Distinctive Feature Representation for Contactless 3D Hand Biometrics using Surface Normal Directions

Kevin H. M. Cheng, Ajay Kumar Department of Computing, The Hong Kong Polytechnic University, Hong Kong

kevin.hm.cheng@connect.polyu.hk, ajay.kumar@polyu.edu.hk

Abstract

Contactless 3D hand biometrics offers hygienic and convenient approaches for biometric recognition. This paper investigates a distinctive feature representation using 3D surface normal information for more accurate 3D hand biometric identification. Prior research on contactless 3D hand biometric identification largely incorporates 3D depth and surface curvature information to recover discriminative features. Our investigation presented in this paper indicates that extracting distinctive features from surface normal information, which can also be directly obtained from low-cost photometric stereo based imaging systems, can offer a computationally simpler alternative and is therefore highly desirable. The directions of neighbouring surface normal vectors can encode frequently observed irregular ridge and valley regions, which can enable more accurate surface feature description. Comparative experimental results presented in this paper validates the effectiveness of the proposed approach.

1. Introduction

Biometric recognition is an important research area with many useful applications such as the immigration inspection, access control and wearable electronics [1]. Face [2, 3], fingerprint [4, 7], and iris [6] are the most popular biometric identifiers for the automated recognition of human identities. Other biometric identifiers such as palmprint [9, 10] and finger knuckle [11-14], which provide easily visible patterns and can be acquired simultaneously, have also emerged for biometric recognition. Beside the choices of biometrics, several research efforts including 3D ear [15, 16], 3D face [17-19], 3D palmprint [20, 23] and 3D fingerprint [21] have shown that using also the 3D domain enables richer information and therefore higher recognition performance. These observations motivate us to address two in-

| | Surface Code [23] | Binary Feature [20] | Collabor. Repres. [27] | Ours |
|--------------------|--------------------------------------|------------------------------------|--|--|
| Source of Info. | Curvature | Depth | Depth | Surface Normal |
| Info. Extracted | Texture Patterns | Texture Patterns | Block- wise Statistics | Texture Patterns with Ori- entation |
| Limitation | Lacks Theoret- ical Support | Lacks Orienta- tion Info. | Fails for Chal- lenging Samples | Larger Tem- plate Size |
| Accuracies | Medium | Medium | Medium | High |

| Table 1: Comparative summary of va | rious feature represen- |
|--------------------------------------|-------------------------|
| tations for 3D finger knuckle and 3D | palmprint recognition. |

teresting biometric problems, i.e. 3D finger knuckle and 3D palmprint recognition.

Among the literature on contactless 3D palmprint recognition, Surface Code [23] and Binary Feature [20], are 3D feature descriptors for palmprint and are also applicable for finger knuckle. Surface Code discretizes the Shape Index [24, 25] using four binary values which are compared using Hamming distance. However, the distance between the encoded feature values lacks theoretical support, which will be further elaborated in Section 2. Binary Feature discretizes a surface type using a binary value computed from ordinal relationship in a small neighbourhood. However, the 3D surface orientation or direction information, which is known to describe local discriminative patterns, has not been considered in both methods. In addition, although learning based approaches [27, 29, 30] have emerged as a popular and effective solution for many computer vision problems, such performance can be largely degraded by the large variance between the distribution of gallery (train) and probe (test) samples. Therefore, this paper attempts to develop a new feature representation of surface texture and directional information from a mathematical perspective.

In order to develop a more robust feature descriptor for 3D hand biometrics, it is crucial to have an insight into the 3D imaging technologies. The photometric stereo reconstruction approach is shown to offer a low cost and accurate reconstruction of surface details [5]. With such an approach, surface normal images can be recovered while depth images can be obtained by the integration of surface normal vectors. If the source surface normal vectors are used instead of the depth map, we can alleviate the need for the additional integration, which also induces reconstruction errors (well-known as the integrability problem [28] in the literature). Existing work on contactless 3D biometric identification largely incorporated 3D depth and surface curvature information to recover discriminative features, while surface normal information, which is also invariant to pose and illumination variations, has not gained adequate attention. Therefore, this paper attempts to investigate the possibility of using surface normal information for more robust 3D feature description. The key contributions from this paper can be summarized as follows:

