Contactless and Partial 3D Fingerprint Recognition using Multi-view Deep Representation

Chenhao Lin, Ajay Kumar

Department of Computing, The Hong Kong Polytechnic University, Hong Kong

Abstract

Contactless 3D fingerprint identification has gained significant attentions in recent years as it can offer more hygienic, accurate and ubiquitous personal identification. Despite such advantages, contactless 3D imaging often results in partial 3D fingerprints as it requires relatively higher cooperation from users during the contactless 3D imaging. Such contactless 3D fingerprint images significantly degrade matching accuracy due to partial 3D fingerprint imaging. This paper proposes an end-to-end contactless 3D fingerprint representation learning model based on convolutional neural network (CNN). The proposed model includes one fully convolutional network for fingerprint segmentation and three Siamese networks to learn multi-view 3D fingerprint feature representation. Contactless partial 3D fingerprint identification is a more challenging problem due to its high degree of freedom during contactless 3D fingerprint acquisition and is also addressed by using proposed model. We therefore investigate multi-view 3D fingerprint recognition and partial 3D fingerprint using proposed approach. Comparative experimental results, presented in this paper using state-of-the-art 3D fingerprint recognition method, demonstrate the effectiveness of the proposed multi-view approach and illustrate a significant improvement of state-of-the-art 3D fingerprint recognition methods.

Keywords: Contactless 3D fingerprint recognition, Partial 3D fingerprint identification, Multi-view CNN.

1. Introduction

Fingerprint is considered as one of the most discriminative biometrics. The high reliability and uniqueness of the fingerprint lead to its widespread use in human authentication [1, 2].

Most of existing fingerprint identification systems available in the market are equipped with contact-based 2D sensors for fingerprint acquisition and recognition [3]. Although these kinds of fingerprint systems have been deployed for a long time, the contact-based acquisition approach by touching the sensors with finger introduces fingerprint deformations and noise which result in the degradation in the fingerprint recognition accuracy [2, 4].

With the rapid development of 3D scanning and reconstruction technologies, contactless 3D fingerprint reconstruction and recognition techniques have been introduced to address the shortcoming of contact-based 2D fingerprint identification system and improve the recognition accuracy [4, 5]. An increasing number of 3D fingerprint identification systems have been developed in recent years [6, 7]. Minutiae-based fingerprint matching approaches including [8] (using 3D minutiae tetrahedron), [9] (using 3D minutiae) and [10, 11] (unwrapped 3D/contactless 2D fingerprint recognition using 2D minutiae) have emerged as the most accurate methods for contactless 3D fingerprints recognition. However, the perspective distortion and reconstruction error during the 3D fingerprint acquisition process can significantly degrade the minutiae extraction accuracy and contactless 3D fingerprint matching performance.

Partial fingerprint identification [12, 13] is a challenging problem and has been investigated for several years. As compared to the contact-based 2D fingerprint acquisition, contactless imaging more often results in partial images which significantly degrades contactless 3D fingerprint matching accuracy. Accurate personal identification using partial 3D fingerprints, which generally degrades from the unintentional 3D translation and rotation of fingers under fixed field of view from the sensors, is more challenging and deserves further investigation or research efforts.

Deep learning technique introduced in recent years has shown its effectiveness in image classification[15], object recognition [16, 17] and feature representation [18]. In particular, convolutional neural network (CNN) approaches have successfully improved state-of-the-art method for 3D object classification and recognition [19, 20]. Meanwhile, the CNN-based approach is also proved to have better feature representative capacity than traditional feature extraction method for biometrics recognition problems [21, 22]. Given the enormous success of CNN-based approaches, we propose to develop such approaches for contactless 3D fingerprint recognition to address the limitations of existing methods and achieve the

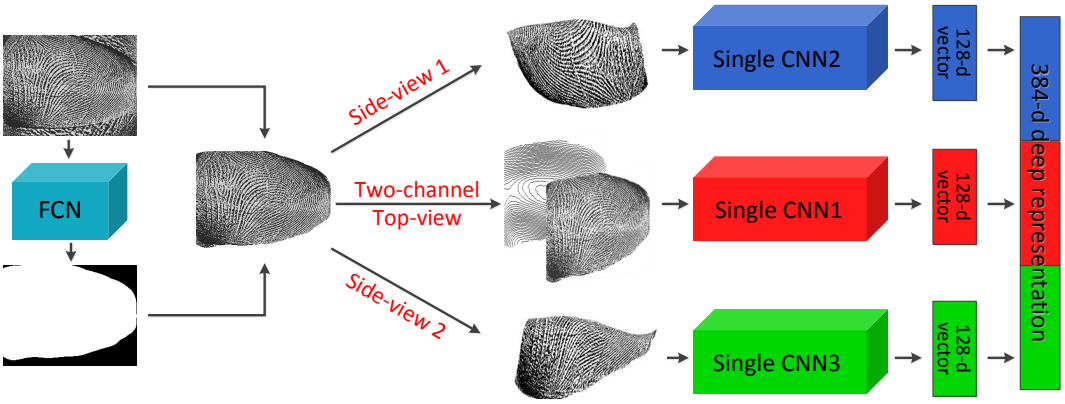


Figure 1: Our model to automatically learn deep 3D fingerprint features for more accurate and partial 3D fingerprint matching

performance improvement of partial 3D fingerprint identification.

1.1. Our Work and Contributions

It is worth noting that the contactless 3D imaging is more prone to partial 3D fingerprint problem, as compared to the respective problem during the contactless 2D fingerprint imaging. This can be largely attributed to the fact that contactless 3D fingerprint imaging process requires acquisition of much more information, e.g. structured light scanning in [11, 23] or multiple 2D images in [9, 24] required for the 3D reconstruction.

Therefore there is strong motivation to address such problem from partial 3D fingerprints for the success of emerging 3D fingerprint technologies. This paper presents a new framework to recover multi-view deep feature representation from contactless 3D fingerprints. The deep features extracted from multiple networks are concatenated and used for contactless 3D fingerprint recognition and partial 3D fingerprint identification. The contour map from 3D fingerprint is also extracted and integrated with multi-channel inputs of proposed multi-view network. A fully convolutional network (FCN) is utilized to automatically segment the fingerprints from background which can result in more accurate matching of 3D and 2D fingerprint. The proposed model for generating 3D fingerprint deep feature representation is illustrated in Figure 1. The comparative summary of different contactless 3D/3D view fingerprint identification methods is illustrated in Table 1.

The proposed model to accurately match 3D fingerprints includes contactless 3D fin-

Table 1: Comparative summary of contactless 3D/3D view fingerprint identification methods

	Database	**Matching strategy**	**Partial 3D fingerprint identification**	**Matched Features**
Reference [9]	Contactless 3D/2D fingerprint data (1920 fingerprints)	3D/2D minutiae matching	No	Hand-crafted
Reference [11]	Contactless/unwrapped 3D view fingerprint data (441 fingerprints)	Unwrapped 2D minutiae matching	No	Hand-crafted
Reference [25]	Contactless 3D fingerprint data (1082 fingerprints)	Curvature features matching	No	Hand-crafted
Reference [24]	Contactless 2D and unwrapped 3D fingerprint data (3000 fingerprints)	Unwrapped 2D minutiae matching	No	Hand-crafted
This Paper	*Contactless 3D/2D and partial 3D fingerprint data (3920 fingerprints)*	*Multi-view deep feature matching*	*Yes*	*Self-learned*

gerprint segmentation network and multi-view 3D fingerprint deep feature representation generation network. The deep feature representation generated from this model is designed to learn and aggregate discriminative ridge and valley details from multiple views of 3D fingerprint images. By integrating the side-view and top-view 3D fingerprint information with respective 3D contour map, the deep feature representation can be enriched with the global shape and texture information. Such deep multi-view 3D fingerprint representation can offer more accurate 3D fingerprint recognition as indicated from rigorous experimental results presented in Section 5 of this paper.

As compared with contact-based partial fingerprint identification, contactless partial 3D fingerprint identification is more challenging and results from its high degree of freedom during the fingerprint acquisition. The key reason for such partial fingerprints, during contactless 3D fingerprint imaging, is due to the limited depth of focus and field of view with commonly used cameras. In this work, partial 3D fingerprint for training the network is defined by 3D rotation of the fingerprints in a plausible range of angles. The deep feature representation from partial 3D fingerprint is generated by using proposed network. This

paper makes first such attempt to address partial 3D fingerprint identification problem and it is important for the success of emerging contactless 3D fingerprint identification. As compared with conventional 3D fingerprint matching methods, our approach can offer significantly improved performance for the partial 3D fingerprint identification.

The comparative experimental results using two different publicly available databases, presented in Section 5.3 of this paper, indicate that our approach can significantly improve the performance over state-of-the-art methods for the contactless 3D fingerprint recognition and partial 3D fingerprint identification. The tests for the statistical significance in Table 3 also illustrate significant improvement in matching performance with the usage of our approach.

1.2. Related Work

Research on contactless 3D fingerprint acquisition and recognition has attracted growing attentions in recent years. Many promising publications have introduced different methods to acquire and match 3D fingerprints [4, 7]. References [26, 27] successfully acquired and recovered microscopic surface geometry details, including palm and fingers, from textured surfaces using elastomeric sensors with a high-resolution camera. However, the subjects skin surface (finger) had to contact with such elastomeric sensor which results in fingerprint deformation. Ultrasound-based approaches have been used in many publications [28, 29, 30, 31] to acquire fingerprint images in 3D space. Although these ultrasound 3D imaging techniques presented 3D fingerprint in high resolution, they failed to extract features or match 3D fingerprint to ensure the effectiveness of their acquisition approach.

