Information Fusion 34 (2017) 1–15

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Adaptive latent fingerprint segmentation using feature selection and random decision forest classification

Anush Sankaran, Aayush Jain, Tarun Vashisth, Mayank Vatsa*, Richa Singh

IIIT-Delhi, Okhla Phase-III, New Delhi, 11002, India

ARTICLE INFO

Article history:
Received 27 November 2015
Revised 12 April 2016
Accepted 16 May 2016
Available online 25 May 2016

Keywords:
Saliency
Random decision forest
Feature selection
Latent fingerprint segmentation

ABSTRACT

Latent fingerprints are important evidences used by law enforcement agencies. However, current state-of-the-art for automatic latent fingerprint recognition is not as reliable as live-scan fingerprints and advancements are required in every step of the recognition pipeline. This research focuses on automatically segmenting latent fingerprints to distinguish between ridge and non-ridge patterns. There are three major contributions of this research: (i) a machine learning algorithm for combining five different categories of features for automatic latent fingerprint segmentation, (ii) a feature selection technique using modified RELIEF formulation for analyzing the influence of multiple category features on latent fingerprint segmentation, and (iii) a novel SIVV based metric to measure the effect of the segmentation algorithm without the requirement to perform the entire matching process. The image is tessellated into local patches and saliency based features along with image, gradient, ridge, and quality based features are extracted. Feature selection is performed to study the contribution of the various category features towards foreground ridge pattern representation. Using these selected features, a trained Random Decision Forest based algorithm classifies the local patches as background or foreground. The results on three publicly available databases demonstrate the efficacy of the proposed algorithm.

1. Introduction

Latent fingerprints are (un)intentional deposition of ridge patterns on a surface that comes in contact with a fingerprint. In law enforcement applications, latent fingerprints are used as a crucial forensic evidence for crime scene analysis. Existing semi-automated procedures for latent fingerprint matching consist of: (1) preprocessing, (2) manual annotation of features such as minutiae and singular points, (3) search for top-k probable matches (typically $k = 50$) using an Automated Fingerprint Identification System (AFIS), and (4) manual verification of the candidate list by forensic experts. The preprocessing stage consists of (i) quality check of latent fingerprints, (ii) segmenting the foreground fingerprint regions from the noisy background, and (iii) enhancing the clarity of ridge patterns. In current approaches, forensic experts have to manually pre-process a large number of fingerprints which is time consuming and not scalable. Therefore, automating these steps can significantly increase the efficiency of AFIS. With the same intent, FBI's Next Generation Identification (NGI) [1] aims at building an automated "lights-out" latent fingerprint matching system.

This research paper focuses on automating the task of latent fingerprint segmentation. As shown in Fig. 1, some of the factors involved in making latent fingerprint segmentation a difficult problem are that fingerprints may be of poor ridge clarity [3] or certain latent fingerprints may not have a clear boundary due to smudges and background noise. A recent survey by Sankaran et al. [4] discusses the challenges involved in automatic latent fingerprint segmentation.

- **Poor ridge clarity:** The varying background surfaces, smudges, and non-linear distortion introduced while lifting, as shown in Fig. 1(a), reduces the ridge information in the print.
- **Structured marks:** Structured markings such as arch, lines, and characters, as shown in Fig. 2(a), have patterns that are similar to ridges, thereby adding to the challenges faced by an automatic latent fingerprint segmentation algorithm.
- **Optimal representation:** The output of fingerprint segmentation can be represented in multiple ways, as shown in Fig. 3. Selecting the optimal representation with maximum information is an essential research challenge.
- **Performance metric:** The performance of the preprocessing stage is usually evaluated as an improvement in rank-k

---

* Corresponding author.

E-mail addresses: anushs@iiitd.ac.in (A. Sankaran), aayush09055@iiitd.ac.in (A. Jain), tarun09051@iiitd.ac.in (T. Vashisth), mayank@iiitd.ac.in (M. Vatsa), rsingh@iiitd.ac.in (R. Singh).

http://dx.doi.org/10.1016/j.inffus.2016.05.002
1566-2535/

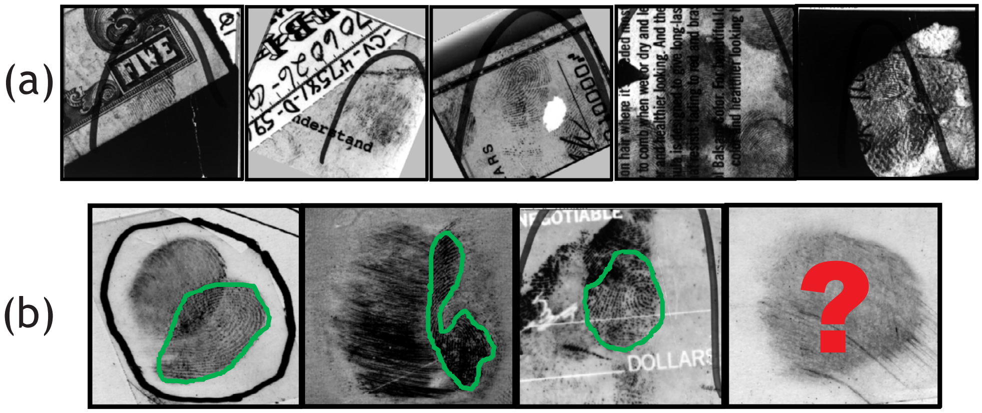
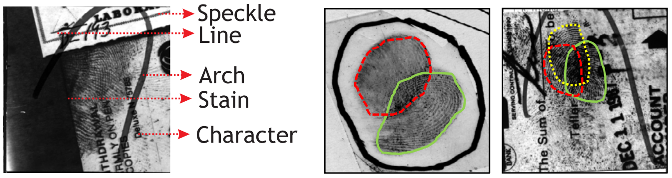