This paper investigates the possibility of using surface normal information instead of depth information for more robust 3D surface feature description and introduces a new 3D feature representation, surface normal direction, for more accurate 3D finger knuckle and 3D palmprint recognition. We can alleviate the need for computing the depth maps by using the source 3D information, i.e. surface normal vectors. The directions of local surface normal vectors can also distinctively describe irregular ridge and valley patterns, which depict discriminative features for biometric recognition. Unlike the two aforementioned methods, our method also extracts orientation or direction information in addition to the texture information. Comparative experimental results presented in section 4 of this paper illustrate superior performance over the evlauted methods on both the contactless 3D finger knuckle and 3D palmprint databases, which validates our theoretical arguments that surface normal information provides robust information, as well as the effectiveness of the proposed feature representation and matcher. Table 1 presents a comparative summary of various feature representations for 3D finger knuckle and 3D palmprint recognition. The recognition accuracies presented in table 1 are averages of the experimental results for 3D finger knuckle and 3D palmprint recognition. In order to ensure the reproducibility of this work, implementation codes are also made publicly available [8].

2. Related Work

The literature of 3D palmprint recognition provided solid foundations for the problem investigated in this paper. The *Surface Code* descriptor [23] is a 3D feature descriptor with a template size of four bits per pixel. *Shape indexes* [24, 25], containing curvature information, are discretized into four binary values. However, such discretization and matching scheme lacks justifications. For example, a surface type can be discretized into nine levels. For such encoding scheme, level 7 is encoded as '0111' while level 8 is encoded as '1000'. The similarity score between these two seemly close-together levels are computed to be 4 using the Hamming distance measure. In contrary, level 0 is encoded as '0000'. Despite this surface type is far different from level 8, the similarity score is computed to be 1. The Binary Feature descriptor [20] is another 3D feature descriptor with a template size of one bit per pixel. Depth images are convoluted with ordinal filters for the computation of feature values. The major strength of this method is its efficiency due to the simple computation and small template size. However, this simple feature descriptor may not fully utilize the discriminative information. For, example, the orientation or direction information, which is known to contain discriminative information, is not considered. The effectiveness of using similar surface type based mehods has been verified in a recent reference on 3D palmprint recognition [32]. Meanwhile, other more recent work [26, 33] also attempted to describe 3D surfaces for finger knuckle recognition. Another work utilized the collaborative representation with L_2 and L_1 norm regularizations [27]. This approach requires the learning of feature vectors from gallery (training) samples. If the variations between the gallery and probe (test) samples are significant, it is challenging to learn the robust feature representations.

Deep learning technologies have been actively investigated and the effectiveness are validated in various applications including object recognition [29-30] and biometrics [34-35]. VGGNet [29] is one of the representatives from Convolutional Neural Network (CNN) approach, which investigated the use of a small kernel size (i.e. 3×3) for extracting robust features. ResNet [30] is another more recent CNN based method, using residual components for enhancing the training of deep neural networks. However, such machine learning methods usually suffer when the varience of statistical distribution between training and testing sets is large. For example, if the acquisition of gallery and probe images is from different cameras, those image variations can constitute the differences between the training sets and testing sets. Furthermore, a certain level of customized development is required on specific biometric problems. For example, a research on iris recognition [34] considers the biometric aspects of using binary templates and the bitshifting strategy for matching the templates. Therefore, deep learning approaches on the problems addressed in this paper are highly interesting and will be promising future research areas.

3. Surface Normal Direction

To begin with the introduction of this feature representation, figure 1 shows an illustration using a convex and a concave case with surface normal vectors in a cross-sectional view in the smallest possible neighbourhood, i.e. 3 in a dimension. The black vectors represent the central vectors of the respective neighbourhood while the red vectors represent the neighbouring vectors. The dotted black vectors are in the same direction as the central vector for better visualization. The core idea to differentiate between the fundamental convex (corresponding to ridges) and concave (corresponding to valleys) cases are to consider whether the pair of neighbouring surface normal vectors are pointing inwards or outwards respective to the central surface normal vector.



Figure 1: Illustration of Surface Normal Direction Feature with a convex and a concave case.