In addition to the development of 3D fingerprint acquisition approach, some commercial systems [32, 33] and promising work [6, 11] have been introduced to study 3D fingerprint recognition problem. However, most of these systems [32, 11] extracted features from unwrapped 2D fingerprint instead of using 3D features which can be beneficial for 3D fingerprint recognition. In references [11, 23], structured light illumination was utilized to recover 3D depth information and reconstruct 3D fingerprint. Fingerprint 2D minutiae features were extracted from unwrapped 3D fingerprint to match each two fingerprints. References [25, 34] reconstructed 3D fingerprint using multiple touchless fingerprint images acquired from multi-view cameras. Some low level of fingerprint features like Scale Invariant Feature

Transformation (SIFT) feature, curvature feature were used as additional features for fingerprint recognition. More recently, the authors in references [8, 9, 4] have proposed low-cost 3D fingerprint identification systems. 3D fingerprints were reconstructed from several contactless 2D fingerprints under different illuminations using photometric stereo techniques. The 3D minutiae features and 3D minutiae tetrahedron features, which can be directly used for 3D fingerprint matching without unwrapping the fingerprints, were successfully extracted from the reconstructed 3D fingerprint point cloud data. Despite all the above successes to advance 3D fingerprint recognition techniques, the perspective distortion and reconstruction error during the 3D fingerprint acquisition process can degrade the feature extraction accuracy and need attention to further improve the 3D fingerprint recognition accuracy.

There have been many promising efforts to improve accuracy for matching contact-based partial 2D fingerprint images in the literature [12, 14, 35, 13, 36]. Reference [35] proposed to use minutiae based matching method in the overlapping areas of partial 2D fingerprints and full 2D fingerprints. Wang et.al. [12] attempted to estimate the missing orientation structures of partial fingerprint which highly improved the retrieval efficiency for partial 2D fingerprint identification. A region-based partial fingerprint matching method was proposed in reference [36]. Pixel-level correlation coefficient between partial and full fingerprint was computed for partial fingerprint identification. The experimental results on FVC dataset [37] demonstrated that better performance can be achieved by their method than other methods. Although the existing methods have improved performance for partial fingerprint identification, there is nil research on partial 3D fingerprint identification.

With the success of CNNs on image recognition tasks, several publications have investigated 3D object recognition using CNN-based approaches [20, 38, 39]. Reference [20] proposed to use 3D CNN on volumetric grids and successfully generated volumetric representations for 3D object recognition. However, the CNN architecture proposed in these approaches were designed for recognizing the global shape of 3D volumetric data instead of recognizing local details of 3D surface like 3D fingerprint. Different CNN-based methods were implemented in [39, 40, 41] to generate 3D deep descriptors from 3D mesh or 3D point cloud data for object recognition or retrieval. As compared with identifying 3D biometrics data (face or fingerprint), the tasks in these references are less challenging because they

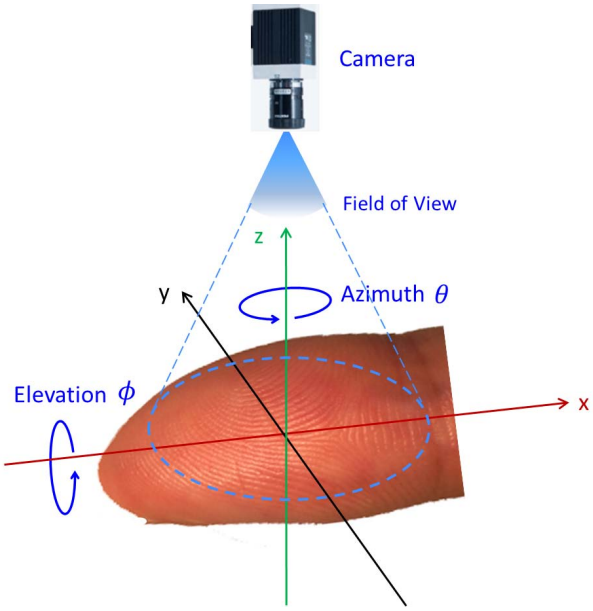


Figure 2: 3D rotation of the fingerprint with respect to the camera. The possible partial 3D fingerprint due to the camera's field of view is also illustrated.

only need to identify or retrieve the 3D data from different categories. References [42, 43] successfully generated deep representations from 3D face using CNN-based approaches and achieved better recognition performance than the methods using traditional face features. These achievements on 3D data recognition, especially on 3D face recognition, have indicated that 3D fingerprints can also be identified using CNN-based approach.

2. MULTI-VIEW 3D FINGERPRINT GENERATION

Contactless partial 3D fingerprint identification is a challenging problem that needs to be addressed for the success of emerging 3D fingerprint systems during deployment. Partial 3D fingerprints are essentially from one of the side-view 3D fingerprint and can be defined by 3D rotation of the 3D fingerprints in a plausible range of elevation and azimuth angles. In this work, we firstly introduce our approach for generating multi-view 3D fingerprint deep feature representation. This approach is then incorporated in experiments to investigate more accurate partial 3D fingerprint identification.

Contactless 3D fingerprint recognition offers significant potential to achieve superior per-

formance than traditional contact-based fingerprint recognition because 3D fingerprint can contain more discriminative information and can be acquired in deformation free manner. Existing approaches for 3D fingerprint recognition include matching 2D minutiae extracted from unwrapped 3D fingerprints [11, 24] and matching 3D minutiae or curvature features directly extracted from 3D fingerprints [8, 4, 25]. The first approaches [11, 24] fail to extract or match the 3D information available in the 3D fingerprint. Due to the acquisition and reconstruction errors, the curvature or 3D minutiae based features extracted using the second approaches [8, 4, 25] may be inaccurate and can result in degraded matching performance. When human experts recognize the 3D fingerprint, they also identify the global shape and local surface details of 3D fingerprint from different perspective angles. Inspired by such references, we investigate to use multi-view deep learning based approach to learn the preferences in human visual system for recognizing 3D fingerprints. As shown in Figure 2, the view angles of 3D fingerprint are defined using (θ, ϕ), which represent the horizontal rotation azimuth and vertical rotation elevation angles. Based on the 3D fingerprint point cloud data, different 3D fingerprint rotation views can be generated using the following rotation matrix,

$$R(\theta, \phi) = R_z(\theta) \times R_x(\phi) \tag{1}$$

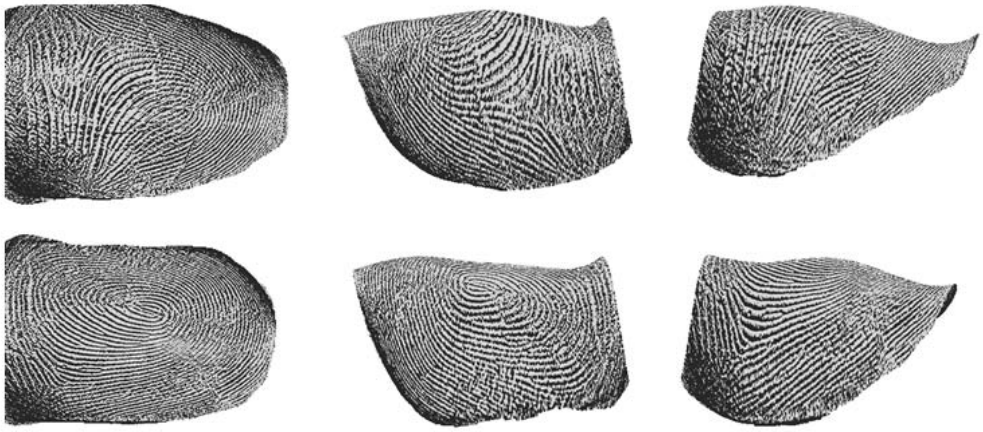
where $R_z(\theta)$ means rotation around the z axes and $R_x(\phi)$ means rotation around the x axes. These matrices can be represented by the following equations,

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \tag{3}$$

Each 3D fingerprint P can be represented using $(\mathbf{x}, \mathbf{y}, \mathbf{z})$, then the rotated fingerprint can be represented using the following equation,

$$\begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \\ \mathbf{z}' \end{bmatrix} = R_z(\theta) \times R_x(\phi) \times \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \end{bmatrix} \tag{4}$$



(a) (b) (c)

Figure 3: (a) Top-view 3D fingerprints (b) Side-view 3D fingerprints with $\theta = 170, \phi = 50$ (c) Side-view 3D fingerprints with $\theta = 10, \phi = 50$.

Three projections are generated from different 3D views that includes one top-view P_t and two side-views (P_{s1} and P_{s2}). The top-view fingerprint is essentially the 3D fingerprint when observed from the top and the two side-view visualization of 3D fingerprints can be defined by different, or a range of, plausible observation angles. In order to avoid the repetition of top-view 3D fingerprint (elevation angle ϕ is too small) and the loss of too much 3D fingerprint details (elevation angle ϕ is too large), and simulate the plausible observations for identifying 3D fingerprints (azimuth angle θ is close to 180 in P_{s1} and close to 0 in P_{s2}), the range of multi-view azimuth angles and elevation angles are selected as follows,

$$\begin{aligned} P &= P_t, (\theta = 0, \phi = 0) \\ P &= P_{s1}, (\theta = 170, \phi = 50) \\ P &= P_{s2}, (\theta = 10, \phi = 50) \end{aligned} \tag{5}$$

The rotation of 3D fingerprint, with respect to the field of view from fixed camera, is illustrated in Figure 2. Figure 3 illustrates sample multi-view images from 3D fingerprint that can be observed from a fixed camera.

