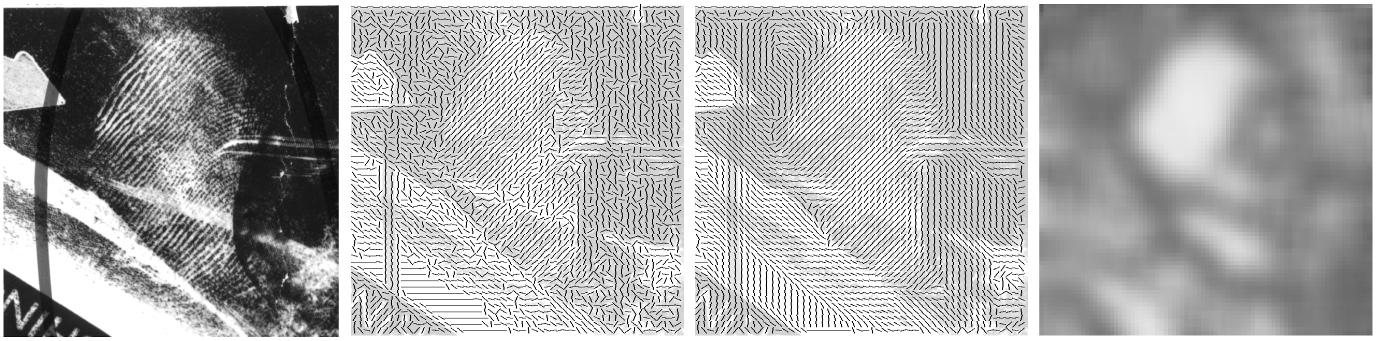


Fig. 6. Essential steps for computing $\vec{a}(x)$. From left to right: original image f, coarse orientation estimation $o(x)$, orientation smoothening $O(x)$ and coherency evaluation $c(x)$.

smoothening:

$$O(x) = \frac{1}{2}\tan^{-1}\left\{\frac{G_\sigma * \sin(2\cdot o(x))}{G_\sigma * \cos(2\cdot o(x))}\right\}$$

where G_σ is a Gaussian smoothing kernel with standard deviation σ.

Finally, the orientation vector $\vec{a}(x)$ in (3) is computed as:

$$\vec{a}(x) = (-\cos O(x), \sin O(x)) \cdot c(x)$$

At regions where the orientation estimation is reliable, the large $c(x)$ value enforces textures along the direction $\vec{a}(x)$ to be fully captured by v, leaving textures of the remaining orientations in u. On the other hand, at regions where $c(x)$ is small and the estimation is not trustworthy, the fidelity term $\frac{1}{2}\int \lambda(x)\,|u-f|\,dx$ becomes dominant and most of the image content will be kept in u. In this way, we can efficiently filter out the structured noises from the texture output v. The process for computing the parameter $\vec{a}(x)$ is illustrated in Fig. 6.

E. Scale Parameter Selection

As discussed in Section III-B, applying one uniform λ value over the entire fingerprint image does not generate satisfactory results. To improve the result, the value of λ should be spatially adaptive. That is, $\lambda(x)$ ought to be adaptively chosen according to the background noise level. Ideally, parameter $\lambda(x)$ should take larger values in regions with much structured noise and be relatively small in fingerprint regions.

To differentiate these regions, we study their characteristics after going through local low-pass filtering. When an input image, f, is locally filtered by a low-pass filter denoted by

$$L_\sigma(\xi) = \frac{1}{1+(2\pi\sigma\,|\xi|)^4},$$

its cartoon and texture components, though both being blurred to some extent, change differently by means of local total variation (LTV), which is defined as

$$LTV(f) = G_\sigma * |\nabla f|,$$

where f is the image region and G_σ is a Gaussian kernel with standard deviation σ. In [33], the author used the relative LTV reduction ratio to differentiate cartoon with textural regions. It was observed that the LTV of textural regions decay much rapidly than that of cartoon regions after low-pass filtering.