Fig. 1. (a) Sample latent fingerprint images from the NIST SD-27 database [2] demonstrating the effect of background information on ridge information and (b) latent fingerprint samples illustrating the problem of segmentation.



(a) Structured noise (b) Overlapped latent fingerprints

Fig. 2. Sample fingerprint images from NIST SD-27 showcasing two specific challenges in latent fingerprint segmentation. (a) presence of structured noise in latent fingerprint background that often resembles ridge like patterns, and (b) overlapped fingerprints result in overlapped ridge information making it difficult to determine the ridge flow of either of the fingerprints.

matching performance [5], [6]. However, as the overall latent print matching involves other complex stages like minutia extraction and matching, there is a desire to evaluate the success of preprocessing stage as-is.

- **Overlapped prints:** Two or more latent fingerprints may be overlapped, as shown in Fig. 2(b). Estimating the orientation of latent fingerprints independently and segmenting individual fingerprints in such cases is also a difficult problem.

Some latent fingerprint segmentation approaches have been developed in literature which are summarized in Table 1. In 2008, Karimi and Kuo [7] proposed an automated latent fingerprint segmentation technique by measuring the variability in the ridge frequency and gradient in the local blocks. The performance of their algorithm was visually demonstrated using two images from the NIST SD-27 database [2]. In 2011, Short et al. [8] proposed a latent fingerprint segmentation technique by cross-correlating latent prints with an ideal template of ridge patterns. The correlation strength identifies ridge-like patterns thereby segmenting the foreground regions. Their algorithm yielded an Equal Error Rate (EER) of 33.8% on the NIST SD-27 database. Choi et al. [9] combined fingerprint orientation tensor and frequency tensor information to segment foreground from background. They showed a rank-1 performance accuracy of 16.28% on the NIST SD-27 and 35.19% on the WVU database. Their approach also yielded a Missed Detection Rate (MDR) of 14.78% and a False Detection Rate (FDR) of 47.99% in NIST SD-27 database. Zhang et al. [5] proposed an Adaptive Directional Total Variational (ADTV) model which is a variation of TV-L2 model. The ADTV model is suitable for decomposing textures with orientation patterns. The orientation pattern forms a defined structure in foreground ridge regions. Their approach yielded a Rank-1 identification accuracy of about 2% on NIST SD-27, with a MDR of 14.10% and FDR of 26.13%. Recently, Cao et al. [6] used a combination of coarse and fine structure ridge dictionary to learn a sparse representation of ridge-like patterns. Rank-1 identification accuracy of 61.24% was reported on NIST SD-27 and 70.16% on WVU database.

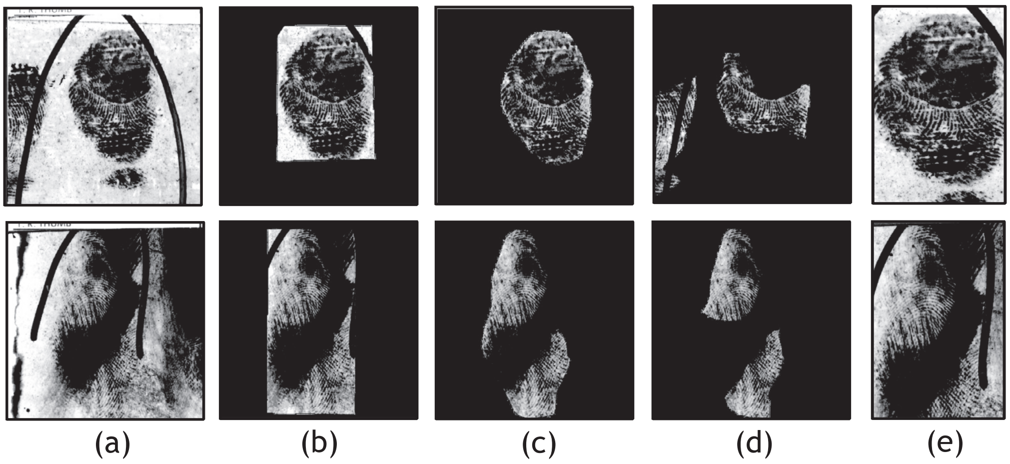


Fig. 3. Sample latent fingerprints from the NIST SD-27 fingerprint database [2] showing segmentation results. (a) Original latent fingerprint images, (b) manually segmented output with a bounding box around the fingerprint region, (c) manually segmented output with exact boundary around the fingerprint region, (d) manually segmented output with only useful ridge information rejecting all the smudgy and noisy (non-informative) regions, and (e) segmented output from *NFSEG* module of NBIS [10].