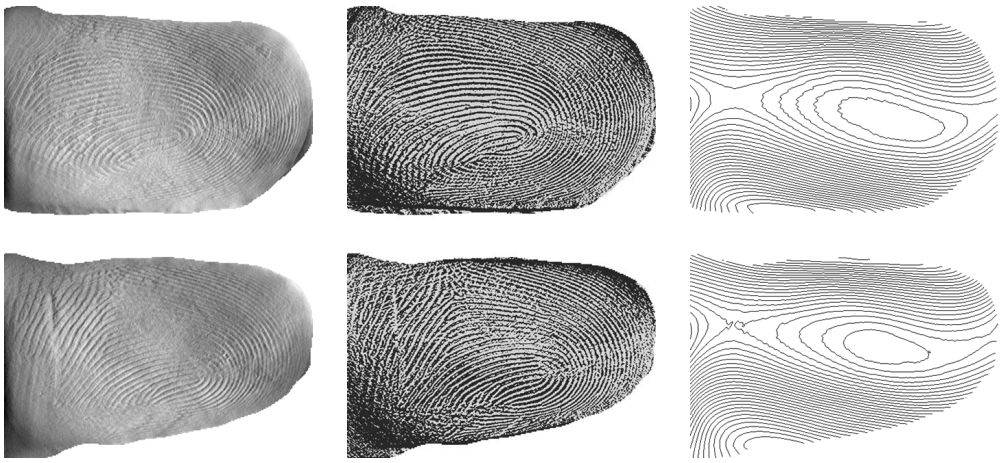
In order to ensure that our network can automatically and efficiently learn the texture details of 3D fingerprint surface, different surface preprocessing approaches are incorporated on 3D fingerprint point cloud data. One straightforward way to represent top-view 3D

fingerprint is to compute the gradient value of 3D fingerprint depth data. For 3D fingerprint depth, the magnitude of the gradient represents how fast the depth changes in a given direction. Therefore local 3D fingerprint ridge and valley details can be represented using the magnitude of gradient value. Another way to illustrate different views of reconstructed 3D fingerprint is to use 3D surface curvature. Curvature feature can be directly extracted from 3D fingerprint surface following the steps described in [44]. Based on the observed computed principle curvature, fingerprint surface shape index [9] can be also computed as the local ridge-valley descriptor. Adaptive histogram equalization is used to normalize and smooth the fingerprint ridge-valley representation. More clear surface details can be observed using the ridge-valley descriptor generated from surface shape index. The experimental results in Section 5 also illustrate that matching 3D fingerprints using ridge-valley descriptor can be more accurate than using depth gradient.

In addition to using fingerprint top-view and side-view images, 3D contour line map from the reconstructed 3D fingerprint images is also used to represent the global shape and depth information. The contour map, which is widely used to represent the 3D topography, can represent steepness of slopes resulting from ridges and valleys of 3D fingerprint surface. Thus, this map is suitable for describing reconstructed 3D fingerprints. For each of the reconstructed 3D fingerprint P, each pixel on P can be represented using (x, y, z), where (x, y) are the 2D location and z is the depth information. The maximum depth can be computed as z_{max} and the minimum depth can be represented as z_{min}. Then the interval z_{int} for generating 3D contour map can be computed as follows,

$$z_{int} = (z_{max} - z_{min})/N \tag{6}$$

where N is the number of contour lines in each contour map. For each reconstructed 3D fingerprint P, the corresponding contour map is generated using the contour lines with computed interval z_{int}. In such a way, different 3D fingerprints that have different depth details are expected to be represented from the variation of the contour map. In this work, we combine 3D fingerprint top-view and corresponding contour map into 2-channel input for robustly training our network. Incorporating 3D fingerprint ridge-valley descriptor with the corresponding contour map can help network learn both texture details and global shape from 3D fingerprint. Figure 4 illustrates different approaches to represent 3D fingerprints



(a) (b) (c)

Figure 4: (a) Top-view 3D fingerprints surface using gradient operation (b) Top-view 3D fingerprints surface ridge-valley representation (c) Top-view 3D fingerprint contour maps.

surface (top-view images) and corresponding contour maps.

3. MULTI-VIEW FINGERPRINT SEGMENTATION

Automated segmentation of the region of interest from the acquired fingerprint images is generally the first step for fingerprint identification. For traditional contact-based fingerprint recognition approach, it is a critical step to remove fingerprint residual or considerable (noisy) background regions. Fingerprint minutiae can also be more accurately extracted from such segmented images. Although there are no fingerprint residual for contactless 3D fingerprint images, the complex background, especially the reflected light from background of 3D fingerprint images, which is generated by using photometric stereo approach [27, 9], can also degrade the feature extraction and recognition performance. In addition, the high quality segmented 3D fingerprints can better describe the shape of 3D surface which is highly desirable for proposed CNN-based approach and can result in more accurate 3D fingerprint matching.

Fully Convolutional Networks proposed in [46] are efficient for the image segmentation. The pixels-to-pixels network architecture is based on existing well-trained deep networks like AlexNet [47], the VGG net [16], and GoogLeNet [17], but the fully connected layer is

replaced by convolutional layer to realize more accurate semantic segmentation. The fully convolutional network attempts to learn the images deep representations based on local spatial input and have achieved state-of-the-art performance for semantic segmentation on many public datasets. Therefore, in this work, this approach is adopted by fine-tuning VGG net to achieve accurate fingerprint segmentation. The foreground and background of 40 top-view fingerprints (P_t) from different subjects in the training dataset are manually marked for training the FCN. The probability p of each pixel (x, y) belongs to foreground or background is estimated. The labels l $(0, 1)$ of each fingerprint are assigned based on the probability using the following equation,

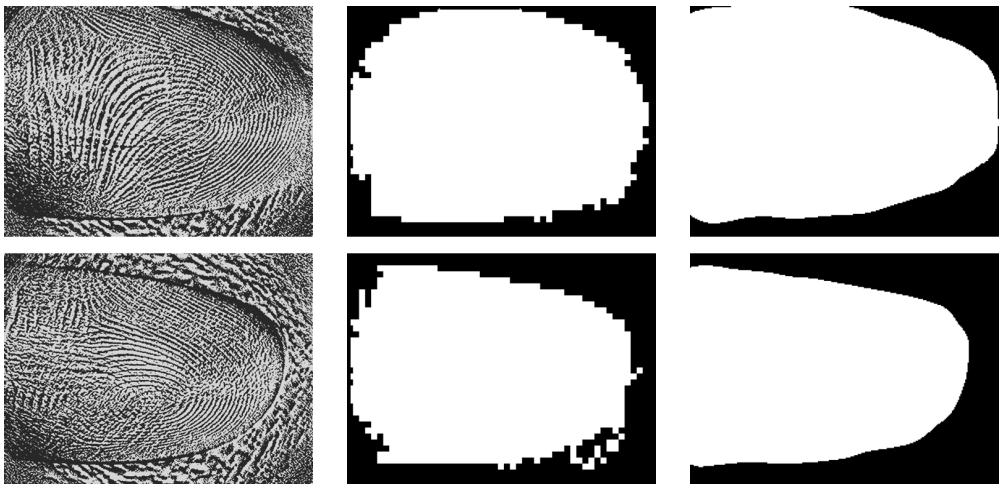
$$l = \begin{cases} 1, \ if \ p\left(P\left(x,y\right)\right) = 1 \\ 0, \ if \ p\left(P\left(x,y\right)\right) = 0 \end{cases} \tag{7}$$

The *softmax* loss function is used to predict the label of each pixel in each top-view fingerprints (P_t). Figure 5 illustrates the segmented masks using FCN approach and traditional ridge-based segmentation method [45]. It can be observed that smoother fingerprint boundaries can be obtained using FCN approach. The comparative results in Section 5 also indicate that segmenting fingerprint using FCN approach can achieve better performance than using traditional methods.

4. MULTI-VIEW FINGERPRINT RECOGNITION NETWORK

Deep convolutional neural network based approaches are widely applied for image classification [17, 47], image recognition and image representation [48]. However, one limitation of conventional CNN is when the number of the samples from each subject is small, even after applying data augmentation, the insufficient training data can easily result in overfitting problem and poor performance. The recognition accuracy can be further improved by using more powerful network structure.

Siamese convolutional neural network [49] is recently introduced and has shown superior performance on image recognition [18, 50], especially on biometrics recognition [22]. Siamese network consists of two CNNs that shares the same network structure and takes the image pairs (matched pairs and unmatched pairs) as the inputs. Much more training data can be generated by using Siamese network than CNN. Sharing weights between two subnets



(a) (b) (c)

Figure 5: (a) Top-view 3D fingerprints ridge-valley representation (b) 3D fingerprints mask generated using method in [45] (c) 3D fingerprint mask generated by fine-tuning FCN.

means fewer parameters are needed for training, which results in the fact that less data are required. To some extent, it can avoid overfitting problem caused by insufficient training data. Furthermore, the image representation produced from each subnet can be stored as the feature vectors. The extracted feature vectors are easier to compare by computing the Euclidean distance and they usually take less storage space than original images, which make the CNN-based approach suitable for the real application deployments. Therefore, in this work the proposed network is trained using Siamese network architecture.

4.1. Multi-view Siamese Network

As introduced in Section 2, three different views including one top-view image and two side-view images generated from the reconstructed 3D fingerprint data are used for training the proposed network. Three Siamese networks (*top-net, side-net1 and side-net2*) with same structure are trained separately using 3D fingerprint top-view and side-view images respectively. Each network includes two single CNNs and the input pairs consists of the fingerprint images from the same or different subjects.

For each CNN from Siamese sub network, the network structure includes six convolution layers and one fully connected layer. The first five layers are followed by max pooling layer.

