This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2022.3233898

Monocular 3D Fingerprint Reconstruction and Unwarping

Zhe Cui, Jianjiang Feng, *Member, IEEE,* Jie Zhou, *Senior Member, IEEE*

Abstract—Compared with contact-based fingerprint acquisition techniques, contactless acquisition has the advantages of less skin distortion, more complete fingerprint area, and hygienic acquisition. However, perspective distortion is a challenge in contactless fingerprint recognition, which changes the ridge frequency and relative minutiae location, and thus degrades the recognition accuracy. We propose a learning-based shape-from-texture algorithm to reconstruct a 3D finger shape from a single image and unwarp the raw image to suppress the perspective distortion. Our experimental results for 3D reconstruction on contactless fingerprint databases show that the proposed method has high 3D reconstruction accuracy. Experimental results for contactless-to-contactless and contactless-to-contact-based fingerprint matching indicate that the proposed method can improve the matching accuracy.

Index Terms—Contactless fingerprint, 3D reconstruction, shape from texture, unwarping.

1 Introduction

Traditional fingerprint acquisition relies on pressing fingers against fingerprint acquisition devices to get ridge-valley images. However, the collected fingerprints often have distortion issues owing to skin deformation, which degrades the fingerprint recognition accuracy [1]. In comparison, contactless fingerprint acquisition system provides a touch-free imaging approach that does not have skin distortion, can obtain a fingerprint with a larger area, and is more hygienic [2].

Early researchers build camera systems to capture contactless fingerprints, then go through conventional fingerprint recognition algorithms [3] [4] [5] [6]. However, because contactless fingerprints are directly captured by cameras, the illumination condition, camera resolution, focus clarity, and other imaging problems result in contactless fingerprints of poorer quality than contact-based fingerprints [7] [8]. To solve quality issues, several deep learning approaches [9] [10] [11] have been introduced recently, and have gained promising results in dealing with various conditions of fingerprint quality.

Another major challenge of contactless fingerprints is that they typically have perspective distortion resulting from camera imaging. The ridge orientation, frequency, and minutiae location are changed while projecting a 3D finger onto a 2D plane, leading to challenges for traditional ridge enhancement, minutiae extraction, and fingerprint matching algorithms [12]. As Fig. 1 shows, perspective distortion degrades the matching performances of both contactless-to-contact-based and contactless-to-contactless matching and needs to be addressed.

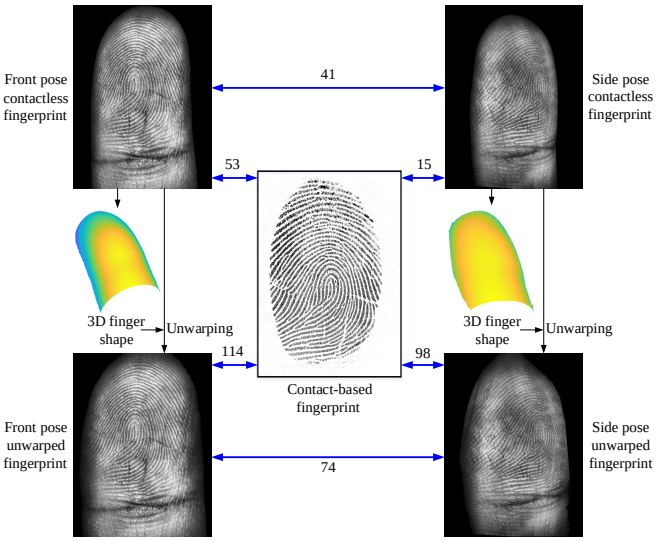


Fig. 1. Given an original contactless fingerprint, the proposed algorithm reconstructs its 3D finger shape and then unwarps it. Using the unwarped images, matching scores of contactless-to-contact-based and contactless-to-contactless matchings are improved significantly (53→114, 15→98, 41→74). The numbers on lines are matching scores by a commercial fingerprint matcher, VeriFinger SDK 12.0 [13].

To solve perspective distortion issues, some researchers have considered finger poses [14] [11]. This method estimates the finger rotation angle from a contactless fingerprint, and then rotates the finger model to the same pose to remove pose variations. The finger model used is either a 2D plane [14] or a simple 3D parametric model [11]. Although such a process can improve the matching accuracy, perspective distortion is not solved because the rectified image is still a projection from the 3D model to the 2D plane.

Another approach [15] considers perspective distortion as a problem similar to skin distortion [16], and computes a thin-plate spline transformation model to undistort con-

Zhe Cui is with the School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China, and also with the Beijing Key Laboratory of Network System and Network Culture, Beijing 100876, China.

Jianjiang Feng and Jie Zhou are with Department of Automation, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China.

e-mail: cuizhe@bupt.edu.cn, jfeng@tsinghua.edu.cn, jzhou@tsinghua.edu.cn.

tactless fingerprints. However, the unwarping process treats contactless fingerprints as 2D images that lack 3D geometric constraints.

Although significant progress has been made in contactless fingerprint matching, the problem of perspective distortion has not yet been resolved. To the best of our knowledge, there is currently no 3D fingerprint reconstruction algorithm that can directly estimate a finger shape from a single contactless fingerprint image. Only by obtaining a 3D model of the finger surface can we explicitly model and correct the perspective distortion.

The ridge-valley pattern of a human finger provides a natural texture clue for inferring a 3D shape from the ridge pattern. However, as the fingerprint ridge frequency is not globally fixed and the ridges in contactless fingerprint images are unclear, especially in the finger border region, conventional shape-from-texture methods [17] cannot achieve good 3D reconstruction results.

This paper proposes a learning-based 3D reconstruction and unwarping method for contactless fingerprints. To the best of our knowledge, this is the first monocular 3D fingerprint reconstruction algorithm for contactless fingerprints that relies on a single fingerprint image. Our method infers a 3D finger shape using fingerprint features learned from the training samples.

Fig. 1 shows the impact of our algorithm on fingerprint matching performance. By reconstructing a 3D finger shape from a single contactless fingerprint and then unwarping the contactless fingerprint to suppress perspective distortion, the unwarped fingerprints have higher matching scores. Thus, the matching performance of contactless fingerprints is improved.

We evaluate our method on three contactless fingerprint databases: PolyU Contactless 3D Fingerprint Database [18], UNSW 2D/3D Fingerprint Database [19], and PolyU Contactless 2D to Contact-based 2D Fingerprint Database [20] for evaluating the 3D reconstruction performance and the contribution of our algorithm to the final matching performance. The reconstruction results show that the proposed reconstruction method achieves state-of-the-art reconstruction accuracy. The matching experiments showed that the proposed method led to the largest matching performance improvement compared with the comparison methods.

In summary, the main contributions of our study are as follows.

- We proposed a 3D fingerprint reconstruction algorithm that can estimate the 3D finger shape from a single contactless fingerprint image.
- We proposed an unwarping algorithm to utilize 3D shape information to unwarp a contactless fingerprint to a flattened fingerprint. The unwarping process decreases perspective distortion, and thus increases the matching performances.
- We conducted extensive experiments to examine the proposed reconstruction and unwarping algorithms on commonly-used contactless fingerprint databases.

The remainder of this paper is organized as follows. Section 2 reviews the recent work on contactless fingerprint. Section 3 introduces the proposed 3D reconstruction and unwarping method. The experimental results are presented in Section 4. Some limitations and discussion of this study are presented in Section 5. Finally, Section 6 concludes the study.

2 Related Work

Contactless fingerprint recognition methods can generally be categorized as 2D-based and 3D-based methods, and 2D-based methods can be further divided according to whether or not 3D finger information is utilized. This section reviews recent studies in these categories.

2.1 Contactless 2D Fingerprint Recognition without 3D Finger Information

In early contactless fingerprint studies, researchers directly use traditional contact-based fingerprint recognition methods [3] [4] [5] to process contactless fingerprint images. Specific devices or general purpose cameras have been used to capture contactless fingerprints [21] [22] [23].

Despite these accomplishments, the accuracy of contactless fingerprint recognition is still affected by the image quality and perspective distortion problems. For image quality issues, some methods utilize high-resolution cameras to improve the image quality [22] [23] [24]. A recent advance is the introduction of deep learning into contactless fingerprint processing [25] [11] and matching [26] [20] [9] [10], which has shown promising results in improving recognition accuracy.

Regarding perspective distortion, pose constraints are commonly considered to achieve better results [22] [14] [11]. The finger pose is estimated from a raw contactless fingerprint image, and the finger is rotated to the same frontal perspective angle in 3D space according to the estimated finger pose. Although normalizing the finger pose is beneficial for fingerprint matching, the corrected fingerprint remains a simple 2D projection of a 3D model. Thus, perspective distortion still exists.

Some methods [15] [27] treat perspective distortion as skin distortion and rectify the distortion by estimating the distortion field [16]. However, this unwarping method is still 2D-based, and does not explicitly consider the 3D finger shape. Additional knowledge of the 2D distortion field is required to train such a correction model, which further limits this approach. In addition, these methods are expressly designed for matching contactless fingerprints with contact-based ones, without the consideration of contactless-to-contactless matching, or the pose variation problems in contactless fingerprint matching.

In general, although 2D-based unwarping can decrease the distortion to a certain level, the 2D-based model does not capture the fact that contactless fingerprints are generated from real 3D fingers. Therefore, recent contactless fingerprint studies have introduced 3D finger information either by 3D acquisition hardware or by 3D reconstruction algorithm.

2.2 3D Fingerprint Acquisition and Recognition

Researchers have proposed multiple approaches for 3D fingerprint reconstruction. A typical method is the multiview 3D reconstruction system [18] [28] [21], in which multiple

cameras with known intrinsic and external parameters are required to obtain a 3D shape. Other 3D reconstruction systems, such as shape from shading [29] [30], shape from focus [31], require multiple shots of a finger to compute a 3D fingerprint, and specially designed acquisition equipment is needed. Some special 3D acquisition methods, such as structured light [32], ultrasound imaging [33], and lasers [34], have also been applied to capture 3D fingerprints.

Although these existing contactless 3D fingerprint reconstruction methods can obtain 3D finger shapes with relatively high precision, they usually require bulky and expensive acquisition equipment and thus have limited usage in real-world applications.

Several researchers [29] [35] [30] have studied feature extraction and matching methods specific to 3D fingerprints. However, it may be difficult for 3D fingerprint recognition systems to benefit from the existing well-optimized 2D fingerprint recognition technologies. A more practical solution is to utilize 3D finger information in inexpensive and common monocular contactless fingerprint acquisition devices.

2.3 Contactless 2D Fingerprint Recognition with 3D Finger Information

A few contactless 2D fingerprint recognition studies [11] [36] have considered the perspective distortion problem in the 3D space. They estimate a 3D finger shape model from a contactless fingerprint and then conduct 3D-to-2D unwarping to decrease the perspective distortion issue. However, the 3D reconstruction models used by these unwarping methods are too simple to obtain an accurate 3D finger shape, thus limiting their performance of unwarping methods. The reconstruction accuracy of these methods is also unknown because of the lack of 3D ground truth.

Using [36] as an example, it uses an ellipse model whose parameters (major and minor axes) are estimated from the finger contour. This idea is similar to the shape from the silhouette method [21], but the latter usually requires 3-5 images from different angles to reconstruct a 3D shape with only limited precision. Compared with the shape from silhouette method [21], the reconstruction accuracy in [36] is further limited, which affects the subsequent unwarping step. Reliable 3D fingerprint reconstruction and unwarping technology for common monocular contactless scenarios requires further research and is the aim of this study.