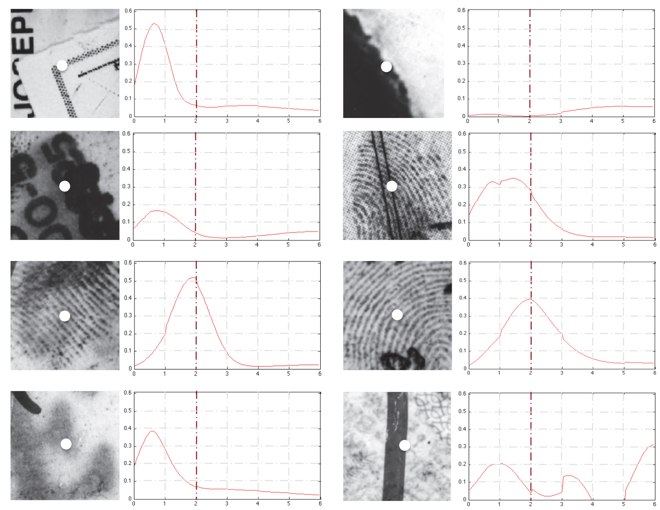


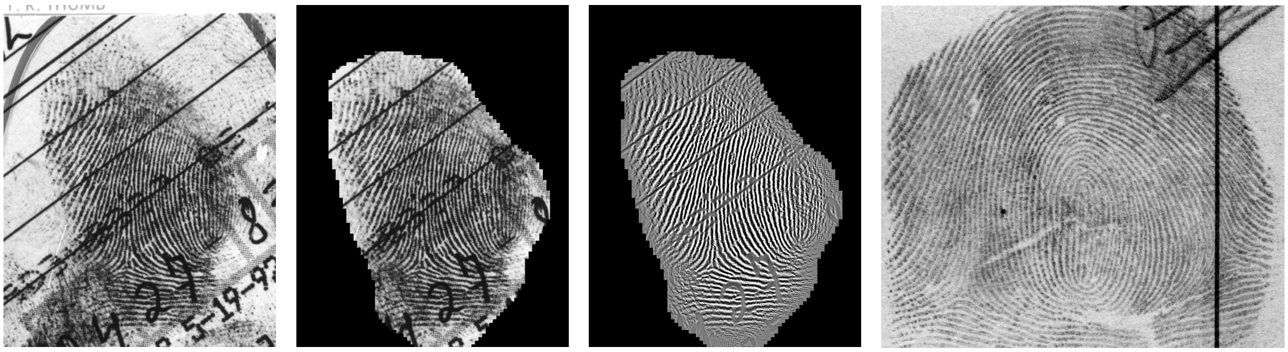
Fig. 7. Plots of $\eta_\sigma(x)$ for several pixels in different latent fingerprint images. It has a sharp peak located near $\sigma = 2.0$ in the fingerprint region while it reaches the maximum at different σ values in other regions.

Though the LTV reduction ratio provides a good measure for separating edgy regions from textural regions, it has limited capability in differentiating textures of different scales (*e.g.,* fingerprints with speckles). To overcome this limitation, we further introduce the differential LTV reduction rate, denoted by η_σ, as

$$\eta_\sigma = \frac{LTV(L_{\sigma+1} * f) - LTV(L_\sigma * f)}{LTV(f)}. \quad (6)$$

For a given local patch, the parameter η_σ describes its structural components' sensitivity to low-pass filtering of scale σ. It provides useful information about the underlying texture structure of a local region. Intuitively, it measures the texture's local oscillatory behavior at a certain spatial scale σ. In Fig. 7, we demonstrate the η_σ values of different textural patches, which are extracted from latent fingerprint images. We observe that the η_σ value of fingerprint regions all reaches local maxima around $\sigma = 2.0$. With a fixed σ value, η_σ will have the largest response for textural components of scales around σ, while the response for textures of other scales will be suppressed.

Based on this observation, we choose the spatially variant coefficient $\lambda(x)$ in (3) as:



(a) Latent print input: Type-1 (b) Latent print input: Type-2 (c) Latent print input: Type-3 (d) Mated rolled fingerprint

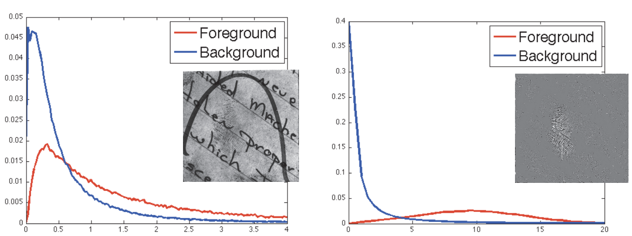
Fig. 8. Input images for latent matching. (a) Type-1: without any segmentation, (b) Type-2: segmentation mask over original image f, (c) Type-3: segmentation mask over texture layer v, (d) the corresponding mated rolled fingerprint.

$$\lambda(x) = \kappa \cdot \frac{1}{\eta_c(x) + \epsilon}, \tag{7}$$

where η_c is the differential LTV reduction rate at $\sigma = c$, which is adjusted to the best response of fingerprint patterns, κ and ϵ are trivial positive constants used for scaling and avoiding zero-division. In our experiments, we observe that $c = 2.0$ gives the optimum value for the latent fingerprint patterns while parameters κ and ϵ are empirically set to 0.5 and 0.01, and used for scaling and avoiding zero-division, respectively.