The input patch size for each CNN is 256×192. The filter size of each convolution layer is same and it is 3×3. 48 feature maps are generated from the first convolution layer. The output numbers of feature map produced from the second to the sixth convolution layer are 64, 96, 128, 256 and 512. The filter size of each max pooling layer is 3×3 and the stride is 2. The fully connected layer outputs 128 feature maps. Each two corresponding convolution layers from two channels CNNs share the same weights. Rectified Linear Unit (ReLU) is applied as the activation function after each convolution layer and the fully connected layer. Contrastive loss function L is designed for measuring the similarities between each input pairs and is widely used in Siamese network. It is can be defined as,

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i\, {d_i}^2) + (1 - y_i)max(\alpha - d_i, 0)^2 \tag{8}$$

where N stands for the number of samples in each batch. Euclidean distance d_i is computed to measure the similarities between each two feature representations $(f(x_{i1}), f(x_{i2}))$, i.e. feature vector of the fully connected layer of input 3D fingerprint samples (x_{i1}, x_{i2}) in our case.

$$d_i = ||f(x_{i1}) - f(x_{i2})||_2 \tag{9}$$

Each input pairs have the labels $y_i \in 0, 1$ indicate whether they belong to the same subject (genuine pair) or different subjects (imposter pair). The network is trained to ensure that the distance between genuine pair is small and to push the imposter pairs away by using such loss function. It also focuses on training the imposter pairs with distance smaller than margin α.

Each Siamese network is separately trained. During the test phase, three fully connected layers, which can be used to describe different features of 3D fingerprints different view images, are concatenated to form one multi-view feature representation. The deep feature f_{all} generated through the whole network, for each of the 3D fingerprints, can be represented as,

$$f_{all} = [F(P_t); F(P_{s1}); F(P_{s2})] \tag{10}$$

where F is the function or the representation from the trained network. This deep feature representation not only includes 3D fingerprint ridge-valley details from different views but also 3D fingerprint shape and depth information. Therefore it is expected to more robustly describe the reconstructed 3D fingerprints presented for the matching.

4.2. Two-channel Top-view Network

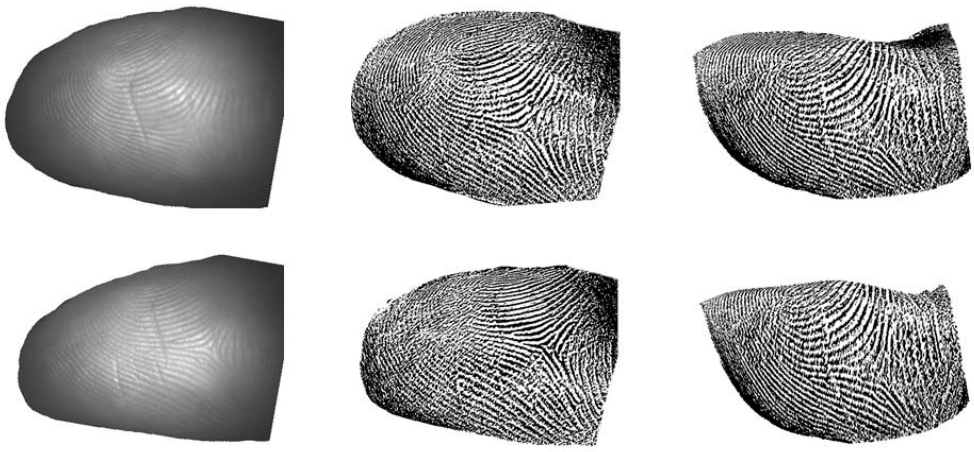
3D fingerprint contour maps generated by using the method described in Section 2 are also incorporated into the proposed network. The contour map consisting of contour lines, which connect points of equal elevation, can be used to indicate the depth variations and global shape of reconstructed 3D fingerprints. In this work, 3D fingerprint top-view images were combined with corresponding contour map into two channels as the inputs of the subnetwork (*two-channel-top-net*). Thus the deep representations generated from this network are expected to be more discriminative with such added information of the global shape and depth variations.

During the test phase, for each of the test or unknown 3D fingerprints, we generate corresponding one top-view fingerprint image and 3D contour map, and two side-view fingerprint images. The respective deep feature representation is generated by concatenating the feature vectors produced from three subnetworks. The matching score S for each two 3D fingerprints (x_{i1}, x_{i2}) can be directly computed using the Euclidean distance between each two respective deep feature representations.

$$S = ||f_{all}(x_{i1}) - f_{all}(x_{i2})||_2 \tag{11}$$

4.3. Partial 3D fingerprint Generation and Identification

Partial fingerprint identification using conventional sensors has always been a challenging problem. Almost all the research in the literature focus on contact-based 2D partial fingerprint identification or latent partial fingerprint identification problems [35, 12, 36]. During the 3D fingerprint acquisition, the unintentional rotation of subject's fingers in 3D space can easily generate partial 3D fingerprint. Such partial 3D fingerprint identification is more challenging and needs attention for the success of emerging 3D fingerprint technologies. Similar to the process for generating side-view fingerprints described in Section 2, we generate partial 3D fingerprint by randomly rotating the 3D fingerprints in a plausible range of elevation and azimuth angle. Such 3D fingerprint images represent partial 3D fingerprints acquired from a fixed camera during 3D fingerprint imaging. Some real partial 3D fingerprints acquired from the same 3D fingerprint imaging setup are also illustrated in Figure 6 for the comparison with simulated partial 3D fingerprints. It can be observed that the partial 3D fingerprints generated from the rotation of 3D fingerprints are quite similar to the



(a) (b) (c)

Figure 6: (a) Partial 3D fingerprint image samples acquired using 3D imaging and (b) corresponding partial 3D fingerprint ridge-valley representation (c) The partial 3D ridge-valley representation from the same 3D fingerprint generated from the rotation of 3D fingerprint with $\theta = 175$, $\phi = 60$ (first row) and $\theta = 170$, $\phi = 50$ (second row).

partial 3D fingerprints directly captured using the camera, which can justify the feasibility of partial 3D fingerprint generation approach for the performance evaluation.

In addition to generating partial 3D fingerprints, we also define partial 3D fingerprint identification protocol. In the literature [12, 36], contact-based 2D partial fingerprints used for the performance evaluation are usually generated from the randomly cropped first impression of each finger and they are matched against the second impression of all fingers for evaluation. In a similar way, partial 3D fingerprints are randomly generated from first 3D fingerprint sample for the performance evaluation. In order to simulate realistic scenarios of unintentional partial 3D fingerprint acquisition and avoid the generation of full 3D fingerprint, the multiple random partial 3D fingerprints are generated from each sample in a plausible or realistic range of rotation angles. The plausible azimuth and elevation rotation angle ranges were from $(180, 60)$ to $(150, 30)$ and $(0, 60)$ to $(30, 30)$. The gallery set contains 3D fingerprints generated from the second sample with two rotation angles $(165, 45)$ and $(15, 45)$. Based on different 3D rotation angles, partial 3D fingerprint can be easily classified into two types. The first type of partial 3D fingerprint (PF_1) is the fingerprint with azimuth

rotation larger than 90 degree and the other type of partial 3D fingerprint's (PF_2) azimuth rotation is smaller than 90 degree.

Partial 3D fingerprints represent incomplete details which make it difficult to extract accurate feature for the 3D fingerprint identification. In order to achieve more accurate partial 3D fingerprint identification, Siamese CNN based architecture proposed in Section 3 is also investigated to generate partial 3D fingerprint deep representation features.

5. EXPERIMENTAL EVALUATION

In this section, results from a range of experiments are reported to evaluate the effectiveness of proposed approach for contactless 3D fingerprint recognition and partial 3D fingerprint identification. The proposed two-channel multi-view Siamese CNN based approach is compared with state-of-the-art fingerprint matching method using minutiae features. The Receiving Operating Characteristic (ROC) curve and respective Equal Error Rate (EER) are generally used for evaluating the performance of biometrics identification system. We also adopt these two evaluation protocols to ascertain the effectiveness of proposed method. In addition, the Cumulative Match Characteristics (CMC) as well as rank-one accuracy are also employed to ascertain recognition performance from the proposed approach.

5.1. Implementation Details

The proposed approach was implemented on a PC with NVIDIA 980Ti GPU for training and testing based on an open-source and with a widely used deep learning framework Caffe [51]. The initial learning rate is set to the same value (0.01) for all the subnetworks and it is dropped by using inverse decay function throughout training. We minimize the loss function using a stochastic approximation of gradient descent. During the training phase, we find that training top-view subnetwork achieves faster convergence than training side-view subnetworks because top-view fingerprints contain more ridge-valley details and have more clear shape. Thus the training process of top-view subnetwork is stopped around 20K iterations and it is stopped around 70K iterations for training side-view subnetworks when the loss fails to decrease.

5.2. Database Description

In order to evaluate the performance and compatibility of the proposed approach, two publicly available databases [24, 52] are employed.

Contactless 3D fingerprint Database. The first database (Database A) [52] consists of 3D fingerprints acquired from 336 different fingers and 160 fingers of them have two-session 3D fingerprints. The resolution of original 3D fingerprints is 2048×1536. In order to avoid the repetition of samples in training and test set, the first session data is used in this work. Each finger has six 3D fingerprint samples. This dataset is divided into training set containing 960 (160×6) 3D fingerprint samples, test set containing the other 960 samples and validation set (96 samples). The top-view and side-view 3D fingerprints are generated by applying the approach described in Section 2.