2.4 Summary

The contactless 2D fingerprint recognition methods based on common cameras have been studied for many years. However, one of its major challenges, the perspective distortion, has not yet been established. Many studies (recent studies are summarized in table 1) are based on the simple assumptions of parametric 3D finger model, thus the corresponding reconstruction and unwarping methods are not able to remove perspective distortion. Specialized 3D acquisition equipments can obtain 3D data with high precision, but those equipments are too bulky and expensive to deploy on a large scale.

Therefore, a 3D fingerprint reconstruction and unwarping method is proposed in this study for robust and low-cost monocular contactless fingerprint processing.

3 Fingerprint Reconstruction and Unwarping

Fig. 2 illustrates the basic idea of this paper using a simplified model. A hemisphere has an equally distributed ridge pattern along the radial direction. However, ridges in its 2D projection viewed from the top are not evenly distributed owing to perspective distortion. The center area is less distorted because this area is perpendicular to the observer. The edge area has a reducing period and significant distortion because the tilt angle there is much larger. Therefore, changes in the ridge period reflect the surface tilt angle towards the observer.

Similar to this simplified model, the local ridge period of the contactless fingerprint clearly changes because of the perspective distortion. The ridge period of the border region of the contactless fingerprint is smaller than that of the center region because of the larger slant angle. The angle of the local surface can be inferred through changes in the ridge period, which is a typical problem of shape from texture [17].

Our method consists of a 3D surface reconstruction from an input image and subsequent image unwarping based on the reconstruction result. The whole algorithm mainly contains four steps: (1) contactless fingerprint preprocessing, (2) finger surface gradient estimation, (3) 3D surface reconstruction from the surface gradient, and (4) unwarping contactless fingerprint according to the reconstructed 3D shape. The flowchart of the proposed method is shown in Fig. 3.

3.1 Image Preprocessing

Raw contactless fingerprint images need to be preprocessed to reduce variations in the input images. Following the processing method in [15], the preprocessing step alters the images in three aspects: contrast, scale, and pose, which correspond to the three sub-modules illustrated in Fig. 4 as enhancement, scaling, and rotation.

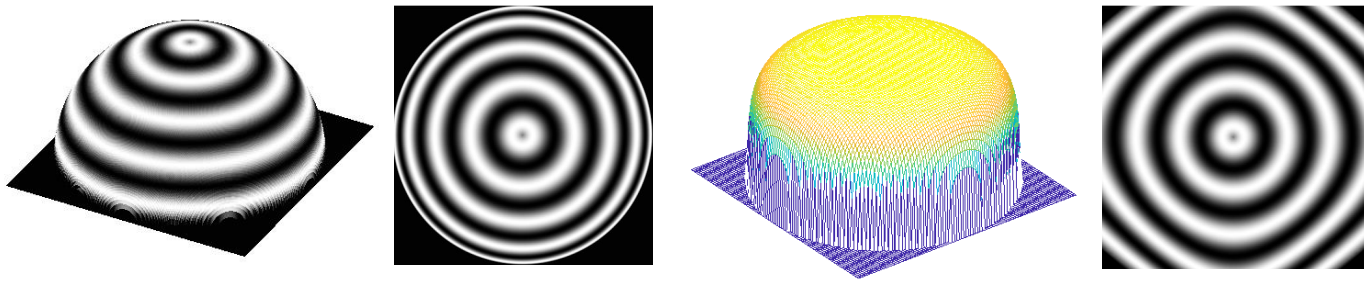
First, we enhance the raw fingerprint image to improve its ridge-valley contrast. Because of the different lighting conditions, image brightness may differ among different images. We first segmented the fingerprint ROI by thresholding the Laplacian edges (>10) to determine the finger contours. After segmentation, we use the contrast-limited adaptive histogram equalization (CLAHE) [38] to enhance the ridge-valley contrast. The grid size of CLAHE is empirically set to 60 pixels.

The next step is scaling. As monocular depth prediction has scale ambiguity issues [39], the scales of contactless images need to be settled. We use the average ridge period to determine the scale parameter. This is based on the observation that the central region of a contactless fingerprint barely has distortion, thus reflecting the actual scale of the original fingerprint. The mean ridge period of the human fingerprint is a constant number (approximately 0.5 mm [40]). Therefore, if we adjust the mean ridge period of the fingerprint image to a constant value, the image-to-real ratio of the fingerprint size can be determined. After scaling, the ratio between the distance on the fingerprint image and the corresponding distance on the real finger was the same for all the fingerprint images.

TABLE 1
Recent 2D contactless fingerprint recognition researches

Authors	Year	3D Reconstruction Method	Perspective Distortion Method	Testing Database	CL-CL	CL-CB
Lin and Kumar [20]	2018	–	2D Unwarping	B	–	✓
Dabouei *et al.* [15]	2019	–	2D Unwarping	A, C	–	✓
Tan and Kumar [11]	2020	Ellipse model	Re-projection at the same 3D pose	A, D	✓	–
Söllinger and Uhl [36]	2021	Ellipse model	3D Unwarping	Not public	–	✓
Grosz *et al.* [27]	2022	–	2D Unwarping	A, B, C	–	✓
Proposed Method	2022	Deep network based surface gradient estimation	3D Unwarping	A, B, E	✓	✓

Database A: UNSW 2D/3D Fingerprint Database [19]
Database B: PolyU Contactless 2D to Contact-Based 2D Fingerprint Database [20]
Database C: ManTech Database [37]
Database D: PolyU 3D Fingerprint Images Database [30]
Database E: PolyU Contactless 3D Fingerprint Database [18]
CL-CL: Contactless-to-contactless fingerprint matching
CL-CB: Contactless-to-contact-based fingerprint matching



(a) A hemisphere with equally distributed pattern (b) Top view image (c) Estimated ridge period (d) Unwarped image

Fig. 2. The basic methodology of our method. (a) is a hemisphere with equally distributed ridges along the radius. (b) is its projected image into a plane, and local ridge periods are changed due to perspective distortion. (c) is the estimated ridge period map from (b) using the short-time Fourier transform, where the change of period corresponds to a tilted surface. (d) is the unwarped image of (b) using an estimated 3D model, where the perspective distortion is reduced.

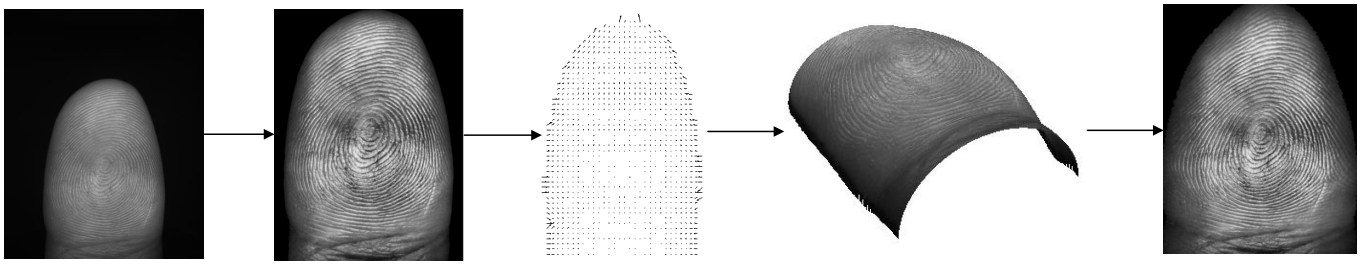


Fig. 3. Flow diagram of the proposed monocular 3D fingerprint reconstruction and unwarping algorithm.

The detailed scaling steps are illustrated in Fig. 4. We determine the geometric center as the ROI center point for each contactless fingerprint. Then, a circle is drawn as the central region to cover approximately 20 percent of the total ROI region. The average ridge period within the central area is computed using a traditional algorithm [41]. The fingerprint image is resized to change the mean ridge period to 10 pixels. By scaling images, the scale ambiguity issue is settled. We scale all ridge periods to 10 pixels because most flat fingerprints have a mean ridge period of 10 pixels under standard 500 ppi. Resizing to 10 pixels is beneficial for further matching contactless fingerprints with contact-based fingerprints.

Finally, the pose variation is handled by rotation. As shown in Fig. 5, the finger may rotate in three directions: *Pitch* around the x axis, *Roll* around the y axis, and *Yaw* around the z axis. The pose correction step mainly deals with *Yaw* angle, as *Yaw* corresponds to a rotation in the 2D image plane and can be directly estimated from contactless images. *Pitch* and *Roll* angle variations are handled by data augmentation for network training.

Yaw angle is obtained by estimating the angle (or slope) of the centerline of the finger region. For each row in the contactless fingerprint, we find its left and right contours

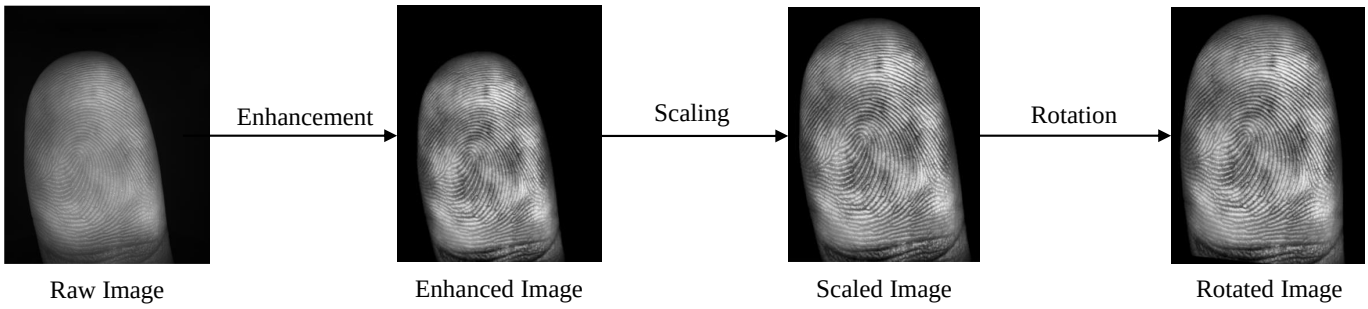


Fig. 4. Key steps for image preprocessing, including enhancement, scaling, and rotation steps.

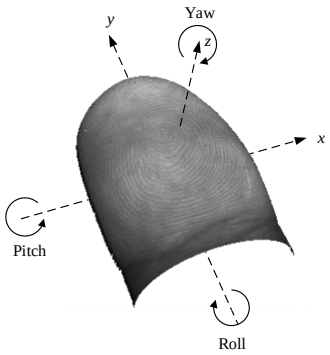


Fig. 5. Three rotation angles of finger pose.

and locate its center point. We use the center points of all horizontal lines to compute a centerline by least-square fitting. Then, the angle between the centerline and vertical direction is the *Yaw* angle. A finger is rotated to make the centerline vertical.

3.2 Surface Gradient Estimation

This step estimates the surface gradients for further 3D reconstruction using the convolutional neural networks. We predict the gradient instead of the depth for three reasons:

- An evident pattern for contactless fingerprints is the change in the ridge period from the center to the edge. This change directly corresponds to the surface title degree or the surface normal direction, which is equivalent to the depth gradient.
- The direct estimation of depth has absolute value uncertainty. The absolute depth could not be determined from the images. However, the gradient reflects the relative changes. Thus, estimating the gradient avoids this problem.
- The following unwarping algorithm is required to calculate the geodesic distance on the 3D surface, which requires a gradient map.