F. Region-of-interest Segmentation and Enhancement

After decomposing the latent fingerprint image using our proposed ADTV model, we have obtained two image layers: 1) cartoon u, which contains the majority of unwanted content (e.g. structured noise, small-scale structures), and 2) texture v, which consists of latent fingerprints and only a small amount of random noise. This decomposition facilitates two procedures: segmentation and enhancement.



(a) Input f (b) Texture v

Fig. 9. Distributions of the variance feature for the foreground and background region in f and v, respectively.

The variance value acts as a key segmentation feature for rolled/plain fingerprints [6]. As discussed in Section II, this feature cannot be directly applied to latent fingerprints due to the presence of structured noise. However, after the cartoon-texture layer decomposition, most high-variance noise components are kept away from the texture layer v, allowing us to use the variance features for segmentation. We verify this point in Fig. 9, where we plot the probability distribution of variance feature at foreground/background regions before and after the ADTV-based decomposition.

In addition, our proposed decomposition scheme is capable of enhancing the fingerprint quality. After decomposition, in the texture layer v, we have removed all the unwanted components that may overlap with the fingerprints. The extracted patterns are less degraded by structured noises and free of illumination effects, leading to enhanced fingerprint quality. In the next section, we will experimentally demonstrate that this enhancement as well as the segmentation will result in better latent matching performance.

IV. Experimental Results

In this section, we evaluate the proposed ADTV model by several experiments. We first show the segmentation results for latent fingerprint images with different quality types. The decomposed texture layer v using the ADTV model is compared with two other TV models: the TV-L1 [14] and TV-L2 [1] model. Then we experimentally examine the impact of our proposed segmentation scheme on the accuracy of feature extractions. Finally, we conduct latent matching experiments to verify whether the segmentation result can indeed lead to higher matching accuracy.

A. Data Preparation

The experiments are conducted on the public domain fingerprint database, NIST SD27, which contains 258 latent fingerprints and their corresponding rolled fingerprints. In this database, fingerprint experts have assigned to each fingerprint one of three quality levels - *good*, *bad* and *ugly*. The numbers of "good", "bad" and "ugly" latent prints are 88, 85 and 85, respectively. In our experiment, we have selected 29, 27 and 27 prints from each category, and experimented on this subset of data. For the matching experiment, we included 27,000 rolled fingerprints from the NIST SD14 database and extended the background database to 27,258 fingerprints, making the problem more realistic and challenging.

The feature extraction and fingerprint matching process are conducted using the commercial matcher Neurotechnology VeriFinger SDK 6.5 [34]. For each latent fingerprint, we compare the results for three inputs:

- *Type-1*: original latent fingerprint without segmentation,
- *Type-2*: segmentation mask applied on original image,
- *Type-3*: segmentation mask applied on texture layer v.