Firstly, surface normal information of 3D hand biometrics can be obtained from 3D reconstruction using a photometric stereo approach. Figure 2(a) shows the surface normal vectors on a 3D finger knuckle surface. Surface normal vectors contain rich information and can be used to accurately describe the discriminative feature of a surface. After that, the surface normal vectors are transformed locally and the surface normal direction feature can be computed according to the core idea illustrated in figure 1. A pair of feature templates is matched by a specially designed matcher for obtaining the similarity score. The details of each procedure will be introduced in the following subsections.

3.1. Surface Normal Transformation

In order to consider whether the neighbouring surface normal vectors are pointing inwards or outwards, we can first transform the central surface normal vector of each neighbourhood of $m \times m$ pointing to the top view direction. For simplicity, we only consider m to be 3. It is equivalent to consider a sliding window with size $m \times m$ on a surface normal vector image. For each sliding window, we have to find a transformation matrix so that the central surface normal vector in the window is transformed to be pointing the top direction, which is a trivial mathematical problem. Let **n** be a unit 3D surface normal vector; $\mathbf{t} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ be the vector pointing the top direction; **R** be a rotational matrix such that:

$$\mathbf{t} = \mathbf{R} \ \mathbf{n} \tag{1}$$

Let **c** and *d* be a cross product and a dot product of **n** and **t** respectively:

$$\mathbf{c} = \mathbf{n} \times \mathbf{t} = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}^T \tag{2}$$

$$d = \mathbf{n} \cdot \mathbf{t} \tag{3}$$

The rotational matrix **R** is computed as:

$$\mathbf{R} = \mathbf{I} + g(\mathbf{c}) + g(\mathbf{c})^2 \frac{1}{d+1}$$
(4)

where I is an identity matrix, $g(\mathbf{c})$ is the skew-symmetric cross-product matrix of \mathbf{c} :

$$g(\mathbf{c}) = \begin{bmatrix} 0 & -c_3 & c_2 \\ c_3 & 0 & -c_1 \\ -c_2 & c_1 & 0 \end{bmatrix}$$
(5)

All the surface normal vectors in the sliding window are transformed by multiplying \mathbf{R} :

$$n_{i,j}' = \mathbf{R} \ n_{i,j} \tag{6}$$

where $n'_{i,j}$ is the transformed vectors, i, j is the indexes representing one of the surface normal vectors in a sliding window. Figure 3 shows a graphical example of the transformation. After the transformation, the central surface normal vector must be pointing the top direction. If a larger neighbourhood is considered, each of the neighboring surface normal vectors can also be transformed by multiplying the same **R**.

3.2. Feature Representation in Four Directions

The transformed surface normal vectors in a sliding window are used for computing a feature value, defined as the Surface Normal Direction feature, in each of the four possible directions. These four directions can be observed in figure 2(b). When considering a pair of neighbouring surface normal vectors in one direction, both vectors pointing inwards to the central vector represents a concave case. Similarly, both vectors pointing outwards from the central vector represents a convex case (illustrated in figure 1). The concave cases describe valley regions while the convex cases describe ridge regions. Let N be a three-dimensional matrix with each entry representing the transformed surface normal vector in a region of $m \times m$, m = 3.

$$\mathbf{N} = \begin{bmatrix} \mathbf{n}_{11} & \mathbf{n}_{12} & \mathbf{n}_{13} \\ \mathbf{n}_{21} & \mathbf{n}_{22} & \mathbf{n}_{23} \\ \mathbf{n}_{31} & \mathbf{n}_{32} & \mathbf{n}_{33} \end{bmatrix}$$
(7)



Figure 2: Illustration of: (a) surface normal vectors on a finger knuckle surface; (b) four directions and their respective backward and forward values; (c) the projection vectors for computing b_3 .