3.2.1 Network Structure

Fig. 6 illustrates the structure of the gradient estimation network. The network uses the preprocessed image and segmentation mask as inputs, and outputs the orientation, period, and gradient maps. The network is inspired by FingerNet [42]: the first part is image normalization to scale image intensities; the second part is the orientation and period feature network, which consists of 3 convolution and max-pooling blocks to extract features of 1/8 input size; the third part is to regress orientation and period maps separately; the final part is to regress gradients from orientation and period features.

Fig. 7 shows the details of the network structure. The first part, image normalization, is completed at the end of image preprocessing and is not included in the network structure. The second part is a VGG-like [43] feature extraction block consisting of six convolution layers and three pooling layers. The extracted features are simultaneously sent to the orientation and period blocks. Both blocks contain three convolution layers and a final regression layer. The features from the orientation and period blocks are concatenated to the gradient regression block. The activation function is ReLU for each layer, except for the three output layers. The orientation output uses softmax activation as it is a classification model with 180 categories. The period and gradient outputs use linear activation during the regression.

Note that an extra mask input is added to activate the output orientation, period, and gradient, because only the finger region is processed. The output orientation, period, and gradient maps have 1/8 size of the inputs during training, and the final outputs are interpolated to the original size during testing.

The output orientation is an N-dimensional vector at each point (x, y) representing the probability distribution, where N is the number of discrete angles of orientation. Here N is set to 180 in our experiment. The final orientation is a probability-weighted estimation [44]:

$$\begin{aligned} d_{\cos}(x,y) &= \frac{1}{N}\sum_{i=0}^{N-1} p(i,x,y)\cos(\frac{360}{N}i), \\ d_{\sin}(x,y) &= \frac{1}{N}\sum_{i=0}^{N-1} p(i,x,y)\sin(\frac{360}{N}i), \\ \mathrm{Ori}(x,y) &= \frac{1}{2}\mathrm{atan2}\left(d_{\sin}(x,y), d_{\cos}(x,y)\right). \end{aligned} \tag{1}$$

The output period is scalar, while the gradient is 2-dimensional: $g_x(x, y)$ and $g_y(x, y)$ represent gradients along the x and y directions, respectively.

3.2.2 Loss Definition

There are three losses corresponding to the three outputs: orientation loss $\mathcal{L}_{\mathrm{Ori}}$, period loss $\mathcal{L}_{\mathrm{Ped}}$, and gradient loss $\mathcal{L}_{\mathrm{Grad}}$.

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2022.3233898

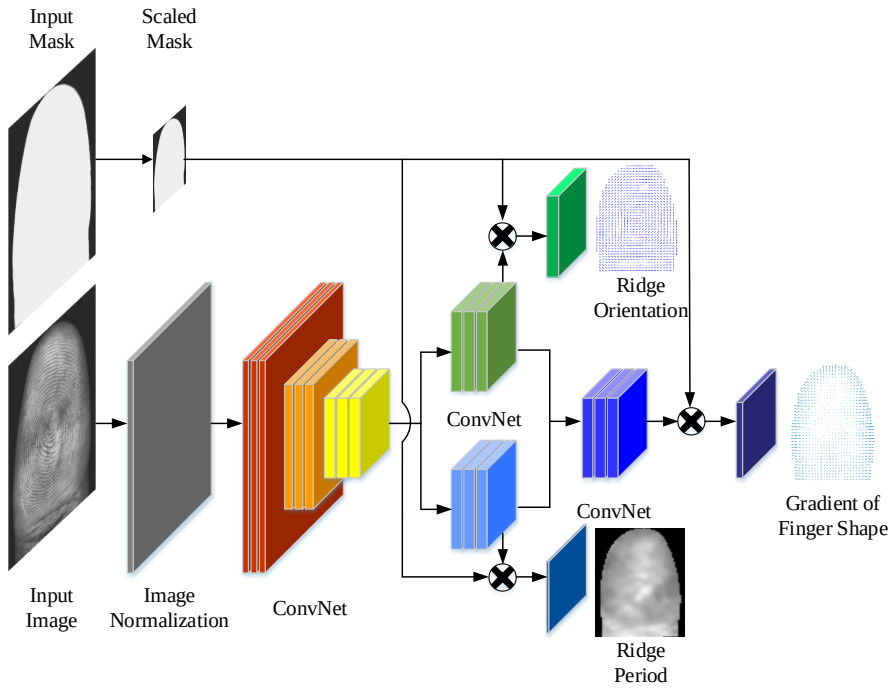


Fig. 6. Structure of gradient estimation network.

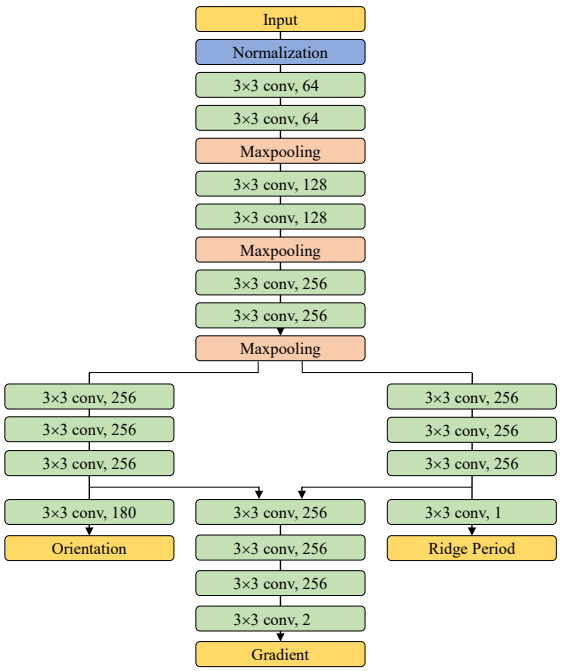


Fig. 7. Details of network structure.

The orientation loss $\mathcal{L}_{\text{Ori}}$ is composed of cross-entropy loss and coherence loss [44]:

$$\begin{aligned}\mathcal{L}_{\text{Ori}} = &-\frac{1}{|M|}\sum_{M}\sum_{i=0}^{N-1}(o^*(i,x,y)\log(o(i,x,y)) \\ &+(1-o^*(i,x,y))\log(1-o(i,x,y))) \\ &+\alpha\frac{|M|}{\sum_M \frac{\|(\overline{d}_{\cos}(x,y),\overline{d}_{\sin}(x,y))\|}{\overline{\text{Ori}}(x,y)}}-1.\end{aligned} \tag{2}$$

$\overline{d}_{\cos}(x,y)$, $\overline{d}_{\sin}(x,y)$, and $\overline{\text{Ori}}(x,y)$ are the smoothing results of $d_{\cos}(x,y)$, $d_{\sin}(x,y)$, and $d(x,y)$ using a 3×3 all one kernel. M is the input mask as activation, and $\sum_M/|M|$ means averaging within the mask region. $o(i,x,y)$ and $o^*(i,x,y)$ are estimation results and ground-truth respectively. α is the weight of coherence loss.

Period loss $\mathcal{L}_{\text{Ped}}$ is composed of regression loss and smoothing loss:

$$\mathcal{L}_{\text{Ped}} = -\frac{1}{|M|}\sum_{M}(p(x,y)-p^*(x,y))^2+\beta\|\nabla p(x,y)\|^2. \tag{3}$$

$p(x,y)$ and $p^*(x,y)$ are estimated period and ground-truth period. β is the weight of smoothing loss.

Gradient loss $\mathcal{L}_{\text{Grad}}$ is:

$$\begin{aligned}\mathcal{L}_{\text{Grad}} = &-\frac{1}{\sum_M w(x,y)}\sum_{M} w(x,y)\|g(x,y)-g^*(x,y)\|^2 \\ &+\gamma\|\nabla g(x,y)\|^2.\end{aligned} \tag{4}$$

$g(x,y)$ and $g^*(x,y)$ are estimated gradient and ground-truth gradient. γ is the weight of smoothing loss. The gradient loss is similar to period loss, but with weight $w(x,y)$:

$$w(x,y) = \exp\left(-\frac{\|g^*(x,y)\|}{\sigma}\right). \tag{5}$$

σ is the weight kernel.

Because gradient data usually have larger values in the marginal region and smaller values in the central region, giving all pixels the same weight using mean squared error in training leads to more attention being paid to marginal area data than on center area data. Therefore, we add weight here to balance the marginal and central areas. The relative errors can be the same by setting larger values with smaller weights.

TABLE 2
Parameters used in loss function

Parameter	α	β	γ	σ	λ_1	λ_2	λ_3
Value	1	1	1	0.5	1	20	100

The total loss is the sum of these three losses:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{Ori}} + \lambda_2 \mathcal{L}_{\text{Ped}} + \lambda_3 \mathcal{L}_{\text{Grad}}. \tag{6}$$

λ_1, λ_2, and λ_3 are the weights of three losses.

All parameters used in the loss function are empirically determined. Table 2 lists the parameter values during the implementation.

3.2.3 Real Training Data

To train the network, as shown in Table 3, we use two contactless fingerprint databases: the PolyU contactless 3D fingerprint database [18][1], and the UNSW 2D/3D fingerprint database [19]. This section introduces the use of the PolyU 3D database.

The PolyU 3D database contains 220 contactless fingerprint images with corresponding ground-truth depth maps obtained using the structured-light technique. We preprocess the raw images using the methods described in Section 3.1, and then extract the orientation and period features using VeriFinger [13]. The orientation and period extracted using VeriFinger are used as the ground truth. Thus, they may not be the true values. This study aims to recover the depth map, and the computing orientation and period are just a reference here.

The original depth data of the PolyU 3D database are very noisy and contain ridge-valley variability, which makes gradients very unstable. Therefore, we smooth the initial depth map using the moving least square [45]. In this study, we aim to estimate a smoothed finger shape without a 3D ridge-valley structure.

For each point (x, y, z), we compute a quadric surface $f(x, y)$ using K nearest neighbors:

$$f(x, y) = a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y + a_6$$
$$\min_a \sum_{i=1}^{K} w_i (z_i - f(x_i, y_i))^2, w_i = \exp\left(-\frac{(x_i - x)^2 + (y_i - y)^2}{\epsilon^2}\right) \tag{7}$$

Solving equation (7) yields a smoothed estimation $f(x, y)$ to replace the original depth z. Subsequently, the gradients are computed as follows:

$$(g_x, g_y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right). \tag{8}$$

When resizing the images during preprocessing, the depth and gradient data are scaled simultaneously. We then crop training patches of size 512×512 from the image, mask, orientation, period, and gradient maps. For each pair of data points in the PolyU 3D database, 100 patches are extracted. Thus, 22,000 sets of training data are obtained from the PolyU 3D database. We use 60% for the training and 40% for the testing.

1. Interested readers may contact Dr. Feng Liu (feng.liu@szu.edu.cn) for the availability of this database.

3.2.4 Synthetic Training Data

Although the PolyU 3D database contains ground-truth depth data, the contactless fingerprints in this database are all of front pose, lacking pose variations. The contactless fingerprints of side pose usually have steeper shapes and more severe perspective distortion. In addition, the lighting condition of the side pose may be different from that of the front pose, which affects image quality. Therefore, side pose data are also needed to enlarge data comprehensiveness.

We use the UNSW database [19] to generate the side pose data. The contactless fingerprints in this database are captured by Surround Imager [21]. Images from three poses are captured simultaneously: left, middle, and right, and different poses are 45 degrees apart. Therefore, contactless fingerprints in this database have sufficient pose variations and can be used to train a network that is suitable for both front and side fingerprints.