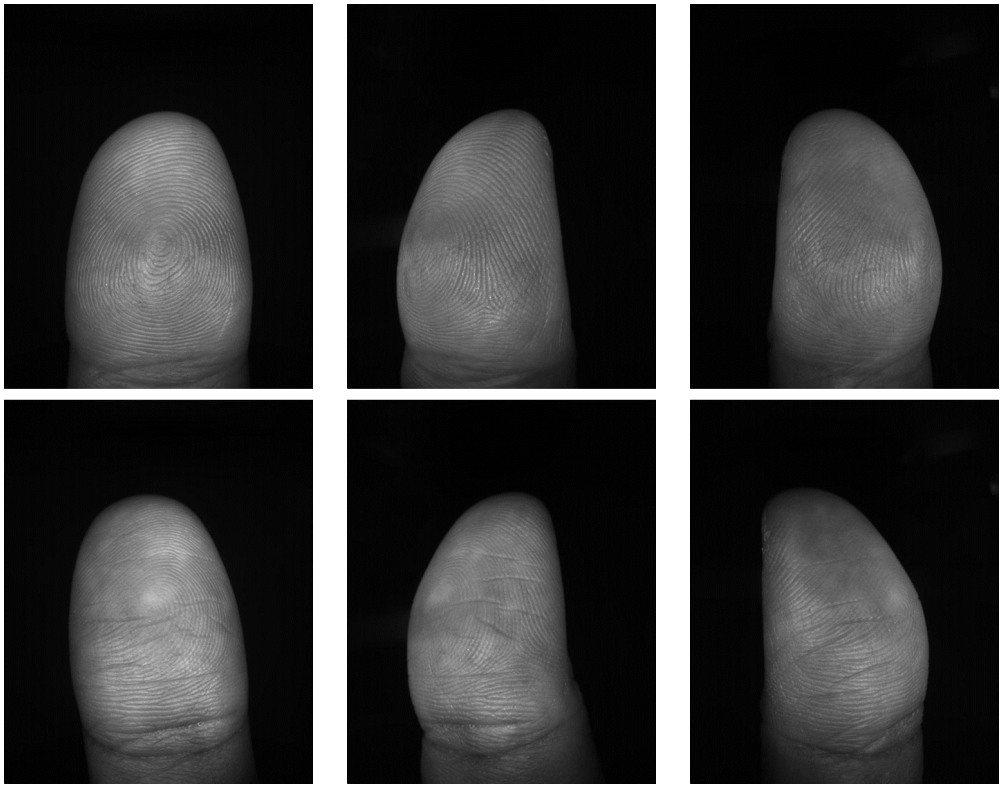
All the inputs fingerprint images are resized into 256×192 resolution for training. In such way, the pattern (ridge details) of the fingerprint images with relatively low resolution can still be identified and using small size images as inputs can result in more efficient training. As indicated in Section 2, we attempt to simulate the two side-view fingerprints (P_{s1} and P_{s2}) with plausible rotation angles ($\theta = 170, \phi = 50$) and ($\theta = 10, \phi = 50$). During the fingerprint acquisition process, the unintentional finger rotation can make the 3D side-view fingerprints from same subject be very different. Hence we apply data augmentation operation on side-view fingerprint training sets. For each side-view fingerprint P_{s1}, four more samples with different rotation angles ($\theta = 170, \phi = 40; \theta = 170, \phi = 60; \theta = 180, \phi = 60; \theta = 160, \phi = 40$) are generated. For each side-view fingerprint P_{s2}, four more samples with different rotation angles ($\theta = 10, \phi = 40; \theta = 10, \phi = 60; \theta = 0, \phi = 60; \theta = 20, \phi = 40$) are generated. The size of each side-view fingerprint set is increased five times after data augmentation. For matching 3D fingerprints using this dataset, 2400 ($160 \times 6 \times 5/2$) genuine matching scores and 457920 ($160 \times 6 \times 159 \times 6/2$) imposter matching scores are generated.

In order to investigate partial 3D fingerprint identification, we randomly generate partial 3D fingerprints following the protocol introduced in Section 4.3 using the test set of this database. Each 3D fingerprint in this database has six samples. The first three samples are used to form gallery set and the other three samples form test set. For each 3D fingerprint sample in gallery set, two partial 3D fingerprints are generated with different but more realistic rotation angles $(165, 45)$ and $(15, 45)$. Test set contains twelve partial 3D

fingerprints generated from the other three fingerprint samples. Each fingerprint in test set generates four random partial 3D fingerprints in a fixed rotation angle range. The plausible azimuth and elevation rotation angle ranges are from $(180, 60)$ to $(150, 30)$ and $(0, 60)$ to $(30, 30)$. In order to simulate the real fingerprint identification process, it is reasonable to assume the types of test fingerprint are unknown. Each test fingerprint will be matched with two types of partial 3D fingerprints in the gallery and the maximum score is selected as the matching score. In such manner, 1920 ($160 \times 12 \times 6/2/3$) genuine matching scores and 305280 ($160 \times 159 \times 12 \times 6/2/3$) imposter matching scores are generated.

Multi-view Contactless Fingerprint Database. The second database (Database B) [24] includes multi-view (three views) contactless fingerprint samples acquired from 1500 fingers using the commercial fingerprint device [32]. For each finger, two fingerprint samples are acquired and each fingerprint sample has three multi-view contactless fingerprints, including one top-view and two side-view fingerprints. Figure 7 illustrates the multi-view fingerprint samples in this database. This database however does not provide full 3D fingerprint data, i.e. fingerprint depth or point cloud, and it is therefore difficult to accurately reconstruct 3D fingerprints from three view contactless fingerprint images in absence of the configuration of their scanner. However these contactless fingerprints can be used for experiments to directly evaluate proposed approaches. Since some fingerprint samples are missing in this database, 1000 fingers are used and divided into training set with 500 fingers and test set with the other 500 fingers. We also generate a small dataset from the other 500 fingers to form validation set. Since the contactless fingerprint images in this database have solid background colors, it's straightforward to segment the fingerprints using conventional fingerprint segmentation method. All the fingerprint images are cropped (aligned) based on the fingerprint center and resized into the same resolution (256×192) as the images in previous dataset.

The fingerprint 3D ridge-valley representation cannot be generated without the knowledge of the fingerprint depth, thus we use conventional Gabor filter [45] based method to enhance the contactless fingerprints. The enhanced fingerprint samples used for training and test are illustrated in Figure 8. In this database, each finger only has two fingerprint samples, which is not sufficient to robustly training the network. Therefore, we augment the dataset by rotating the fingerprint by 5 degrees in 2D space. This step quadruples the

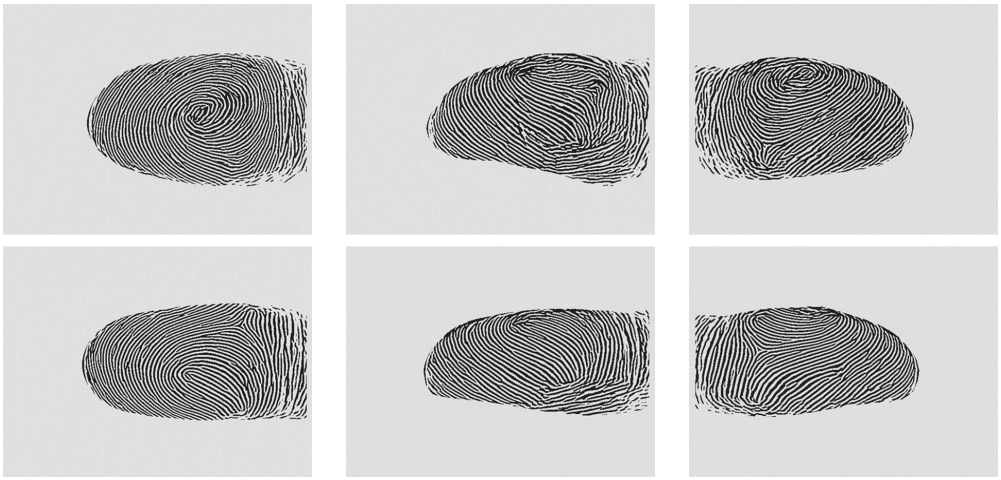


(a) (b) (c)

Figure 7: (a) Top-view contactless fingerprint samples (b) (c) Two side-view contactless fingerprint samples.

size of the training set. For matching multi-view contactless fingerprints using this dataset, 500 ($500 \times 2 \times 1/2$) genuine matching scores and 499000 ($500 \times 499 \times 2 \times 2/2$) imposter matching scores are generated.

Three different approaches are incorporated to train this database. The weights of pre-trained network generated from the first database can be directly used on this database. However, the different image representation of the fingerprints from two databases is expected to offer relatively low recognition accuracy. For this reason, we also adapt the other approaches, i.e. fine-tuning the proposed network and training the proposed network using this new database. The experimental results in Section 5.3 illustrate that fine-tuning or training the proposed network achieves better performance than directly using the network weights.



(a) (b) (c)

Figure 8: (a) Enhanced top-view contactless fingerprint samples (b) (c) Two enhanced side-view contactless fingerprint samples.

5.3. Experimental Results

In this section, we present a series of comparative experimental results to validate the effectiveness of the proposed method on contactless 3D fingerprint recognition and partial 3D fingerprint identification.

Firstly several experiments were performed to evaluate the influence from the proposed image preprocessing, segmentation operations and two-channel network on 3D fingerprints recognition using dataset A. The first experimental results that shown in Figure 9 illustrate that better performance can be achieved by using 3D fingerprint ridge-valley descriptor instead of using 3D fingerprint depth gradient for matching multi-view 3D fingerprints. The EERs for matching one top-view and two side-view depth gradient fingerprints are 3.72%, 3.73% and 4.72%. Respective error rate reduces to 1.83%, 2.04% and 2.49% when using 3D fingerprint ridge-valley descriptor. Then the effects of different segmentation approaches were evaluated and illustrated in Figure 10. We can observe that more accurate and smooth shape can be generated by using FCN-based approach than using conventional method for fingerprint segmentation, which also resulted in more accurate 3D top-view and side-view fingerprint matching. The corresponding EERs are also reduced from 2.98% to 1.83% for top-view fingerprint matching, from 2.30% to 2.04% and from 3.19% to 2.49%

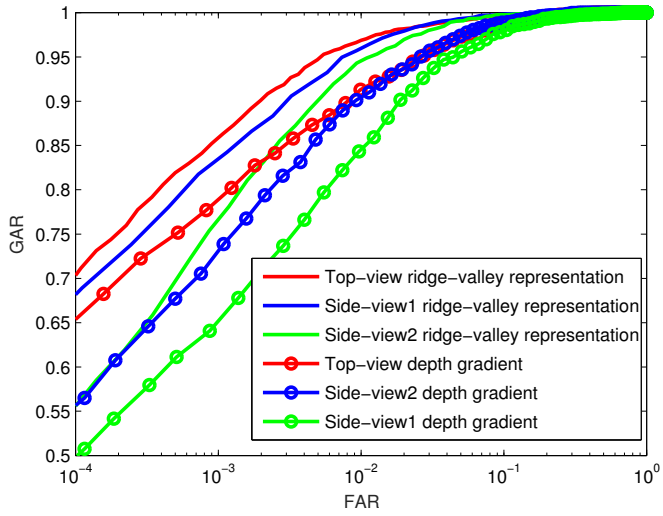


Figure 9: Comparative ROC curves for different views 3D fingerprint matching using different fingerprint representation.

for side-views fingerprint matching. The results for matching multi-view 3D fingerprints using different preprocessing approaches, segmentation methods and two-channel network were also illustrated in Figure 11. As indicated from the comparative results, matching 3D fingerprints using two-channel networks with FCN-based segmentation achieves best performance with the EER of 0.64%. In addition, we also compared proposed method with existing state-of-the-art deep learning based approach, i.e. fine-tuning *Resnet* [48] on different view 3D fingerprints with FCN-based segmentation. In order to fine-tune *Resnet*, we cropped the images into 1536×1536 based on the fingerprint center and resized them into 224×224. From the comparative results illustrated in Figure 12, we can observe that proposed method achieved better performance.