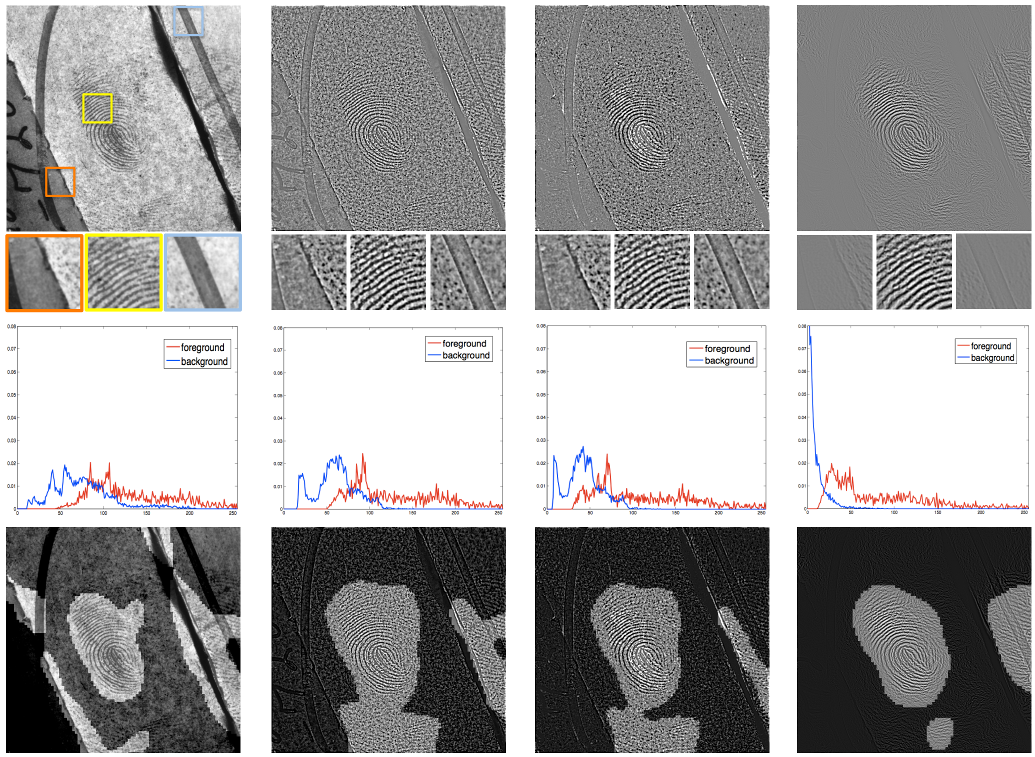


Fig. 10. Performance comparison of the proposed ADTV model and two other TV-based models. First row: original image f, texture output v of TV-L2 [1], TVL1 [14] and the proposed ADTV model. Second row: distribution of variance feature in the foreground and background areas. Third row: the segmentation result based on variance feature.

As mentioned previously, our proposed ADTV model has two functionalities: segmentation and texture enhancement. We evaluate the effectiveness of segmentation by comparing the corresponding results of Type-1 and Type-2, while comparing results of Type-2 and Type-3 provides evaluation for the impact of texture enhancement.

B. Comparison with other TV-based models

In Fig. 13-15, we present the segmentation results using our proposed ADTV scheme for latent prints of good, bad and ugly quality types, respectively. Visual inspection shows that the proposed ADTV segmentation scheme provides satisfactory results.

In Fig. 10, we compare the texture output v of our proposed ADTV model with two classical TV models: TV-L1 [14] and TV-L2 [1]. The variance distribution of foreground/background regions as well as the corresponding segmentation result are shown in the bottom part of Fig. 10, respectively. In fingerprint regions, our proposed ADTV model is able to extract all essential fingerprint texture with clear ridge information, while the results of other TV model still contains some background noise. In regions with structured noise, boundary and speckle noise can be clearly observed in the texture layer v obtained by other TV-based model. In contrast, our proposed ADTV model is capable of filtering out these background noise signals from the texture output v.

C. Feature Extraction Accuracy

Without segmentation, the performance of latent matching is very poor due to the high number of unreliable features. There are two types of features that are essential for fingerprint matching: singular points (core, delta) and minutiae. Traditional feature extraction algorithms perform poorly on latent fingerprints, especially at regions with much structured noises. Some areas of the noise are often miss-identified as useful fingerprint features, which could heavily affect the accuracy of the fingerprint matching stage. Therefore, with the help of accurate segmentation, we can remove the unwanted structured noise components and thereby decrease the number of erroneous features.

We manually segment each latent fingerprint image, and use VeriFinger SDK 6.5 for feature extraction. All extracted features points that fall within the manual segmentation region are used as the ground truth. Then we calculate the number of true features points that were missing as well as the number of false feature points detected, for two scenarios: with and without segmentation. Experimental results are given in Table I. For latent inputs without any segmentation, though none of the true features points are missing, 30 - 40% of the detected features are erroneous. On the other hand, for inputs after segmentation, the false feature point ratio has decreased to less than 5%, while the missing features points have only slightly increased by 5 - 10%. In the next session, we will show that the improvement in feature extractions will lead to better matching performance.

D. Fingerprint Matching Results

The ultimate goal of segmentation is to successfully match the input latent fingerprint with the corresponding plain/rolled fingerprint in a large database. In the preceding sections, we

TABLE I
FEATURE EXTRACTION ACCURACY WITH AND WITHOUT SEGMENTATION OF ADTV MODEL

	Without Segmentation	Segmentation with ADTV model
Missed Minutiae	0.0%	9.5%
Missed Singular Points	0.0%	4.2%
False Minutiae	34.2%	5.3%
False Singular Points	45.3%	0.0%

have visually demonstrated the effectiveness of the segmentation scheme using our proposed ADTV model, and shown that the segmentation does improve the accuracy of feature extractions. In this section, we conduct matching experiment to verify whether the segmentation result can indeed lead to improved matching accuracy.