Figure 3: Surface normal vectors over a neighbourhood region of a pixel: (a) original; (b) transformed.

where $\mathbf{n}_{ij} = [n_{ij}^x \ n_{ij}^y \ n_{ij}^z]^T$, $i, j \in \{1, 2, 3\}$. Note that $\mathbf{n}_{22} = [0 \ 0 \ 1]^T$ because of the transformation procedure presented in section 3.1. We firstly consider the *horizon-tal* direction as the first direction. We define b_1, f_1 be a backward and a forward value for the first considered direction (*horizontal*) respectively. Figure 2(b) shows the four considered directions (*horizontal*, vertical and two diagonals) and their respective spatial location for the concerning backward and forward values. While the backward and forward values can be first computed from the transformed surface normal vectors, we can utilize those values for the computation of the final feature value. For the first considered direction (*horizontal*), only the horizontal dimension (*x*-dimension) in the surface normal vectors are considered for the computation. We directly obtain the backward and forward values of this direction as follows:

$$b_1 = n_{21}^x$$
, $f_1 = n_{23}^x$ (8)

Similarly, we define b_2 , f_2 be a backward and a forward value for the second considered direction (*vertical*) respectively. For this direction, only the vertical dimension (*y*-dimension) in the surface normal vectors are required for

the computation:

$$b_2 = n_{12}^y$$
 , $f_2 = n_{32}^y$ (9)

For the other two *diagonal* directions, both the horizontal dimension (x-dimension) and the vertical dimension (ydimension) in the surface normal vectors are required for the computation. We first consider the direction from top left to bottom right. We define b_3 , f_3 be a backward and a forward value for this diagonal direction respectively. In order to compute b_3 and f_3 , the projection vector on the diagonal direction is required. For example, consider the computation of b_3 (illustrated in Figure 2(c)). Let a be the shortest distance between the coordinate (n_{11}^x, n_{11}^y) and the line y = x, or x - y = 0. The distance a is:

$$a = \frac{|n_{11}^x - n_{11}^y|}{\sqrt{1+1}} \tag{10}$$

The magnitude of b_3 is:

$$|b_3| = \sqrt{(n_{11}^x)^2 + (n_{11}^y)^2 - a^2}$$
(11)

The sign of b_3 is:

$$sgn(b_3) = sgn([n_{11}^x, n_{11}^y] \cdot [1, 1])$$
(12)

Combining equations (10)-(12) we can compute b_3 as:

$$b_{3} = \sqrt{(n_{11}^{x})^{2} + (n_{11}^{y})^{2} - (\frac{n_{11}^{x} - n_{11}^{y}}{\sqrt{2}})^{2}} \cdot sgn(n_{11}^{x} + n_{11}^{y})$$
(13)

Similarly, we can also compute the forward value f_3 as:

$$f_{3} = \sqrt{(n_{33}^{x})^{2} + (n_{33}^{y})^{2} - (\frac{n_{33}^{x} - n_{33}^{y}}{\sqrt{2}})^{2}} \cdot sgn(n_{33}^{x} + n_{33}^{y})$$
(14)

where sgn is a sign function. The last considered direction is from the top right to the bottom left. We define b_4 , f_4 be a backward and a forward value for this diagonal direction respectively. The line equation of this direction is y = -x, or x + y = 0. Similar to the last diagonal direction we can compute b_4 as:

$$b_4 = \sqrt{(n_{13}^x)^2 + (n_{13}^y)^2 - (\frac{n_{13}^x + n_{13}^y}{\sqrt{2}})^2 \cdot sgn(n_{13}^x - n_{13}^y)}$$
(15)

and f_4 as:

$$f_4 = \sqrt{(n_{31}^x)^2 + (n_{31}^y)^2 - (\frac{n_{31}^x + n_{31}^y}{\sqrt{2}})^2 \cdot sgn(n_{31}^x - n_{31}^y)}$$
(16)

After obtaining all the four pairs of backward and forward values, we can compute the final feature values for each of the four directions respectively. The feature values are computed as follows:

$$p_{k} = \begin{cases} 1, & b_{k} > 0 \text{ and } f_{k} < 0 \\ 3, & b_{k} < 0 \text{ and } f_{k} > 0 \\ 2, & otherwise \end{cases}$$
(17)

where $k \in \{1, 2, 3, 4\}$ represents the four considered directions (*horizontal*, *vertical* and two *diagonals*). The pixelwise feature value p_k represents a concave (valley) case when $p_k = 1$; a convex (ridge) case when $p_k = 3$; an ambiguous (uncertain) case when $p_k = 2$. The computation of this pixel-wise feature value is repeated for all the pixels in a surface normal image. Finally, we can obtain a feature template with a dimension of the image height × the image width × four.