Because this database has no ground-truth 3D fingerprint data, we design a shape from the silhouette algorithm (see Supplementary Material) to generate finger shape models from images of three different angles. This dataset was captured using the scanner of TBS Inc. [21], which uses a shape-from-silhouette to reconstruct the 3D finger shape. Therefore, the shape-from-silhouette algorithm is suitable for UNSW database. This algorithm is only used to obtain a 3D finger shape from all three views, which is not the monocular fingerprint reconstruction algorithm in our manuscript.

The UNSW database contains 9,000 images from 150 people $\times$ 10 fingers $\times$ 3 angles $\times$ 2 captures. As in [15], we use the last 3,000 images from the last 50 people for training the network, others for the matching experiment. We use the abovementioned reconstruction method to simultaneously obtain shape models from three images of different poses. Therefore, we get $50 \times 10 \times 3 \times 2 = 3000$ shape models. As we can get 3,000 data from UNSW database, which are much larger than the 220 data from PolyU 3D database, we only generate 5 patches from each data, and a total number of 15,000 patches of size 512×512 are extracted for fine-tuning the network.

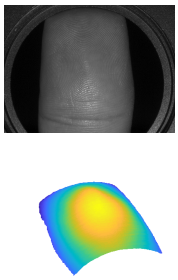
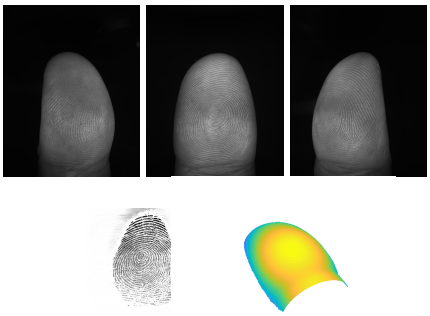
3.3 Surface Reconstruction

We use the gradient output from the abovementioned network to reconstruct the surface. Because gradients are derivatives of depths, depths can be recovered by integration. For reconstruction accuracy, we design a specific integration method (**Algorithm 1**).

The basic idea of the reconstruction algorithm is to integrate outward from the center to the periphery. This is because the gradients in the central region are small, whereas those in the peripheral region are large. Integration from the center to the periphery can reduce numerical errors. The integration starting point, or center point, is defined as the minimum gradient value.

Because the line integral $\int g_x(x, y)dx + g_y(x, y)dy$ from $(0, 0)$ to (x, y) has infinite integral paths, we choose $(0, 0) \rightarrow (x, 0) \rightarrow (x, y)$ and $(0, 0) \rightarrow (0, y) \rightarrow (x, y)$ as two paths to compute z_1 and z_2, and use their average as the final result. The trapezoidal integral is used for the numerical calculation.

TABLE 3
Contactless fingerprint databases used in this paper.

Database	PolyU Contactless 3D Fingerprint Database [18]	UNSW 2D/3D Fingerprint Database [19]	PolyU Contactless 2D to Contact-based 2D Fingerprint Database [20]
Example			
Image type	Contactless 2D fingerprints & ground-truth depth data by structured light	Contactless 2D fingerprints & contact-based 2D fingerprints & 3D finger shapes reconstructed by our shape-from-silhouette algorithm	Contactless 2D fingerprints & contact-based 2D fingerprints
Description	220 contactless fingerprints & corresponding depth data acquired by structured light	9,000 contactless fingerprints from 1500 fingers × 3 angles × 2 captures & mated 6,000 contact-based fingerprints from 1500 fingers × 4 captures	2,016 contactless fingerprints from 336 fingers × 6 captures & mated 2,016 contact-based fingerprints from 336 fingers × 6 captures
Training data	22,000 patches from 220 data, 13,200 for training network	15,000 patches from the last 3,000 contactless fingerprints, 9,000 patches for training network	–
Experiment	Reconstruction accuracy	Reconstruction accuracy & Matching accuracy	Matching accuracy
Experiment data	8,800 patches for reconstruction experiment	6,000 patches for reconstruction experiment & first 6,000 fingerprints for matching experiment	2,976 pairs of fingerprints for matching experiment

Algorithm 1 Surface Reconstruction Algorithm

Input: Gradients g_x, g_y
Output: Depth z

Find integration starting point (zero point) (c_x, c_y), which has the minimum gradients value.
Integration along $x-y$ path:
for x **do**
  $z_1(x, c_y) = \int_{c_x}^{x} g_x(x, c_y)dx$
  for y **do**
    $z_1(x, y) = \int_{c_y}^{y} g_y(x, y)dy$
Integration along $y-x$ path:
for y **do**
  $z_2(c_x, y) = \int_{c_y}^{y} g_y(c_x, y)dy$
  for x **do**
    $z_2(x, y) = \int_{c_x}^{x} g_x(x, y)dx$
$z(x, y) = \frac{z_1(x,y)+z_2(x,y)}{2}$
return z

Algorithm 1 describes the reconstruction process. The estimated gradient g_x, g_y is used in the calculation. First, we determine the integration starting point (zero point) (c_x, c_y) of the smallest gradient value, which is usually located at the center of the fingerprint.

Subsequently, a numerical integral is conducted. As mentioned above, two integral paths are used. Integral z_1 first moves along the x axis and then switches to the y axis. The integral z_2 is the opposite. The final result z is their average to improve accuracy.

3.4 Unwarping

Contactless fingerprint unwarping simulates the finger rolling action to unfold the finger surface into a flat plane. We develop an arc length-based unwarping algorithm (**Algorithm 2**). The integration method is similar to the reconstruction algorithm, but replaces the integrating gradient value with the arc length.

Algorithm 2 Contactless Fingerprint Unwarping Algorithm

Input: Gradients g_x, g_y, contactless image $I_{\text{in}}(x, y)$
Output: Unwarped image $I_{\text{out}}(x, y)$

Find integration starting point (zero point) (c_x, c_y), which has the minimum gradients value.
Compute horizontal arc length u and vertical arc length v:
for x **do**
  $u(x, y) = \int_{c_x}^{x} \sqrt{1 + g_x(c_x, y)^2}dx$
for y **do**
  $v(x, y) = \int_{c_y}^{y} \sqrt{1 + g_y(x, y)^2}dy$
Transform image I_{in} to I_{out} according to correspondences $(x, y) \rightarrow (u + c_x, v + c_y)$
return I_{out}

The basic idea of the proposed unwarping algorithm is to use the arc length as a new coordinate system. Therefore, the distance between the two points on the surface is

TABLE 4
Average time cost of each step

Step	Time cost (s)
Image Preprocessing	0.97
Surface Gradient Estimation	0.46
Surface Reconstruction	0.68
Unwarping	1.07
Overall	3.18

maintained after unwarping. We use the gradient to unwarp the image instead of using the recovered depth to further reduce the numerical error, as depth z may already contain numerical errors.

Algorithm 2 describes the steps involved in the unwarping algorithm. First, we find the integration starting point the same as in the reconstruction algorithm. The curve distance is then calculated using the basic curve length integration formula with gradients, and the arc length of each point is used as its new coordinate for image unwarping. A point at (x, y) is moved to (c_x+u, c_y+v) after transformation. Bilinear interpolation is applied to the image transform.

3.5 Implementation Details

The proposed gradient estimation network is trained using 2 Nvidia 1080Ti GPUs using the Adam [46] optimizer with default parameters (learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, decay rate = 0). The training loss function is composed of the orientation, period, and gradient loss. Detailed network and loss function descriptions are provided in Section 3.2.1 and 3.2.2.

The network is first trained on the PolyU 3D database and then fine-tuned on the UNSW database. A total of 13,200 samples from PolyU 3D and 9,000 samples from UNSW are used for training. As the network output has 1/8 size of input size, the groundtruth orientation, period, and gradient are down-sampled by 1/8 ($512 \times 512 \rightarrow 64 \times 64$) for computing losses. The training batch-size is set to 16. The network converges after 50 epochs.

For testing, the output orientation, period, and gradient are up-sampled to the original size using bilinear interpolation. Estimation errors are computed between the interpolated outputs and ground truth. 8,800 samples from PolyU 3D and 6,000 samples from UNSW are used for testing.

Table 4 lists the average time cost of each step of processing a single contactless fingerprint. The surface gradient estimation step is implemented in Python and 2 Nvidia 1080Ti GPUs, while the other steps are implemented in MATLAB and an Intel Xeon E5-2640 2.5GHz CPU.

4 Experiment

This section presents experiments on the proposed reconstruction and unwarping algorithm. Sections 4.1-4.4 introduce the datasets used in the experiments, the experiments on reconstruction accuracy, unwarping performance, and matching performance, respectively.

4.1 Datasets

As summarized in Table 3, three databases are used to evaluate the proposed method. These databases are publicly available with ground-truth 3D depth data or strong view-angle differences, which are suitable for training and testing the proposed network. In addition, these databases have been commonly used in previous studies, and we can fairly compare our method with previous studies.

PolyU 3D database [18] contains contactless fingerprints along with their ground-truth depth maps. This database is used for the reconstruction accuracy experiments. As there is only one image for each finger, matching experiments cannot be performed using this database.

UNSW database [19] contains contactless and corresponding contact-based fingerprints. The ground-truth depth maps are generated as mentioned in Section 3.2.4. Both the reconstruction and matching experiments are performed using this database.

PolyU CL2CB database [20] also contains contactless and corresponding contact-based fingerprints. However, because we could not obtain the ground-truth 3D depth maps of this dataset, only matching experiments are conducted.

4.2 3D Reconstruction Accuracy

4.2.1 Evaluation Protocols

The proposed method is quantitatively evaluated using 8,800 samples from the PolyU 3D database and 6,000 samples from the UNSW database. Tables 5 and 6 show the quantitative results of the mean errors of the estimated orientation, period, gradient, and depth for the two databases. The definitions of the four output errors are as follows.

$$\begin{aligned} e_{\text{Ori}} &= \frac{1}{|M|}\sum \min(|\text{Ori}_{\text{out}} - \text{Ori}_{\text{true}}|, \\ &\quad 180 - |\text{Ori}_{\text{out}} - \text{Ori}_{\text{true}}|), \\ e_{\text{Ped}} &= \sqrt{\frac{1}{|M|}\sum (\text{Ped}_{\text{out}} - \text{Ped}_{\text{true}})^2}, \\ e_{\text{Grad}} &= \sqrt{\frac{1}{|M|}\sum \|\text{Grad}_{\text{out}} - \text{Grad}_{\text{true}}\|_2^2}, \\ e_{\text{Dep}} &= \sqrt{\frac{1}{|M|}\sum (\text{Dep}_{\text{out}} - \text{Dep}_{\text{true}})^2}. \end{aligned} \tag{9}$$

Ori_{out}, Ped_{out}, Grad_{out}, and Dep_{out} are the output orientation, period, gradient, and depth value. Ori_{true}, Ped_{true}, $\text{Grad}_{\text{true}}$, and Dep_{true} are the ground-truth orientation, period, gradient, and depth value. $1/|M|$ indicates the average within the mask region.

The average error is the average value of Equation (9) for all test samples. The average weighted error means adding weight using Equation (5) when computing Equation (9) to balance the center and edge areas. There are several notes on the evaluation:

- The orientation and period maps are intermediate outputs used to examine whether the network properly extracts fingerprint texture features. The gradient output is the final required result, while the orientation and period outputs are just referencing here.

TABLE 5
Evaluation results on PolyU 3D database

Errors	Orientation (Degree)	Period (Pixel)	Gradient (mm/Pixel)	Depth (mm)
Average	2.6927	0.2624	0.4359	0.1802
Weighted	2.8698	0.2722	0.0035	0.1506

TABLE 6
Evaluation results on UNSW database

Errors	Orientation (Degree)	Period (Pixel)	Gradient (mm/Pixel)	Depth (mm)
All[1]	2.7056	0.3376	0.1599	0.4844
All[2]	2.5170	0.3287	0.0073	0.3540
Front Pose[1]	2.5297	0.3355	0.0679	0.3236
Front Pose[2]	2.6690	0.3471	0.0046	0.1626
Side Pose[1]	2.7935	0.3386	0.2059	0.5648
Side Pose[2]	2.4410	0.3194	0.0086	0.4512

1: Average error.
2: Weighted average error.