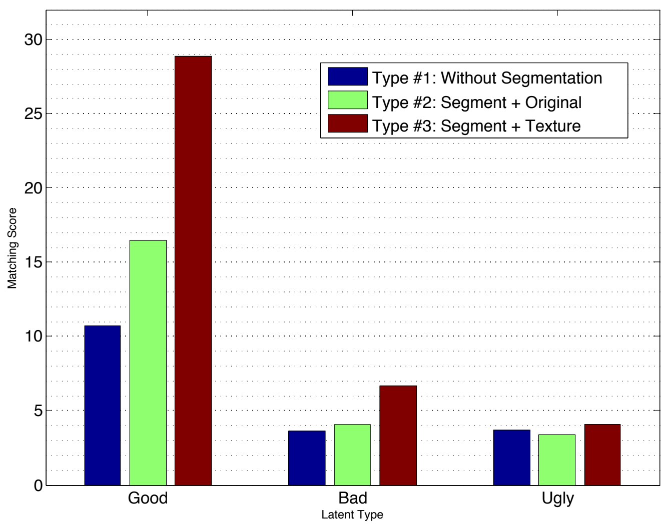


Fig. 11. Average latent matching score for latent inputs of all three quality types (good, bad, ugly).

We first use VeriFinger SDK 6.5 to match all three input types with their corresponding mated rolled fingerprint. The average matching scores are computed for three quality types separately (see Fig. 11). For input latents of good quality, the average matching scores increase significantly after segmentation, and further improvement can be observed after texture extraction using our ADTV model, which acts similar to fingerprint enhancement. However, for latent prints of bad and ugly qualities, no significant improvements can be observed. For these input images, the fingerprint patterns are so degraded that even with reliable segmentation and enhancement, latent matching still remains to be extremely challenging.

The Cumulative Match Characteristic (CMC) curves of the three input types to the fingerprint matcher are shown in Fig. 12. Each input is searched against the background database of 27,258 rolled fingerprint. The CMC curve plots the rank-k identification rate against k, for $k = 1, 2, 3, ..., 100$. The rank-k identification rate indicates the proportion of times the mated rolled fingerprint appears in the top k matches. Due to the poor matching scores of bad and ugly latent prints, we only include the matching results of latent prints with good quality. As shown in Fig. 12, automatic matching performance is significantly improved when Type-2 and Type-3 are used as input to the matcher. The rank-1 identification ratio has increased from 9.8% (Type-1) to 18.9% (Type-2) and 30.2% (Type-3).

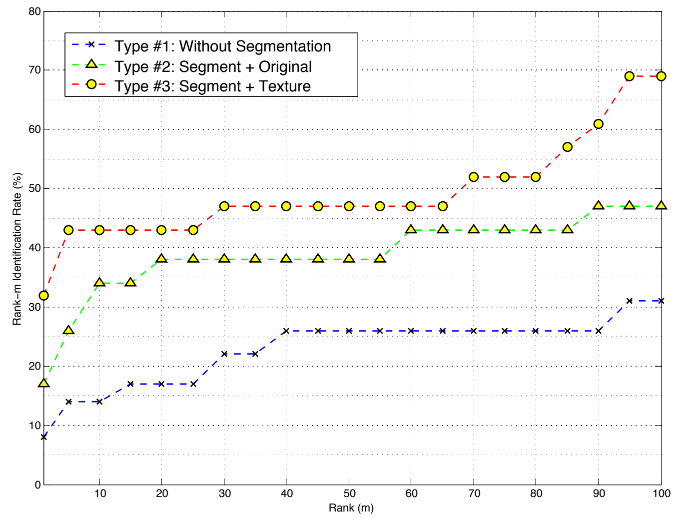


Fig. 12. Cumulative matching curve (CMC) of all three latent input types.

V. CONCLUSIONS AND FUTURE WORK

While current automated fingerprint identification systems have achieved high accuracy in matching rolled/plain prints, latent matching still remains to be a challenging problem and requires much human intervention. The goal of this work is to achieve accurate latent segmentation, which is an essential step towards achieving automatic latent identification. Existing fingerprint segmentation algorithms performs poorly on latent prints, as they are mostly based on the assumptions that are only applicable for rolled/plain fingerprints.

In this paper, we have proposed the Adaptive Directional Total Variation (ADTV) model as an image decomposition scheme that facilitates effective latent fingerprint segmentation and enhancement. Based on the classical Total-Variation model, the proposed ADTV model differentiates itself by integrating two unique features of fingerprints, scale and orientation, into the model formulation. The proposed model has the ability to decompose a single latent image into two layers and locate the essential latent area for feature matching. The two spatially varying parameters of the model, scale and orientation, are adaptively chosen according to the background noise level and textural orientation, and effectively separate the latent fingerprint from structured noises in the background. Experimental results show that the proposed scheme provides