In summary, when the backward value is larger than zero and the forward value is less than zero, it indicates that both neighboring surface normal vectors are pointing inwards, which corresponds to a concave/valley case. Similarly, the cases of ridge and uncertainty can also be encoded, which will result in four feature images, each with three encodings, for the four considered directions as feature templates for matching. Therefore, when comparing a pair of feature images, the encoding resulted from each direction can be compared individually. For a more efficient implementation, the magnitude of the forward and backward values can be ignored. Figure 4 presents plots of concave (in red) and convex (in blue) features on 3D images, while the ambiguous (uncertain) features remains in grey. These visualizations demonstrate the detection of the concave and convex regions on both 3D finger knuckle and 3D palmprint surfaces.

3.3. Matching Pairs of Feature Templates

The Surface Normal Direction feature consists of four values per pixel, with each value to be either 1, 2, or 3. In order to maximize the performance of using such feature representation, a specific similarity matrix is also developed. It

is expected that all three cases (i.e. concave, convex, and ambiguous) provide helpful information for discriminating identities. For a pair of pixels from the probe and gallery templates, they are considered to be similar if their values are the same. Let A and B be two templates of surface normal direction feature with size $M \times N \times 4$. Let a_{ijk} and b_{ijk} ($i \in [1, M], j \in [1, N], k \in [1, 4]$) be the feature value (i.e. 1, 2, or 3) in A and B respectively. Let h be a distance function:

$$h(a_{ijk}, b_{ijk}) = \frac{|a_{ijk} - b_{ijk}|}{2}$$
(18)

where $h \in \{0, 0.5, 1\}$. A zero response from h indicates that same feature values are obtained from a pair of templates (considered to be similar), while a unity response from h indicates that both concave and convex case occur in the pair of templates (considered to be not similar). The remaining case, response from h equals 0.5, indicates the case that ambiguous case is obtained in one template, while either concave or convex case is obtained in another template. Those cases are quite unstable and would degrade the recognition performance. Therefore, those cases are ignored for the similarity computation. We assume that each of the four considered directions have a equal importance. Let o_1 and o_0 be the occurrence of h equals 1 and 0 respectively, for a pair of templates A and B. The comparison score s for comparing the similarity between this pair of templates is computed as:

$$s = \frac{o_1}{o_1 + o_0}$$
(19)

Ideally, s equals 0 indicates the most similar case while s equals 1 indicates the most dissimilar case. By partially ignoring the cases of h equals 0.5, the similarity scores are more robust. To account for pose variations, rotational or translational shifting can be applied on probe images. The minimum of the matching scores from matching the shifted versions of the probe images with the gallery images are the final matching score.

4. Experiments and Results

The performance of our proposed method is evaluated using the verification, closed-set and open-set identification experiments. The results are presented using receiver operating characteristics (ROC) curve, cumulative match characteristics (CMC) curve, and false negative identification rate (FNIR) versus false positive identification rate (FPIR) curve. The FNIR and FPIR can be computed as follows (corrected equations in [22]):

$$FPIR(T) = \frac{1}{K} \sum_{i=1}^{K} H(T - s_{i1})$$
 (20)

$$FNIR(T) = 1 - \frac{1}{M} \Sigma_{i=1}^{M} H(T - s_{ic})$$
 (21)



Figure 4: Sample concave (valley, in red) and convex (ridge, in blue) features on depth images: (a)-(d) finger knuckle, direction 1-4; (e)-(h) palmprint, direction 1-4.



Figure 5: A pair of sample images acquired from the same subject in the contactless 3D finger knuckle database.

where T is the threshold; K is the number of searches for non-enrolled images; M is the number of searches for enrolled images; s_{i1} is the score of rank 1 of i^{th} search; s_{ic} is the score of the true class of i^{th} search; H is the unit step function.