Then several experiments were performed to compare the proposed approach with traditional 3D or 2D minutiae based methods using two databases. In order to ensure fairness in comparison and to achieve the best performance of conventional minutiae based method, in database A, the minutiae were extracted from the fingerprint images with 512×384 resolution and the resolution of fingerprint images in database B is 512×640. We followed the approaches in [8, 9] to extract 3D minutiae and 3D minutiae tetrahedron from ridge-valley top-view 3D fingerprints, 3D fingerprint curvature images and depth gradient top-view 3D

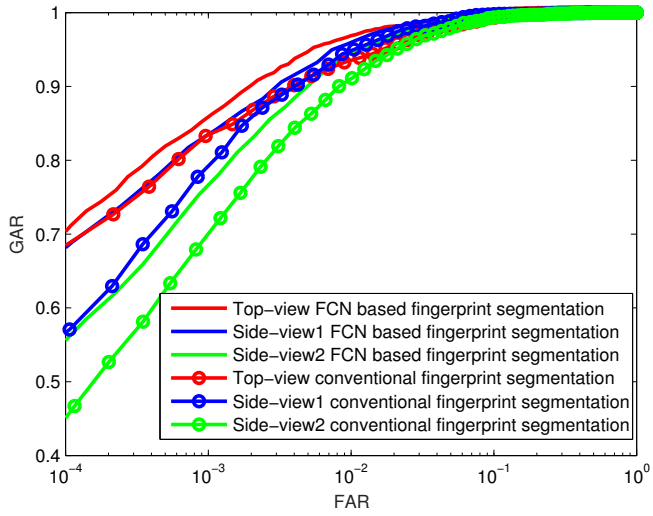


Figure 10: Comparative ROC curves for different views 3D fingerprint matching using different fingerprint segmentation methods.

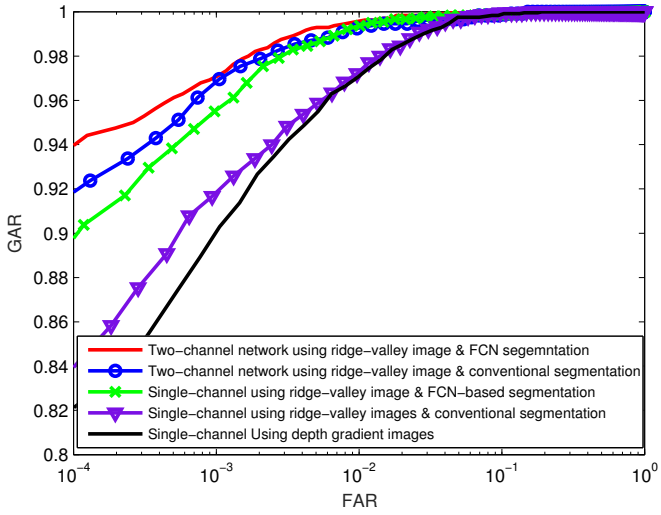


Figure 11: Comparative ROC curves for multi-view 3D fingerprint recognition using different processing methods and networks.

fingerprints for comparison using database A. In addition, minutiae were also extracted from ridge-valley side-view fingerprints for matching and the score fusion was applied to combine the matching score generated from one top-view and two side-view fingerprints. The fingerprint 2D minutiae extracted can be represented as $m = [x, y, \theta, q, t]$, where x, y are minutiae

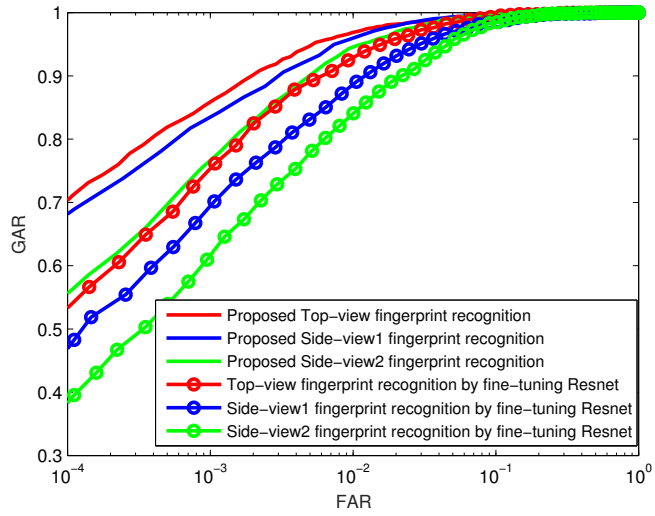


Figure 12: Comparative ROC curves for different views 3D fingerprint matching by using proposed method and fine-tuning *Resnet*.

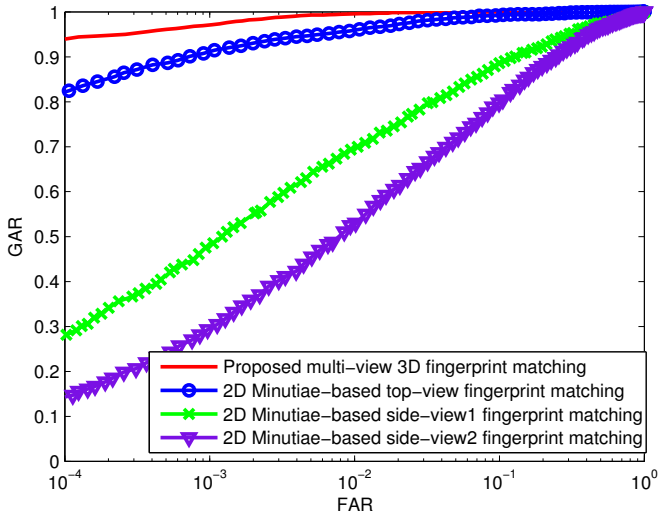


Figure 13: Comparative ROC curves for 3D fingerprint recognition using proposed methods and minutiae-based methods on multi-view fingerprints.

location, θ is the minutia direction, q is the minutia quality and t is minutia type and the 3D minutiae [9] extracted can be represented as $m = [x, y, z, \theta, \phi, q, t]$, where z is minutia height, ϕ is the 3D minutia orientation. The average number of minutiae extracted from top-view, side-view1 and side-view2 fingerprints is 65.73, 16.46 and 19.18 respectively. Since

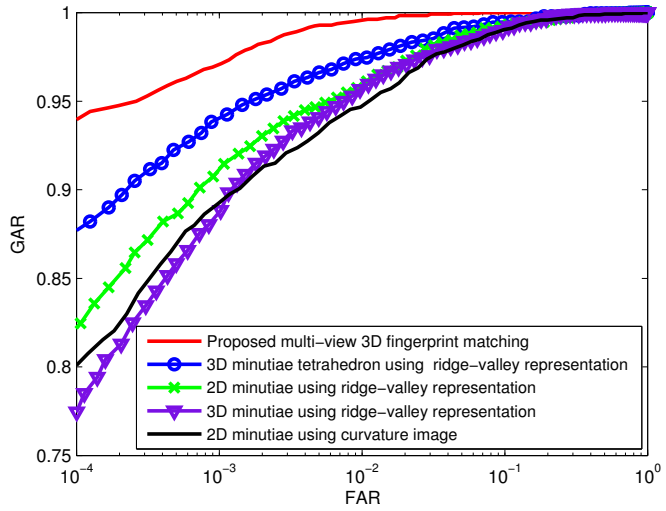


Figure 14: Comparative ROC curves for multi-view 3D fingerprint recognition using proposed method and minutiae-based method on database A.