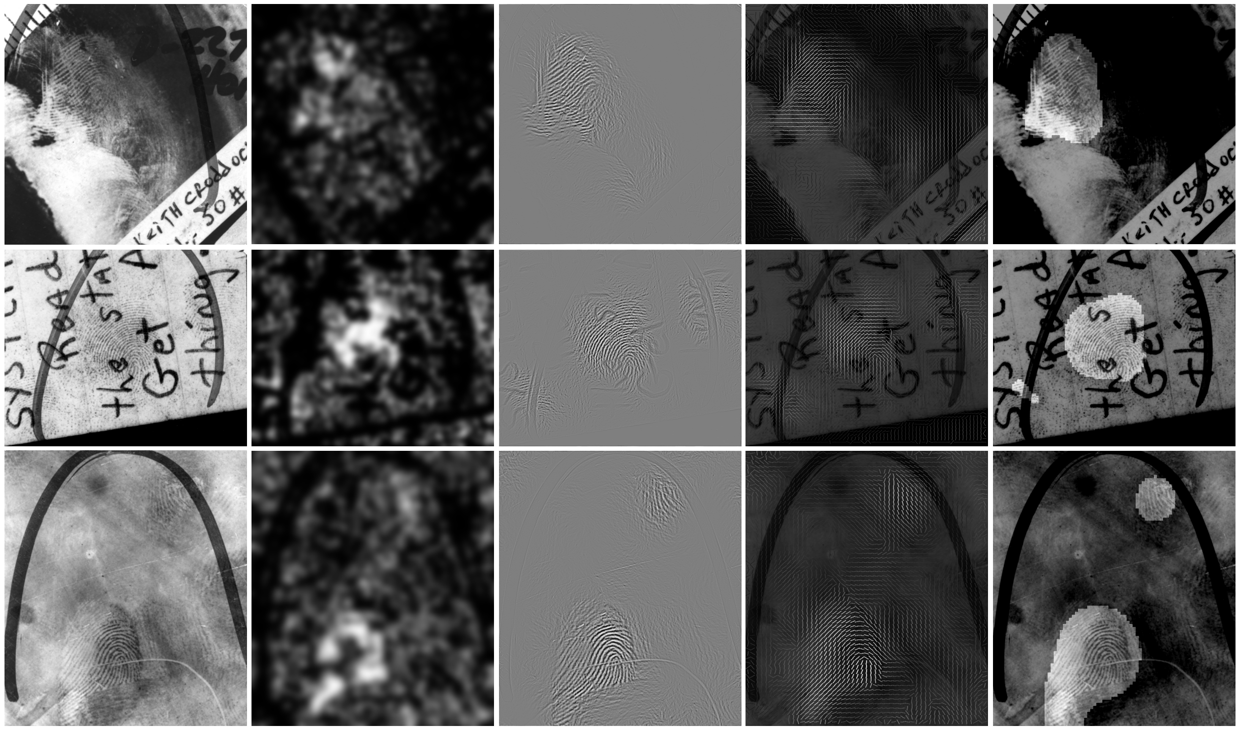


Fig. 13. Experimental results of latent fingerprints with *good* quality. From left to right: original image f, scale parameter $\lambda(x)$, orientation vector $\vec{a}(x)$, texture output v and the final segmentation result.

effective segmentation and enhancement. The improvements in feature detection accuracy and latent matching further justifies the effectiveness of the proposed scheme.

The proposed scheme can be regarded as a preprocessing technique for automatic latent fingerprint recognition. It also has strong potential to be applied on other relevant applications, especially for processing images with oriented textures. This study can be further extended along the following directions:

1) The effectiveness of the proposed scheme is related to the accuracy of orientation estimation. When the estimated orientation is unreliable, fingerprint patterns may not be fully extracted to the texture layer v, leading to poor segmentation and enhancement results. In addition, the positions of singular points were not taken into consideration by our proposed model. Though only very few singular points appear in each latent image, additional detection and processing techniques need to be introduced for handling regions surrounding the singular points.
2) Experimental results on latent matching shows that our proposed scheme shows significant improvement only for latent images of good quality, while not much improvement could be observed for bad and ugly latent prints. These challenging images still require manual processing from fingerprint experts.
3) For a latent image of size 768×768 pixels, it would take about one minute for the proposed model to converge. Although this computation complexity is acceptable for the processing of latent fingerprints, speed-up solutions are desirable to enable more efficient processing.

Acknowledgment

The authors would like to thank Dr. Xian Tang, Dr. Lester Li and other researchers at 3M Cogent Inc. for their support in data collection and useful discussion.

References

[1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, November 1992.

[2] A. K. Jain and J. Feng, "Latent fingerprint matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, pp. 88–100, January 2011.