- The gradient and depth values from the PolyU 3D database are acquired using a structured light device. Therefore, they are sufficiently accurate as ground-truth data but contain only front pose data. Meanwhile, the depth data from the UNSW database are synthetic, as mentioned in Section 3.2.4. However, they contain both front and side pose data.
- The units of error of the four outputs are different: the orientation error is in degree; the period error is in pixel; the depth error is in mm; and the gradient error is in mm/pixel.
- The estimated depth map has scale uncertainty. The reconstructed depth from Section 3.3 is the integration result from zero point, thus the minimum value in the finger surface is always zero. Therefore, the ground-truth depth map is subtracted by a number to make the depth value of the corresponding lowest point zero. Subsequently, the depth error between them is computed.

4.2.2 Evaluation Results

Tables 5 and 6 present the evaluation results for the Poly3D and UNSW databases, respectively. From Table 5 and 6, we can see that:

- Adding weights barely causes a difference in orientation and period estimation errors. However, the gradient and depth errors are reduced after adding weights, especially for gradients. As Fig. 8 shows, most depth errors occur near the edge region, whereas the center region has a small error and is more reliable. This is also a motivation for designing our reconstruction and unwarping algorithm, in which the calculation should start from the center area to reduce error.
- A large part of the side pose fingerprints has larger gradient and depth errors than the front pose fingerprints. This is because the side pose images are usually more curved. Therefore, it is more challenging to estimate the shapes and obtain larger errors.

TABLE 7
Numerical errors of reconstruction algorithm on PolyU 3D database

Reconstruction Error	Depth (mm)
Complete Region	2.5210
Region without Margin	0.0141

To further investigate the accuracy of the proposed reconstruction algorithm, we conduct a numerical experiment. Table 7 lists the numerical errors of the reconstruction algorithm on the PolyU 3D database. Here, numerical errors refer to the use of the ground-truth gradient to reconstruct the depth and compute the differences between the ground-truth depth and the reconstructed depth. This is an evaluation of the numerical precision of our reconstruction algorithm because both the input and target are ground-truth data. The complete region in Table 7 indicates the computation of the RMSE within the entire fingerprint region, whereas the region without a margin means removing three pixels from the fingerprint border.

Clearly, from Table 7, nearly all of the numerical errors occur in the border region. This is because numerical differential and integration are used in our algorithm, and they are unstable near the border region as sudden changes in depth happen here. Removing 3-pixel wide border barely affects the final result. In our following experiments, all results are computed without the border region.

As shown in Table 7, the numerical error is relatively small ($\sim$ 0.01 mm) compared with the reconstruction error in the following experimental section ($\sim$ 0.2 mm). Therefore, the numerical error of the reconstruction algorithm is neglected in the following reconstruction experiment. The reconstruction errors mainly originate from gradient estimation errors.

Fig. 8 shows some qualitative results of reconstruction errors. Three examples of front, right, and left contactless fingerprints are displayed. We can see from the heatmaps of reconstruction errors that most errors are relatively small ($<$ 0.5 mm). Large reconstruction errors ($>$ 2 mm) happen mostly in the border region owing to numerical integration errors or foreground segmentation errors. The mean reconstruction errors are very small if we remove these border outliers, which is consistent with the evaluation results in Table 5, 6, and 7.

Fig. 9 shows an example of the reconstruction and unwarping results for different poses from the same finger. Our method can successfully process fingerprints that vary in yaw angle, which benefits from the preprocessing step that corrects fingerprint to the same pose.

4.2.3 Ablation Study of Network Structure

The proposed network simultaneously learns gradient, orientation field, and ridge period. However, the orientation field and ridge period are only involved in the training process and are not used in testing. The orientation field and ridge period serve as auxiliary tasks to assist in learning the gradient, which is the main advantage of the multi-task learning structure [47].

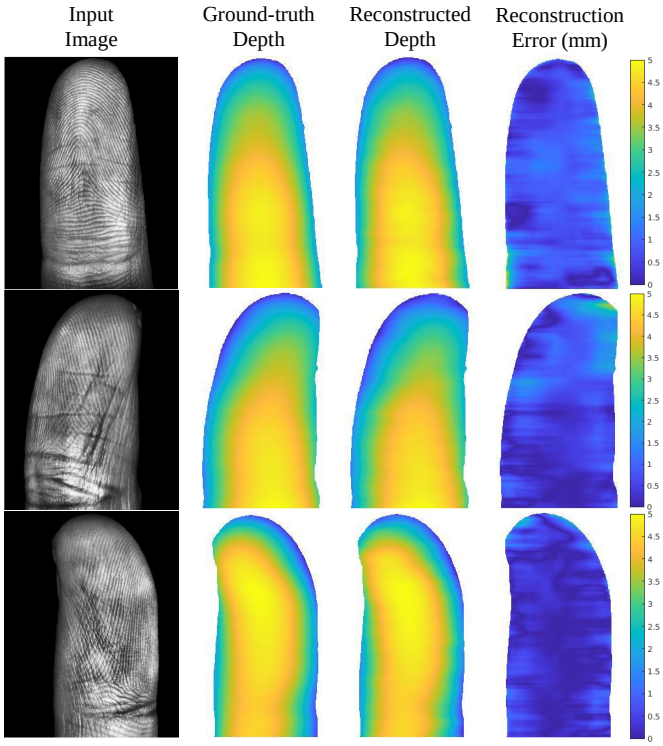


Fig. 8. Illustration of reconstructed 3D finger shape and reconstruction error. Most reconstruction errors occur near the border region.

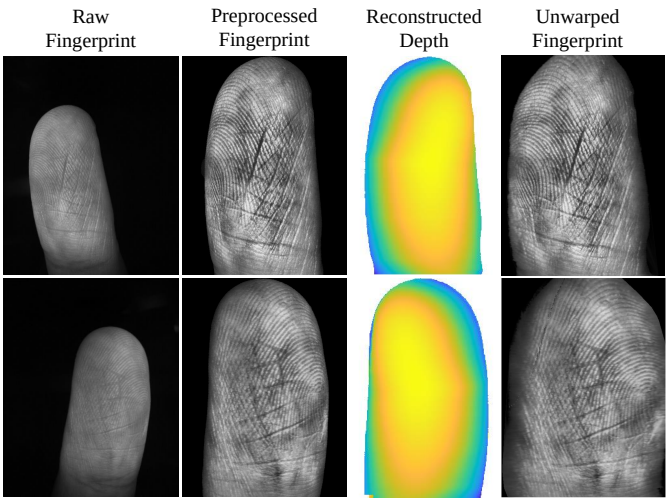


Fig. 9. Illustrations of reconstruction and unwarping results of different poses from the same finger.

We conduct an ablation test on the network structure to examine the advantage of jointly learning the orientation field and the ridge period with a gradient. As shown in Fig. 10, three network structures are compared. Plain network is the proposed network without the orientation and period branches, and only the gradient estimation network is utilized. Parallel network means estimating ridge orientation, period, and gradient parallelly, while the three branches in our network are connected. Fig. 10 shows the skeletons of different networks. The three networks are all trained on the training set described in Section 4.1, and tested on the PolyU 3D database. The three networks are trained with the same batch size and epochs.

TABLE 8
Reconstruction errors on PolyU 3D database by different network structures

Errors	Plain Network	Parallel Network	Our Network
Gradient (mm/Pixel)	0.7054	0.4446	0.4359
Depth (mm)	0.3502	0.2607	0.1802

Table 8 shows the reconstruction errors in the PolyU 3D database for different network structures. Our network has better gradient estimation and depth reconstruction results, which proves that using orientation and period is beneficial for gradient estimation. This is probably because the changes in ridge orientation and period directly reflect the finger surface inclination based on the shape from texture theory [17]. Estimating ridge orientation and period can be viewed as a pre-task for estimating the gradients. They also serve as middle layer outputs to determine whether our network extracts meaningful fingerprint features.

4.3 Unwarping Performance

4.3.1 *Unwarping Influences on Minutiae Matching*

We study the effect of the proposed unwarping algorithm on contactless fingerprint matching. Fig. 11 shows the histogram of increases in mated minutiae pairs after unwarping on the UNSW database. For a pair of mated contactless/contact fingerprints, the minutiae pairs are computed using VeriFinger. It can be seen from Fig. 11 that the number of mated minutiae pairs increase after unwarping.

Fig. 12 shows the histogram of increases in matching scores after unwarping using VeriFinger on the UNSW database. Fig. 12 shows that the matching scores of the corresponding contactless/contact fingerprints increase after unwarping. Therefore, we can conclude that the proposed unwarping algorithm improves the performance of contactless fingerprint matching.

4.3.2 *Comparison of Unwarping Algorithms*

We conduct comparison experiments for three unwarping methods: 1) non-parametric unwarping [45] [48], 2) parametric unwarping [36], and 3) our method.

Non-parametric unwarping [45] [48] maintains local structures by minimizing the distance changes after unwarping:

$$e = \sum_{j \in N_i} \left(d_{ij} - r_{ij}\right)^2, \tag{10}$$

where point j is in the neighborhood of point i, d_{ij} is the current distance between i and j after unwarping in 2D image space, and r_{ij} is the original distance in 3D space. Minimizing the energy function in Equation (10) requires multiple iterations to reach the equilibrium state. Thus, this method is very time-consuming.

Parametric unwarping [49] [36] estimates a parametric 3D model from contactless fingerprint and then unwarps the fingerprint according to the estimated 3D model to reduce perspective distortion. Cylindrical [49] and ellipsoidal [11]

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2022.3233898

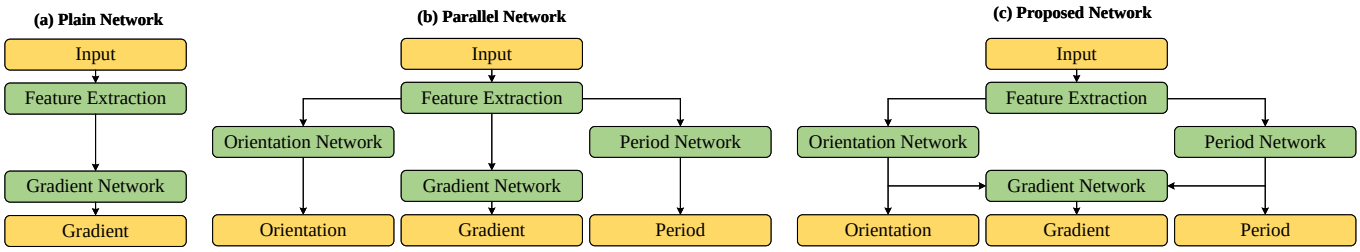
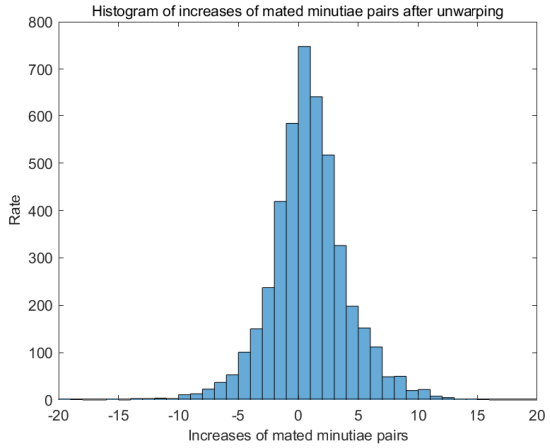
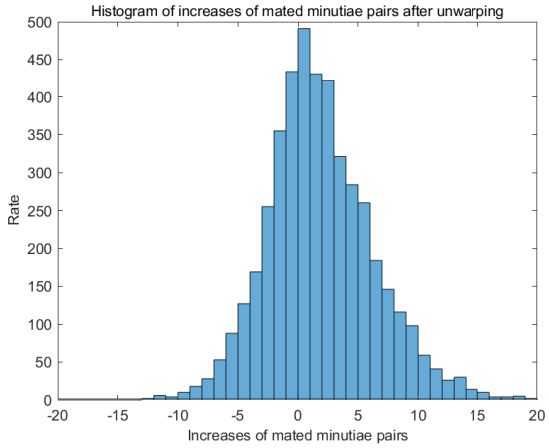


Fig. 10. Illustrations of different network structures.