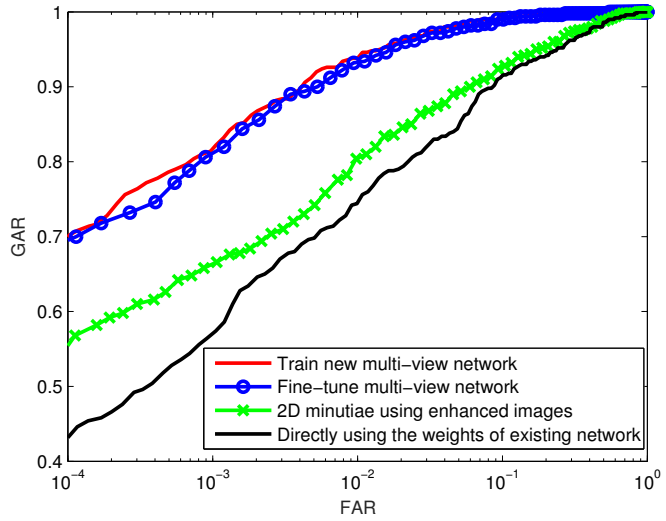


Figure 15: Comparative ROC curves for multi-view 2D fingerprint recognition using proposed method and minutiae-based method on database B.

the minutiae extracted from side-view fingerprints are relatively inaccurate, the matching performance is not expected to be improved by the combination of top-view and side-view fingerprints matching. Figure 13 illustrates the comparative results for matching 3D fingerprints using proposed method and minutiae-based methods on multi-view ridge-valley

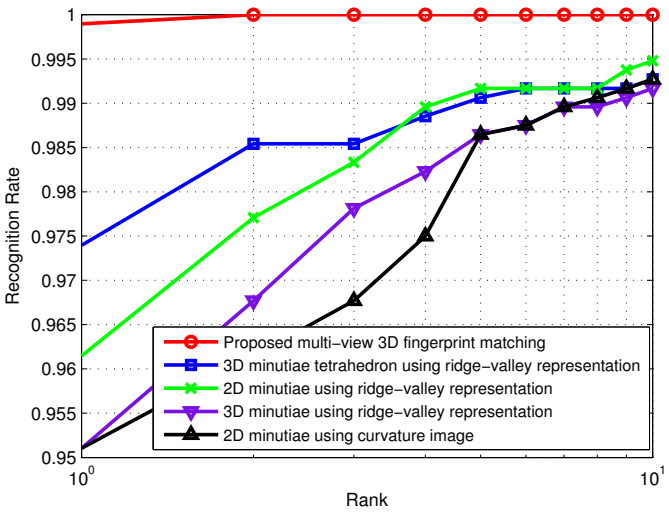


Figure 16: Comparative CMC curves for multi-view 3D fingerprint matching using proposed method and minutiae-based method on database A.

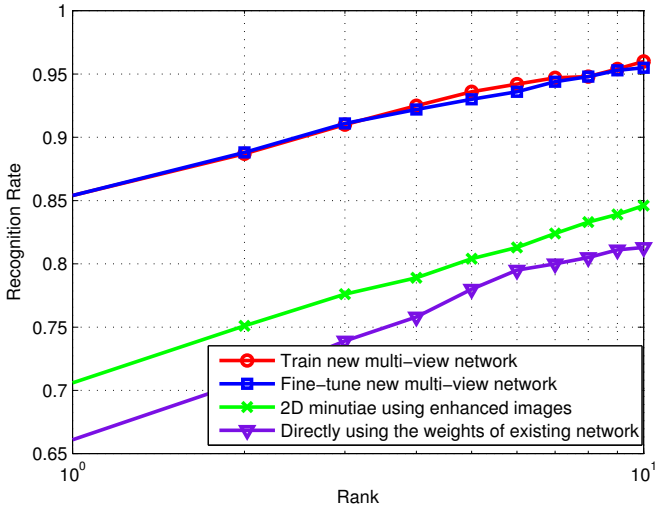


Figure 17: Comparative CMC curves for multi-view contactless 2D fingerprint matching using proposed method and minutiae-based method on database B.

fingerprints. The comparison in Figure 14 also illustrates the performance for matching 3D fingerprint using proposed method and several popular minutiae-based approaches. It can be observed that the proposed method achieves superior performance than other minutiae-based methods. The EER is reduced from 1.86% to 0.64% by using proposed method.

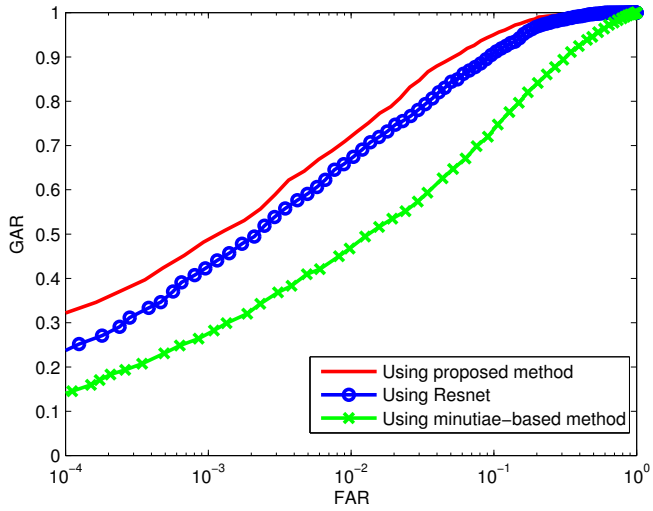


Figure 18: ROC curves for partial 3D fingerprint identification using proposed method, minutiae-based method and *Resnet* on database A.

Table 2: Comparative summary of performance from contactless 3D/2D fingerprint identification methods

Experiments	**Equal Error Rate**	**Rank-one accuracy**
2D minutiae approach on dataset A	2.39%	96.14%
3D minutiae approach on dataset A	2.64%	95.11%
3D minutiae tetrahedron approach on dataset A	1.86%	97.40%
Proposed single-channel multi-view method on dataset A	**0.81%**	**96.25%**
Proposed two-channel multi-view method on dataset A	**0.64%**	**99.89%**
2D minutiae approach on dataset B	8.35%	70.60%
Fine-tune proposed single-channel multi-view on dataset B	**3.02%**	**85.40%**
Train single-channel multi-view on dataset B	**2.84%**	**85.40%**

We also implemented different approaches to match multi-view contactless 2D fingerprints using database B. Firstly, we directly used the weights from the proposed single-

Table 3: Significance of the difference between the areas under the ROC curves

Criterion	***p-value***	***z-statistic***
Multi-view 3D fingerprint matching on dataset A	0.00238	2.8327
Multi-view 2D fingerprint matching on dataset B	0.000013	4.2145
Partial 3D fingerprint identification (compare with using minutiae-based method)	<0.00001	15.3855
Partial 3D fingerprint identification (compare with using *Resnet*)	0.001329	3.2097

channel multi-view network for testing. However the enhanced fingerprints in database B are different with the ridge-valley representation fingerprints in database A, which resulted in relatively poor recognition performance. The EER is 9.12%. Then we attempted to fine-tune the proposed network and directly train the network using the database B. The network architecture is same as the single-channel multi-view network. From the comparative results in Figure 15, it can be seen that better performance can be achieved by fine-tuning the network or training a network. The EERs are decreased into 3.07% (fine-tuning network) and 2.84% (training new network). In addition, the proposed method also performed better than minutiae-based method. The EER is 8.35% using minutiae-based method. 3D fingerprint verification problem on two databases was also investigated using CMC curves which are illustrated in Figure 16 and Figure 17. From the results we can see that the proposed method offers superior performance for contactless 3D and 2D fingerprint matching than minutiae-based approaches on both two databases. Table 2 illustrates the EER values and rank-one accuracy of different approaches on contactless 3D/2D fingerprint matching.

In order to evaluate the effectiveness of the proposed approach on partial 3D fingerprint identification, we followed the protocol described in Section 5.2 and performed a set of experiments using proposed method, conventional minutiae-based method and fine-tuning *Resnet* [48] for comparison. The ROC curves for partial 3D fingerprint identification using database A are illustrated in Figure 18. The proposed method also outperforms fine-tuning *Resnet* approach and minutiae-based method. The EER is reduced from 17.33% (minutiae-based) and 9.58% (fine-tuning *Resnet*) to 7.27% (proposed).

In addition, the statistical significance test (*p-value* and *z-statistic*) [53, 54] of the area under the ROC curve was also performed to ascertain the significance of the performance improvement using proposed approaches. For each two ROC curves, the significance level (*p-value*) indicates a significance difference between two AUCs (area under curves) when the p value is smaller than 0.05 ($p\text{-}value < 0.05$). *z-statistic* was also used to indicate the difference between the two AUCs. The significance of the difference between the areas under two ROC curves with 95% confidence interval are illustrated in Table 3, where the ROC curves are generated using proposed method and the state-of-the-art minutiae based method. For partial 3D fingerprint identification, we also compared proposed method with fine-tuning *Resnet* approach. From the results, we can see that the improvements using proposed method are statistically significant.

5.4. Space and Time Complexity Analysis

The proposed implementation was run on an Intel i5-2500 3.3GHz CPU with NVIDIA 980Ti GPU. It requires about 3 hours to train the top-net and about 6 hours for training each side-net. During the test phase, the deep feature representations are generated by forwarding operations and the matching scores are generated by simply computing the Euclidean distance between each two feature representations. As compared with minutiae-based approaches, which need relatively complicated feature extraction and time-consuming minutiae alignment processes, the proposed method offers lower time complexity for feature extraction and matching. In addition, the 384-dimension vectors extracted from proposed network, to represent each of the 3D fingerprint deep feature, are also acceptable and suitable for storage. As a result, the proposed multi-view 3D fingerprint recognition system can also meet expectation of the large scale deployment. Table 4 illustrates the comparison of matching complexity and computation time, for the feature extraction and matching using different approaches.

6. CONCLUSIONS AND FURTHER WORK

This paper proposes an end-to-end CNN-based model to learn discriminative contactless 3D fingerprint feature representation for the contactless 3D fingerprint recognition and partial 3D fingerprint identification. A pre-trained fully convolutional network is fine-tuned to

Table 4: Comparison of time and space complexity of different 3D fingerprint recognition methods

Approaches	**Matching complexity**	**Feature extraction**	**Matching**	**Template size**
2D minutiae	2D minutiae alignment (High)	$1.69s$	$1.16s$	$329 - d$ vector
3D minutiae	3D minutiae alignment (High)	$2.26s$	$1.43s$	$483 - d$ vector
Proposed method	**Euclidean distance (Low)**	**0.01s**	$\mathbf{9.2 * 10^{-5}s}$	**384 − d vector**

obtain more accurate and robust segmentation of 3D fingerprint. Then 3D fingerprint feature representations, which contain more discriminative fingerprint texture details, shape and depth information, are learned from proposed two-channel multi-view CNN. Contactless partial 3D fingerprint identification is also introduced and investigated using proposed model in this work. The comparative results illustrated in Section 5 indicate that the proposed method performs significantly better than the minutiae-based approaches and other popular CNN-based approach for contactless 3D fingerprint recognition and partial 3D fingerprint identification on two publicly available databases. Superior recognition accuracy, low time complexity and low template storage requirement for the proposed method make it attractive for large scale deployment.