[3] B. M. Mehtre, N. N. Murthy, S. Kapoor, and B. Chatterjee, "Segmentation of fingerprint images using the directional image." *Pattern Recognition*, vol. 20, no. 4, pp. 429–435, 1987.

[4] B. Mehtre, "Segmentation of fingerprint images - a composite method," *Pattern Recognition*, vol. 22, no. 4, pp. 381–385, 1989.

[5] N. K. Ratha, S. Chen, and A. K. Jain, "Adaptive flow orientation-based feature extraction in fingerprint images." *Pattern Recognition*, vol. 28, no. 11, pp. 1657–1672, 1995.

[6] A. M. Bazen and S. H. Gerez, "Segmentation of fingerprint images," in *ProRISC 2001 Workshop on Circuits, Systems and Signal Processing*, 2001, pp. 276–280.

[7] J. Feng, S. Yoon, and A. K. Jain, "Latent fingerprint matching: Fusion of rolled and plain fingerprints," in *Proceedings of the Third International Conference on Advances in Biometrics*, ser. ICB '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 695–704.

[8] F. Chen, J. Feng, A. K. Jain, J. Zhou, and J. Zhang, "Separating overlapped fingerprints." *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 346–359, 2011.

[9] S. Yoon, J. Feng, and A. K. Jain, "On latent fingerprint enhancement," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 7667, Apr. 2010.

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

[11] N. Short, M. Hsiao, A. Abbott, and E. Fox, "Latent fingerprint segmentation using ridge template correlation," in *Proceedings of the International Conference on Imaging for Crime Detection and Prevention*. London, UK: IEEE, 2011.

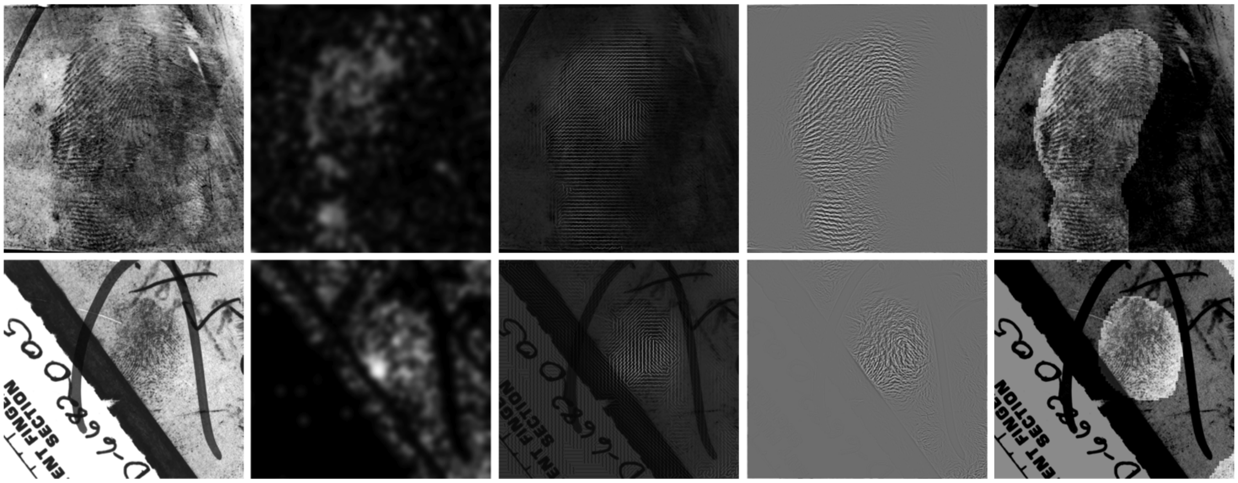


Fig. 14. Experimental results of latent fingerprints with *bad* quality. From left to right: original image f, scale parameter $\lambda(x)$, orientation vector $\vec{a}(x)$, texture output v and the final segmentation result.