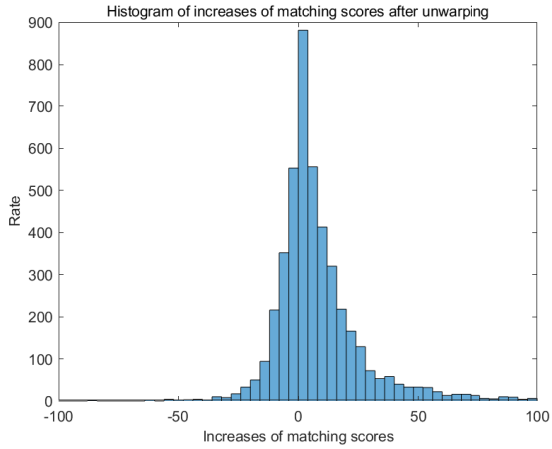


(a) Contactless-to-contact-based fingerprint matching

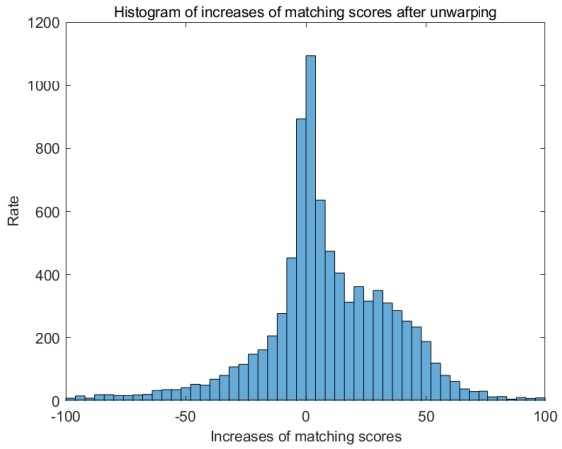


(b) Contactless-to-contactless fingerprint matching

Fig. 11. Histogram of increases of mated minutiae pairs after unwarping on UNSW database.



(a) Contactless-to-contact-based fingerprint matching



(b) Contactless-to-contactless fingerprint matching

Fig. 12. Histogram of increases of matching scores after unwarping on UNSW database.

[36] models are commonly used in estimation, and we select [36] as the baseline for the parametric unwarping method.

Fig. 13 shows the results of the different unwarping algorithms. All three methods can get reasonable unwarping results. Non-parametric unwarping is based on local structures, and it is difficult to preserve long-term information. Moreover, its time cost is very high due to iterations. Parametric unwarping is based on very strong prior shape assumption and may not be suitable for all fingerprints. Our unwarping method relies on no parametric 3D shape, which is more generic and achieves good matching performances.

4.4 Matching Accuracy

4.4.1 Matching Protocols

Matching experiments are performed on UNSW database and PolyU CL2CB database. Two types of matching experiments are conducted: contactless-to-contact-based matching, and contactless-to-contactless matching. Matching scores are generated using VeriFinger 12.0 [13]. Additional matching results by the MCC minutiae matcher [50] are

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2022.3233898

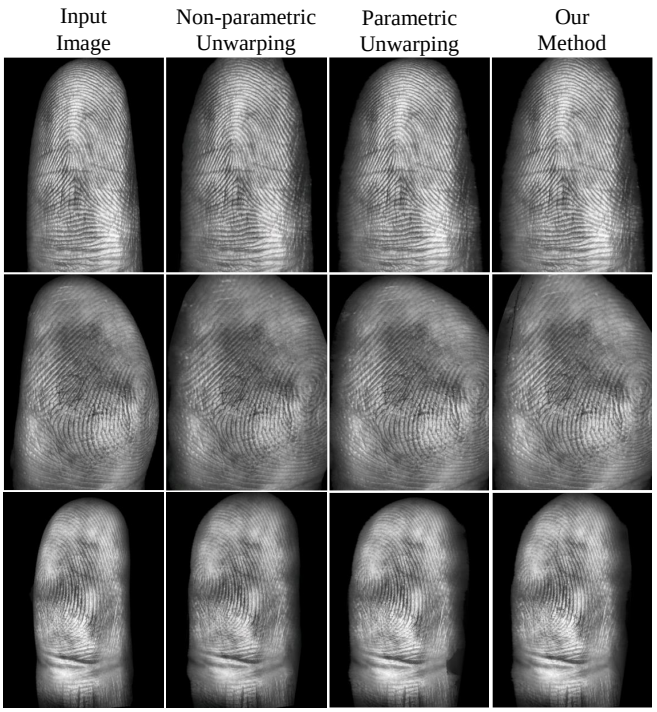


Fig. 13. Examples of different unwarping algorithms.

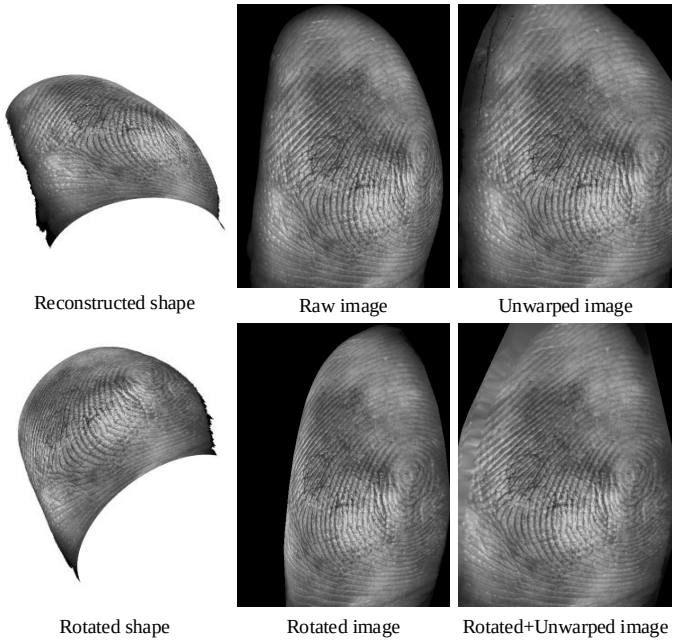


Fig. 14. Illustration of fingerprint unwarping and rotation in matching experiment. The first row is the reconstructed 3D shape, raw image, and unwarped image. The second row is the rotated 3D shape, rotated image, and rotated+unwarped image.

TABLE 9
Equal error rates of matching results on UNSW database by different unwarping algorithms

Matching Experiment	No Unwarping	Non-parametric Unwarping [45]	Parametric Unwarping [36]	Our Method
CL-CB	15.79%	14.43%	15.53%	**14.03%**
CL-CL	23.95%	20.53%	22.04%	**20.18%**
Time cost (s)	–	140	0.98	1.07

provided in the supplementary material, as the matching performance of MCC is not as good as that of VeriFinger, and MCC has not been utilized in previous studies.

UNSW database. Matching experiments are performed on the first 1,000 fingers of the UNSW database. For contactless-to-contact-based fingerprint matching, as there are two captures for each contactless fingerprint and four captures for each contact-based fingerprint, we use the first capture of contactless and contact fingerprint for efficiency. Therefore, a total number of $3,000 \times 1,000$ pairs of matching is performed. As the 3,000 contactless fingerprints are from three different poses, we further divide the contactless-to-contact-based matching into three groups according to pose: front, right, and left. Each group contains $1,000 \times 1,000$ pairs, where only 1,000 pairs are genuine matching, and the remaining pairs are impostor matching.

For contactless-to-contactless matching, we use the first capture of each contactless fingerprint as the reference fingerprint, and the second capture as the input fingerprint. Thus, $3,000 \times 3,000$ pairs of matching are performed. Noticing that matching between fingerprints of different poses from the same finger is also considered genuine matching, there are 9,000 pairs of genuine matching in total, and the rest are impostor matching.

In both matching experiments, the aforementioned matching is conducted on four types of images: raw images, unwarped images, rotated images, and rotated+unwarped images. The rotation here refers to the pose compensation in [14] and [11], which rotates contactless fingerprints of different poses to the same viewing angle. In our experiment, after estimating the 3D shape from the contactless fingerprint, we rotate all 3D shapes to the same front pose, then re-project the 3D finger onto the 2D plane to obtain a rotated image.

The contactless fingerprints in the UNSW database are captured by Surround Imager [21] cameras placed 45° apart. Therefore, the rotation angle is considered a known value of ±45°, and is used in rotation compensation. Note that for front pose fingerprints, their rotated images are the same as raw images. Fig. 14 shows four images of the unwarped results of a side pose fingerprint in the matching experiments.

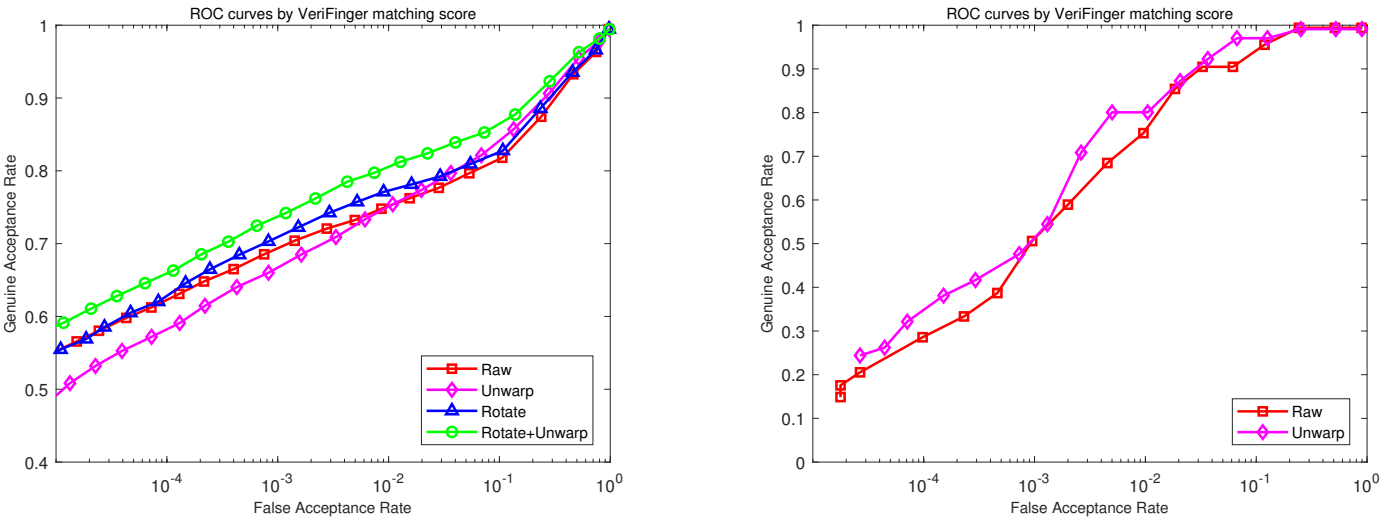
PolyU CL2CB database. The PolyU CL2CB database contains 2,016 contactless fingerprints and the corresponding contact-based fingerprints from 336 fingers × 6 captures. For contactless-to-contact-based matching, there are $336 \times 6 \times 6 = 12,096$ genuine matchings and $336 \times 6 \times 335 \times 6 = 4,052,160$ impostor matchings. For contactless-to-contactless matching, there are $336 \times 15 = 5,040$ genuine matchings and $336 \times 6 \times 335 \times 6 = 4,052,160$ impostor matchings.