Although the proposed multi-view CNN-based approach achieves superior performance for partial 3D fingerprint and contactless fingerprint recognition, the identification accuracy can still be improved by incorporating larger training set, better architecture and training strategies including translation-based data augmentation. Performance evaluation on large scale public 3D fingerprint database, and the development of large scale partial 3D fingerprint database acquired under more realistic imaging environment, are highly desirable and are part of our further work.

References

[1] A. K. Jain, A. Kumar, Biometric recognition: an overview, in: Second Generation Biometrics: The Ethical, Legal and Social Context, Springer, 2012, pp. 49–79.

[2] D. Maltoni, D. Maio, A. Jain, S. Prabhakar, Handbook of fingerprint recognition, Springer Science & Business Media, 2009.

[3] C. Watson, M. Garris, E. Tabassi, C. Wilson, R. McCabe, S. Janet, K. Ko, Nist biometric image software (2011).

[4] A. Kumar, Contactless 3D Fingerprint Identification, Springer, 2018.

[5] C. Lee, S. Lee, J. Kim, A study of touchless fingerprint recognition system, in: Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), Springer, 2006, pp. 358–365.

[6] Y. Chen, G. Parziale, E. Diaz-Santana, A. K. Jain, 3d touchless fingerprints: compatibility with legacy rolled images, in: Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the, IEEE, 2006, pp. 1–6.

[7] R. D. Labati, A. Genovese, V. Piuri, F. Scotti, Touchless fingerprint biometrics: a survey on 2d and 3d technologies, Journal of Internet Technology 15 (3) (2014) 325–332.

[8] C. Lin, A. Kumar, Tetrahedron based fast 3d fingerprint identification using colored leds illumination, Pattern Analysis and Machine Intelligence, IEEE Transactions on (2017) 1–10.

[9] A. Kumar, C. Kwong, Towards contactless, low-cost and accurate 3d fingerprint identification, Pattern Analysis and Machine Intelligence, IEEE Transactions on 37 (3) (2015) 681–696.

[10] G. Parziale, E. Diaz-Santana, R. Hauke, The surround imagertm: A multi-camera touchless device to acquire 3d rolled-equivalent fingerprints, in: International Conference on Biometrics, Springer, 2006, pp. 244–250.

[11] Y. Wang, L. G. Hassebrook, D. L. Lau, Data acquisition and processing of 3-d fingerprints, Information Forensics and Security, IEEE Transactions on 5 (4) (2010) 750–760.

[12] Y. Wang, J. Hu, Global ridge orientation modeling for partial fingerprint identification, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (1) (2011) 72–87.

[13] T.-Y. Jea, V. S. Chavan, V. Govindaraju, J. K. Schneider, Security and matching of partial fingerprint recognition systems, in: Proceedings of SPIE, Vol. 5404, 2004, pp. 39–50.

[14] Q. Zhao, D. Zhang, L. Zhang, N. Luo, High resolution partial fingerprint alignment using poreCvalley descriptors, Pattern Recognition 43 (3) (2010) 1050–1061.

[15] N. van Noord, E. Postma, Learning scale-variant and scale-invariant features for deep image classification, Pattern Recognition 61 (1) (2017) 583–592.

[16] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

[17] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[18] S. Zagoruyko, N. Komodakis, Learning to compare image patches via convolutional neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 4353–4361.

[19] C. Shi, C. M. Pun, Superpixel-based 3D deep neural networks for hyperspectral image classification, Pattern Recognition 74 (1) (2018) 600–616.

[20] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, J. Xiao, 3d shapenets: A deep representation for volumetric shapes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1912–1920.

[21] Y. Sun, X. Wang, X. Tang, Hybrid Deep Learning for Face Verification, Pattern Analysis and Machine Intelligence, IEEE Transactions on 38 (10) (2016) 1997–1997.

[22] Y. Sun, Y. Chen, X. Wang, X. Tang, Deep learning face representation by joint identification-verification, in: Advances in neural information processing systems, 2014, pp. 1988–1996.

[23] M. Troy, L. Hassebrook, V. Yalla, R. Daley, Non-contact, 3d fingerprint scanner using structured light illumination, in: Proc. of SPIE Vol, Vol. 7932, 2011, pp. 79320C–1.

[24] W. Zhou, J. Hu, I. Petersen, S. Wang, M. Bennamoun, A benchmark 3d fingerprint database, in: Fuzzy Systems and Knowledge Discovery (FSKD), 2014 11th International Conference on, IEEE, 2014, pp. 935–940.

[25] F. Liu, D. Zhang, L. Shen, Study on novel curvature features for 3d fingerprint recognition, Neurocomputing 168 (2015) 599–608.

[26] M. K. Johnson, E. H. Adelson, Retrographic sensing for the measurement of surface texture and shape, in: Computer Vision and Pattern Recognition, 2009. IEEE Conference on, pp. 1070–1077.

[27] M. K. Johnson, F. Cole, A. Raj, E. H. Adelson, Microgeometry capture using an elastomeric sensor, in: ACM Transactions on Graphics (TOG), Vol. 30, ACM, 2011, p. 46.

[28] R. G. Maev, E. Y. Bakulin, E. Y. Maeva, F. Severin, High resolution ultrasonic method for 3d fingerprint representation in biometrics, in: Acoustical Imaging, Springer, 2008, pp. 279–285.

[29] A. Savoia, G. Caliano, A. Iulat, C. Longo, A. Caronti, R. Carotenuto, M. Pappalardo, Design and fabrication of a cmut probe for ultrasound imaging of fingerprints, in: Ultrasonics Symposium (IUS), IEEE, 2010, pp. 1877–1880.

[30] N. Lamberti, G. Caliano, A. Iula, A. S. Savoia, A high frequency cmut probe for ultrasound imaging of fingerprints, Sensors and Actuators A: Physical 172 (2) (2011) 561–569.

[31] A. Iula, A. Savoia, G. Caliano, Capacitive micro-fabricated ultrasonic transducers for biometric applications, Microelectronic Engineering 88 (8) (2011) 2278–2280.

[32] TBS, `http://www.tbs-biometrics.com/`, accessed Oct, 2012.

[33] FlashScan, `http://www.FlashScan3D.com/`, accessed Oct, 2012.

[34] F. Liu, D. Zhang, 3d fingerprint reconstruction system using feature correspondences and prior estimated finger model, Pattern Recognition 47 (1) (2014) 178–193.

[35] T.-Y. Jea, V. Govindaraju, A minutia-based partial fingerprint recognition system, Pattern Recognition 38 (10) (2005) 1672–1684.

[36] O. Zanganeh, B. Srinivasan, N. Bhattacharjee, Partial fingerprint matching through region-based similarity, in: Digital lmage Computing: Techniques and Applications (DlCTA), International Conference on, IEEE, 2014, pp. 1–8.

[37] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, A. K. Jain, Fvc2002: Second fingerprint verification competition, in: Pattern recognition, 2002. Proceedings. 16th international conference on, Vol. 3, IEEE, 2002, pp. 811–814.

[38] C. R. Qi, H. Su, M. Nießner, A. Dai, M. Yan, L. J. Guibas, Volumetric and multi-view cnns for object classification on 3d data, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 5648–5656.

[39] H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller, Multi-view convolutional neural networks for 3d shape recognition, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 945–953.

[40] Y. Fang, J. Xie, G. Dai, M. Wang, F. Zhu, T. Xu, E. Wong, 3d deep shape descriptor, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2319–2328.

[41] K. Guo, D. Zou, X. Chen, 3d mesh labeling via deep convolutional neural networks, ACM Transactions on Graphics (TOG) 35 (1) (2015) 3.

[42] D. Kim, M. Hernandez, J. Choi, G. Medioni, Deep 3d face identification, arXiv preprint arXiv:1703.10714.

[43] W. AbdAlmageed, Y. Wu, S. Rawls, S. Harel, T. Hassner, I. Masi, J. Choi, J. Lekust, J. Kim, P. Natarajan, et al., Face recognition using deep multi-pose representations, in: Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on, IEEE, 2016, pp. 1–9.

[44] A. Belyaev, Mesh smoothing and enhancing. curvature estimation, Saarbrucken.

[45] L. Hong, A. Jain, Fingerprint enhancement, in: Automatic Fingerprint Recognition Systems, Springer, 2004, pp. 127–143.

[46] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, Pattern Analysis and Machine Intelligence, IEEE Transactions on 39 (4) (2017) 640–651.

[47] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, p. 1106–1114.

[48] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[49] S. Chopra, R. Hadsell, Y. LeCun, Learning a similarity metric discriminatively, with application to face verification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, IEEE, 2005, pp. 539–546.

[50] B. Kumar, G. Carneiro, I. Reid, et al., Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 5385–5394.

[51] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: Convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM international conference on Multimedia, ACM, 2014, pp. 675–678.

[52] Weblink for downloading *PolyU Contactless to Contact-based Fingerprint Database*, `http://www.comp.polyu.edu.hk/~csajaykr/fingerprint.htm/`.

[53] E. R. DeLong, D. M. DeLong, D. L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, Biometrics (1988) 837–845.

[54] E. B. Sudderth, A. Torralba, W. T. Freeman, A. S. Willsky, Describing visual scenes using transformed objects and parts, International Journal of Computer Vision 77 (1-3) (2008) 291–330.