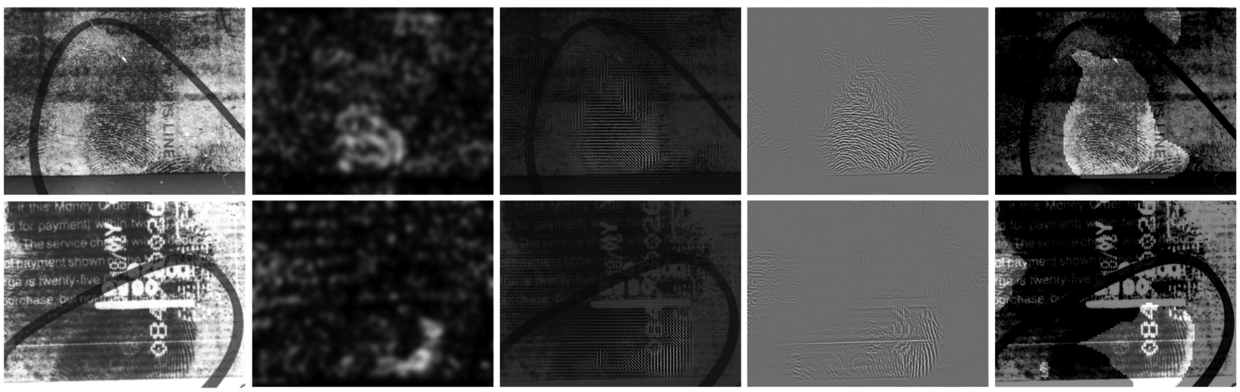


Fig. 15. Experimental results of latent fingerprints with *ugly* quality. From left to right: original image f, scale parameter $\lambda(x)$, orientation vector $\vec{a}(x)$, texture output v and the final segmentation result.

[12] W. Yin, D. Goldfarb, and S. Osher, "The total variation regularized l1 model for multiscale decomposition," *Multis. Model. Simul.*, vol. 6, 2006.

[13] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition–modeling, algorithms, and parameter selection," *Int. J. Comput. Vision*, vol. 67, no. 1, pp. 111–136, Apr. 2006.

[14] T. Chan and S. Esedoglu, "Aspects of total variation regularized l1 function approximation," *SIAM Journal on Applied Mathematics*, vol. 65, 2004.

[15] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, "Total variation models for variable lighting face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1519–1524, 2006.

[16] J. Zhang, R. Lai, and C.-C. J. Kuo, "Latent fingerprint segmentation with adaptive total variation model," in *Proceedings of the International Conference on Biometrics*, 2012 (to appear).

[17] ——, "Latent fingerprint detection and segmentation with a directional total variation model," in *Proceedings of the International Conference on Image Processing (ICIP)*, 2012 (to appear).

[18] H. Fleyeh, D. Jomaa, and M. Dougherty, "Segmentation of low quality fingerprint images," *International Conference on Multimedia Computing and Information Technology (MCIT-2010)*, vol. March 2-4, 2010.

[19] Z. Shi, Y. Wang, J. Qi, and K. Xu, "A new segmentation algorithm for low quality fingerprint image," in *Proceedings of the Third International Conference on Image and Graphics*, ser. ICIG '04. Washington, DC, USA: IEEE Computer Society, 2004, pp. 314–317.

[20] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split bregman method: Segmentation and surface reconstruction," *Journal of Scientific Computing*, vol. 45, no. 1-3, 2009.

[21] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM J. Img. Sci.*, vol. 2, pp. 323–343, April 2009.

[22] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Mode. Simul.*, vol. 4, 2005.

[23] C. Wu and X. Tai, "Augmented lagrangian method, dual methods and split bregman iterations for rof model, vectorial tv and higher order models," *SIAM J. Imaging Science*, vol. 3, 2010.

[24] E. Esser, "Applications of lagrangian-based alternating direction methods and connections to split bregman," *UCLA CAM Report (09-31)*, 2009.

[25] S. Setzer, "Split bregman algorithm, douglas-rachford splitting and frame shrinkage," in *Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision*, ser. SSVM '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 464–476.

[26] J. Cai, S. Osher, and Z. Shen, "Split bregman methods and frame based image restoration," *Multiscale Model. Simul.*, vol. 8, 2009.

[27] R. Glowinski and P. L. Tallee, "Augmented lagrangian and operator-splitting methods in nonlinear mechanics," *SIAM*, 1989.

[28] M. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, 1969.

[29] X. Tai, J. Hahn, and G. Chung, "A fast algorithm for euler's elastic model using augmented lagrangian method," *SIAM J. Imaging Science*, vol. 4, 2011.

[30] C. Wu, J. Zhang, and X.-C. Tai, "Augmented lagrangian method for total variation restoration with non-quadratic fidelity," *Inverse Problems and Imaging (IPI*, vol. 4, 2011.

[31] L. Hong, Y. Wan, and A. Jain, "Fingerprint image enhancement: Algorithm and performance evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 777–789, 1998.

[32] 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.

[33] A. Buades, T. M. Le, J.-M. Morel, and L. A. Vese, "Fast cartoon + texture image filters," *Trans. Img. Proc.*, vol. 19, pp. 1978–1986, August 2010.

[34] Neurotechnology Inc., VeriFinger, http://www.neurotechnology.com, 2012.