As there are no pose differences in the contactless fingerprints of the PolyU CL2CB database, matching experiments are run only on raw images and unwarped images.

4.4.2 *Contactless-to-Contact-based Fingerprint Matching*

Fig. 15 shows the ROC curves of contactless-to-contact-based fingerprint matching on the UNSW and PolyU CL2CB databases. The proposed unwarping algorithm successfully

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2022.3233898



(a) UNSW database (b) PolyU CL2CB database

Fig. 15. ROC curves of contactless-to-contact-based fingerprint matching on two databases.

TABLE 10
Equal error rates of contactless-to-contact-based fingerprint matching on UNSW database

Pose	Raw	Unwarp	Rotate	Rotate +Unwarp
Front	4.00%	**3.59**%	4.00%	**3.59**%
Right	19.31%	17.07%	17.84%	**16.47**%
Left	21.67%	18.70%	20.31%	**16.01**%
All	15.79%	14.03%	14.98%	**12.72**%

TABLE 11
Equal error rates of matching results on PolyU CL2CB database

Matching Experiment	Raw	Unwarp
Contactless-to-Contact-based	5.44%	**5.00**%
Contactless-to-Contactless	7.77%	**5.95**%

TABLE 12
Matching performances of the proposed method

Database	Matching Experiment	Our Method EER(%)	Previous SOTA EER(%)
UNSW	CL-CB	3.59	7.71 [15]
PolyU CL2CB	CL-CB	5.00	7.93 [10]
UNSW	CL-CL	14.12	14.27 [11]
PolyU CL2CB	CL-CL	5.95	–

improves the matching accuracy of contactless fingerprints, which proves that our unwarping algorithm is effective in reducing perspective distortion and improving matching performance. Table 10 shows the detailed matching accuracy of contactless fingerprints with different poses on the UNSW database, and Table 11 shows the results on the PolyU CL2CB database. The unwarping method improves the matching accuracy of all poses on the UNSW database.

We compare our results with [15], which reported state-of-the-art performance on the UNSW database. In [15], only front pose fingerprints are used in the matching experiment, which are much easier to match than the fingerprints of other poses. Therefore, we compare [15] with our front pose matching results. Meanwhile, we compare with [10] on the PolyU CL2CB database. As shown in Table 12, our method is superior to previous methods for contactless-to-contact-based matching on both databases.

It should be noted that a recent study on contactless-to-contact-based matching [27] (submitted at similar time as this paper) has reported high performances (EER=0.46% for PolyU, EER=6.81% for UNSW). They utilized DeepPrint [51] as supervision to guide the network to learn features from contactless fingerprints, which has been trained on very large datasets to improve performances, whereas our methods are trained only on a very limited amount of data for 3D reconstruction.

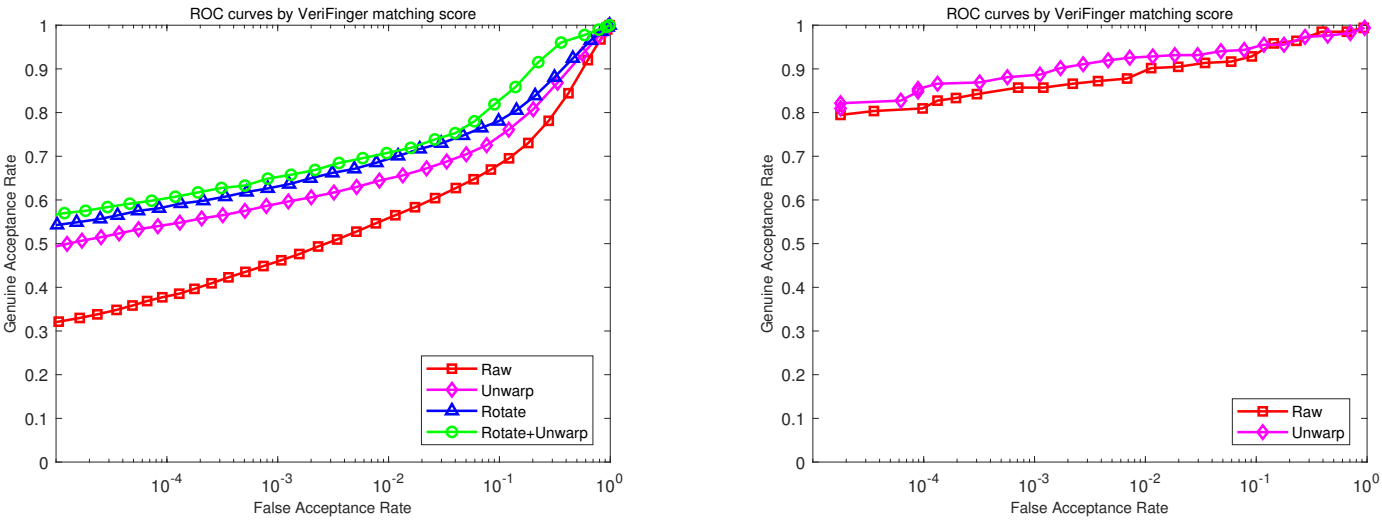
4.4.3 *Contactless-to-Contactless Fingerprint Matching*

Fig. 16 shows the matching results of contactless-to-contactless fingerprints on the UNSW and PolyU CL2CB databases. Matchings between contactless fingerprints does not use pose information to facilitate matching. Namely, images of different poses are considered independent samples. As we can see, unwarping increases the matching performances.

For the UNSW database, because input fingerprints and reference fingerprints both have three poses, there are nine combinations of matching in total. The detailed matching results of these combinations are presented in Table 13. The proposed unwarping algorithm improves the matching performances for all poses. Also, the matching results on the PolyU CL2CB database are listed in Table 11. As we can see, unwarping improves the matching accuracy.

For contactless-to-contactless matching, we compare our method with the state-of-the-art method in [11], which utilizes pose compensation to deal with perspective distortion. Table 12 compares our matching result with [11] that both use VeriFinger [13] to compute the matching score. The results in Table 12 show the average matching results for all poses. As we can see, our method outper-

This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2022.3233898



(a) UNSW database (b) PolyU CL2CB database

Fig. 16. ROC curves of contactless-to-contactless fingerprint matching on two databases.

TABLE 13
Equal error rates of contactless-to-contactless fingerprint matching on UNSW database

Pose	Raw	Unwarp	Rotate	Rotate +Unwarp
Front-Front	8.12%	**6.88**%	8.12%	**6.88**%
Front-Right	16.58%	13.38%	12.92%	**9.79**%
Front-Left	19.55%	16.75%	14.53%	**11.63**%
Right-Front	16.28%	13.12%	12.98%	**10.13**%
Right-Right	12.75%	10.97%	10.50%	**9.27**%
Right-Left	44.03%	38.27%	32.35%	**25.67**%
Left-Front	18.69%	16.73%	14.60%	**11.18**%
Left-Right	43.93%	37.13%	31.60%	**25.94**%
Left-Left	9.99%	8.80%	8.92%	**8.42**%
All	23.95%	20.18%	17.74%	**14.12**%

forms [11] in terms of matching accuracy. As for the PolyU CL2CB database, there is currently no previous contactless-to-contactless matching algorithm on this database to be compared with.

5 Limitations and Future Work

The proposed 3D reconstruction and unwarping method can recover a 3D finger shape from a single contactless fingerprint image and obtain a rolled fingerprint without perspective distortion. The experimental results show that the proposed algorithm achieves good reconstruction accuracy and increases matching performances. However, the current algorithm still has some limitations that require further exploration.

Reconstruction precision. The current reconstruction algorithm recovers a smoothed finger shape without a detailed ridge-valley structure. Although reconstructing 3D ridge-valley structure is a more powerful 3D reconstruction technology, a 3D shape with smoothed surface is sufficient for unwarping a contactless fingerprint to a rolled fingerprint. Even if a 3D ridge-valley structure can be reconstructed, it must be converted back to a 2D ridge-valley pattern for compatibility with the existing 2D fingerprint systems. The ridge-valley structure is already conveyed by the 2D texture image. It is not yet clear whether the reconstruction of a 3D ridge-valley structure can benefit real applications.

Dataset precision. The PolyU 3D and UNSW datasets used for training the gradient estimation network have limited 3D accuracy, particularly the UNSW dataset, whose ground-truth depth map is reconstructed using images from three view angles by shape from the silhouette method. However, these databases are the most commonly used publicly available contactless fingerprinting databases. We must utilize these databases to compare the results with those of previous studies on these databases. Our method can benefit from more accurate 3D fingerprint for training. However, our method is already able to decrease the perspective distortion and increase the matching performance of contactless fingerprints, even when using 3D fingerprint data that are not as accurate. We plan to study how the accuracy of 3D training data affects the final matching performance.

Dataset complexity. The acquisition of contactless fingerprints in certain application scenarios may be more complex and flexible, such as collection by smartphones, thereby increasing the requirements for the adaptability of the proposed algorithm. In this study, the proposed network was trained on two datasets, in which the camera parameters were not significantly different. Therefore, an algorithm with the current parameters may have problems if directly applied to fingerprint photos taken from long distances. A simple solution is to collect a dataset similar to the target application scenario and retrain the network. The two datasets in our study were collected in typical fingerprint recognition scenarios, and the proposed method performed well under these normal situations. In the future, more training data under different imaging settings can be collected to explore the generalization of our method to significantly different camera parameters.

Scale ambiguity. Generally, it is very difficult to completely determine the scale of monocular 3D reconstruction

methods without changing the hardware or adding certain clues. Our solution is using the mean ridge period value to adjust the scale. However, the mean ridge periods for different people may differ slightly. For example, the ridge period of a baby's fingerprint is smaller than that of an adult's fingerprint. Therefore, the reconstructed 3D fingerprint may have some scale ambiguity, depending on the small differences between a certain finger's ridge period and the average ridge period among all people. Although unifying ridge periods of all fingerprints does not harm genuine matches, it may lead to increases in impostor matching scores. Therefore, much larger datasets are required to address this problem.

Skin distortion. The current algorithm solves perspective distortion for contactless fingerprints; however, skin distortion also needs to be considered when matching contact-based fingerprints. This study deals with perspective distortion for contactless fingerprint, which results from projecting a 3D finger onto a 2D image. The proposed algorithm is effective for both contactless-to-contactless matching and contactless-to-contact-based matching, as shown in our experiments. Because skin distortion is very diverse depending on the direction and force of the pressing finger, our unwarping method does not consider skin distortion. To further improve the similarity between contactless and contact-based fingerprints, distortion rectification of contact-based fingerprints can be studied [15] [16].

Feature extraction and matching. The focus of this study is the geometric aspect of contactless fingerprints, and image quality also deserves attention. The feature extraction and matching methods used in the matching experiment are still conventional methods developed for contact-based fingerprints. Feature extraction and matching are certainly important for the overall performance. In the future, we plan to study particular feature extraction and matching methods suitable for contactless fingerprints.

6 Conclusion

This paper proposes a new approach for contactless 3D fingerprint reconstruction and unwarping using deep learning. Our method contains four steps: contactless fingerprint preprocessing, network for surface gradient estimation, 3D shape reconstruction, and the contactless fingerprint unwarping. Reconstruction experiments on contactless fingerprint databases show that the proposed method has low reconstruction error and high unwarping quality. Matching experiments prove that the proposed method can reduce perspective distortion and is beneficial for contactless fingerprint matching.

Acknowledgments

We are grateful to Dr. Feng Liu for providing us with the PolyU 3D database of contactless fingerprint images and the corresponding depth data. We would also like to thank the research groups for making their fingerprint data publicly available.

This work was supported in part by the National Natural Science Foundation of China under Grant 61976121 and 62206026.

References

[1] D. Maltoni, D. Maio, A. K. Jain, and J. Feng, *Handbook of Fingerprint Recognition*. Springer, 2022.
[2] A. Kumar, *Contactless 3D Fingerprint Identification*. Springer, 2018.
[3] Y. Song, C. Lee, and J. Kim, "A new scheme for touchless fingerprint recognition system," in *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2004, pp. 524–527.
[4] C. Lee, S. Lee, and 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)*, 2006, pp. 358–365.
[5] B. Y. Hiew, A. B. Teoh, and Y.-H. Pang, "Touch-less fingerprint recognition system," in *IEEE Workshop on Automatic Identification Advanced Technologies*, 2007, pp. 24–29.
[6] A. Kumar and Y. Zhou, "Contactless fingerprint identification using level zero features," in *CVPR Workshop on Biometrics*, 2011, pp. 114–119.
[7] W. Zhou, J. Hu, S. Wang, I. Petersen, and M. Bennamoun, "Performance evaluation of large 3D fingerprint databases," *Electronics Letters*, vol. 50, no. 15, pp. 1060–1061, 2014.
[8] G. Fiumara, E. Tabassi, P. Flanagan, J. Grantham, K. Ko, K. Marshall, M. Schwarz, B. Woodgate, and C. Boehnen, "Nail to nail fingerprint challenge: Prize analysis," *NISTIR 8210*, 2018.
[9] S. Chopra, A. Malhotra, M. Vatsa, and R. Singh, "Unconstrained fingerphoto database," in *CVPR Workshop on Biometrics*, 2018, pp. 517–525.
[10] C. Lin and A. Kumar, "A CNN-based framework for comparison of contactless to contact-based fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 662–676, 2019.
[11] H. Tan and A. Kumar, "Towards more accurate contactless fingerprint minutiae extraction and pose-invariant matching," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3924–3937, 2020.
[12] A. K. Jain, L. Hong, and R. Bolle, "On-line fingerprint verification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 302–314, 1997.
[13] Neurotechnology Inc., VeriFinger SDK 12.0. http://www.neurotechnology.com.
[14] R. D. Labati, A. Genovese, V. Piuri, and F. Scotti, "Contactless fingerprint recognition: A neural approach for perspective and rotation effects reduction," in *IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, 2013, pp. 22–30.
[15] A. Dabouei, S. Soleymani, J. Dawson, and N. M. Nasrabadi, "Deep contactless fingerprint unwarping," in *International Conference on Biometrics (ICB)*, 2019.
[16] X. Si, J. Feng, J. Zhou, and Y. Luo, "Detection and rectification of distorted fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 555–568, 2015.
[17] D. Blostein and N. Ahuja, "Shape from texture: Integrating texture-element extraction and surface estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 12, pp. 1233–1251, 1989.
[18] F. Liu and D. Zhang, "3D fingerprint reconstruction system using feature correspondences and prior estimated finger model," *Pattern Recognition*, vol. 47, no. 1, pp. 178–193, 2014.
[19] W. Zhou, J. Hu, I. Petersen, S. Wang, and M. Bennamoun, "A benchmark 3D fingerprint database," in *International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2014, pp. 935–940.
[20] C. Lin and A. Kumar, "Matching contactless and contact-based conventional fingerprint images for biometrics identification," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2008–2021, 2018.
[21] G. Parziale, E. Diaz-Santana, and R. Hauke, "The surround imager: A multi-camera touchless device to acquire 3D rolled-equivalent fingerprints," in *International Conference on Biometrics (ICB)*, 2006, pp. 244–250.
[22] C. Stein, C. Nickel, and C. Busch, "Fingerphoto recognition with smartphone cameras," in *International Conference of Biometrics Special Interest Group (BIOSIG)*, 2012, pp. 1–12.
[23] A. Sankaran, A. Malhotra, A. Mittal, M. Vatsa, and R. Singh, "On smartphone camera based fingerphoto authentication," in *International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2015, pp. 1–7.

[24] A. Malhotra, A. Sankaran, A. Mittal, M. Vatsa, and R. Singh, "Fingerphoto authentication using smartphone camera captured under varying environmental conditions," in *Human Recognition in Unconstrained Environments*, 2017, pp. 119–144.

[25] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, and F. Scotti, "A novel pore extraction method for heterogeneous fingerprint images using convolutional neural networks," *Pattern Recognition Letters*, vol. 113, pp. 58–66, 2018.

[26] C. Lin and A. Kumar, "Multi-siamese networks to accurately match contactless to contact-based fingerprint images," in *International Joint Conference on Biometrics (IJCB)*, 2017, pp. 277–285.

[27] S. A. Grosz, J. J. Engelsma, E. Liu, and A. K. Jain, "C2CL: Contact to contactless fingerprint matching," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 196–210, 2022.

[28] R. D. Labati, A. Genovese, V. Piuri, and F. Scotti, "Toward unconstrained fingerprint recognition: A fully touchless 3D system based on two views on the move," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 2, pp. 202–219, 2015.

[29] A. Kumar and C. Kwong, "Towards contactless, low-cost and accurate 3D fingerprint identification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 3438–3443.

[30] C. Lin and A. Kumar, "Tetrahedron based fast 3D fingerprint identification using colored LEDs illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 3022–3033, 2017.

[31] G. Abramovich, K. Harding, S. Manickam, J. Czechowski, V. Paruchuru, R. Tait, C. Nafis, and A. Vemury, "Mobile, contactless, single-shot, fingerprint capture system," in *Biometric Technology for Human Identification VII*, vol. 7667, 2010, p. 766708.

[32] Y. Wang, L. G. Hassebrook, and D. L. Lau, "Data acquisition and processing of 3D fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 750–760, 2010.

[33] A. Baradarani, R. G. Maev, and F. Severin, "Resonance based analysis of acoustic waves for 3D deep-layer fingerprint reconstruction," in *IEEE International Ultrasonics Symposium (IUS)*, 2013, pp. 713–716.

[34] J. Galbally, G. Bostrom, and L. Beslay, "Full 3D touchless fingerprint recognition: Sensor, database and baseline performance," in *International Joint Conference on Biometrics (IJCB)*, 2017, pp. 225–233.

[35] F. Liu, D. Zhang, and L. Shen, "Study on novel curvature features for 3D fingerprint recognition," *Neurocomputing*, vol. 168, pp. 599–608, 2015.

[36] D. Sollinger and A. Uhl, "Optimizing contactless to contact-based fingerprint comparison using simple parametric warping models," in *International Joint Conference on Biometrics (IJCB)*, 2021.

[37] L. Ericson and S. Shine, "Evaluation of contactless versus contact fingerprint data phase 2 (version 1.1)," *DOJ Office Justice Programs, I. ManTech Adv. Syst. Int., Fairmont, WV, USA, Tech. Rep*, vol. 249552, 2015.

[38] K. Zuiderveld, "Contrast limited adaptive histograph equalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 2, pp. 150–162, 1994.

[39] D. Eigen, C. Puhrsch, and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," in *Advances in Neural Information Processing Systems (NIPS)*, vol. 27, 2014, pp. 2366–2374.

[40] R. T. Moore, "Analysis of ridge-to-ridge distance on fingerprints," *Journal of Forensic Identification*, vol. 39, no. 4, pp. 231–238, 1989.

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

[42] Y. Tang, Y. Liu, and J. Feng, "FingerNet: An unified deep network for fingerprint minutiae extraction," in *International Joint Conference on Biometrics (IJCB)*, 2017, pp. 108–116.

[43] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *The 3rd International Conference on Learning Representations (ICLR2015)*, 2015.

[44] M. Kass and A. Witkin, "Analyzing oriented patterns," *Computer Vision, Graphics, and Image Processing*, vol. 37, no. 3, pp. 362–385, 1987.

[45] A. Fatehpuria, D. L. Lau, and L. G. Hassebrook, "Acquiring a 2D rolled equivalent fingerprint image from a non-contact 3D finger scan," in *Biometric Technology for Human Identification III*, vol. 6202, 2006, p. 62020C.

[46] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, vol. 5, 2015.

[47] S. Ruder, "An overview of multi-task learning in deep neural networks," *arXiv:1706.05098*.

[48] S. Shafaei, T. Inanc, and L. G. Hassebrook, "A new approach to unwrap a 3-D fingerprint to a 2-D rolled equivalent fingerprint," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, 2009, pp. 1–9.

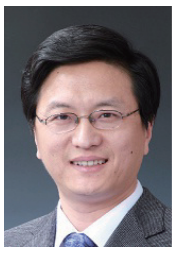
[49] Y. Chen, G. Parziale, E. Diaz-Santana, and A. K. Jain, "3D touchless fingerprints: Compatibility with legacy rolled images," in *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, 2006.

[50] R. Cappelli, M. Ferrara, and D. Maltoni, "Minutia cylinder-code: A new representation and matching technique for fingerprint recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2128–2141, 2010.

[51] J. J. Engelsma, K. Cao, and A. K. Jain, "Learning a fixed-length fingerprint representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 6, pp. 1981–1997, 2021.

Zhe Cui received the B.S. degree from the Department of Physics, Tsinghua University in 2014, the M.Eng. and Ph.D. degrees from the Department of Automation, Tsinghua University, in 2017 and 2021, respectively. He is currently an Associate Professor with the School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include fingerprint recognition, computer vision, and pattern recognition.

Jianjiang Feng received the B.Eng. and Ph.D. degrees from the School of Telecommunication Engineering, Beijing University of Posts and Telecommunications, China, in 2000 and 2007, respectively. From 2008 to 2009, he was a Post-Doctoral Researcher with the PRIP Laboratory, Michigan State University. He is currently an Associate Professor with the Department of Automation, Tsinghua University, Beijing. His research interests include fingerprint recognition and computer vision.



Jie Zhou received the B.S. and M.S. degrees from the Department of Mathematics, Nankai University, Tianjin, China, in 1990 and 1992, respectively, and the Ph.D. degree from the Institute of Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China, in 1995. From 1995 to 1997, he served as a Post-Doctoral Fellow with the Department of Automation, Tsinghua University, Beijing, China. Since 2003, he has been a Full Professor with the Department of Automation, Tsinghua University. In recent years, he has authored over 100 papers in peer-reviewed journals and conferences. Among them, over 40 papers have been published in top journals and conferences, such as the IEEE PAMI, TIP, and CVPR. His current research interests include computer vision, pattern recognition, and image processing. He received the National Outstanding Youth Foundation of China Award. He is an Associate Editor of the IEEE Transactions on Pattern Analysis and Machine Intelligence, the International Journal of Robotics and Automation, and two other journals.