4.1. Validation using Contactless 3D Finger Knuckle Database

The HKPolyU 3D finger knuckle images database [33] provides two-session 3D images from 130 subjects. For the verification and identification experiments, 105 subjects with two sessions are used. An evaluation protocol for two sessions' data is adopted, i.e. first session data is used as the training sets and second session data is used as the testing sets. This protocol generates $105 \times 6 = 630$ genuine and $105 \times 6 \times 104 = 65520$ imposter matching scores.

We first present the experimental result of ablation studies in figure 6 showing that the feature templates of each direction can be used alone for personal recognition. However, when using the two templates from direction 1 and 2 together, the recognition performance is better than using each of the template alone. Furthermore, when using the four templates together, the recognition performance is the best.

The effectiveness of our proposed method is validated



Figure 6: Experimental results of ablation studies on the contactless 3D finger knuckle database.

by comparing with the baseline methods. We have selected two baseline methods which were developed for extracting 3D surface features for palmprint recognition. These two methods (*Surface Code* [23] and *Binary Feature* [20]) have been reported to be effective in extracting surface valley and ridge patterns and therefore it is reasonable to adopt these methods as the baselines. The depth images required for these methods are computed using Frankot Chellappa algorithm [28] and the parameters are optimized for achieving the best performance.

The two session images from this database has been acquired using two different camera lenses, so that those images contain a large variance of statistical distribution. Therefore, it is challenging for learning based approaches. For example, figure 5 shows a pair of sample images from the same subject in this dataset. We can observe that these images contain the same biometric pattern but are visually very different. From our preliminary experimentations on learning based methods, including *VGGNet* [29] and *ResNet* [30], those methods can learn very well on the



(a) (b) (c) (c) FNIR versus FPIR.

gallery samples (i.e. the first session images) but perform poorly on the probe samples (i.e. the second session images). Therefore, those methods are not selected as baselines for the performance comparison in this section.

Figure 7 shows the comparative experimental results using our method and the two baseline methods. For verification scenarios (Figure 7(a)), our proposed method significantly outperforms both baselines. The improvement of performance is more obvious for tight security situations, where the false acceptance rate is very small. For closed-set identification scenarios (Figure 7(b)), the performance of our method and Surface Code are similar, while our method produces a higher rank-1 accuracy. However, unenrolled identities may be presented to practical systems, therefore evaluation of identification rates with open set produces more reliable results for practical situations. For this evaluation (Figure 7(c)), 105 subjects (80%) are considered as enrolled users while another 25 subjects (20%) are considered as unenrolled users. Our method outperforms both baselines by producing a lower false negative identification rate. Similar to the ROC curves, the improvement of performance is more obvious for tight security situations.

4.2. Validation using Contactless 3D Palmprint Database

The HKPolyU Contactless 2D/3D Palmprint Database [23] provides 1770 palmprint images acquired from 177 subjects in two sessions. Five 3D images are available for each subject per session. Figure 8 shows a pair of sample images from the same subject in this dataset. The evaluation protocol generates 885 (177 \times 5) genuine and 155760 (177 \times 176 \times 5) imposter matching scores. For the open set identification evaluation, 142 subjects (80%) are considered as enrolled users while the remaining 35 subjects (20%) are considered as unenrolled users.

The required surface normal images for the proposed method are computed by simply taking the gradient from the available 3D depth images in this database. Our method is compared with three state-of-the-art methods, *Binary Feature* [20], *Surface Code* [23], collaborative representa-



Figure 8: A pair of sample images acquired from the same subject in the contactless 3D palmprint database.

tion with L2-norm regularizations (CR_L2) [27]. Similar to the finger knuckle experiments, VGGNet [29] and ResNet [30] do not perform well on the probe samples (i.e. the second session images) and are therefore not selected as the promising baseline for the performance comparison in this dataset. It can be observed that the experimental results presented in [20] incorporates a masking procedure. However, the details are not clearly described in the paper. In order to ensure fairness in comparison, our evaluations on all methods and our method are without masks. The parameters are also optimized for achieving the best performance. For *CR_L2*, we first investigate the variations between their reported database and our evaluated database. Since both databases provide 3D depth images with the same resolution (square size images with 128 pixels), it is reasonable to employ the same parameters provided along with CR_L2, which is already optimized for their reported database.

Figure 9 presents the comparative experimental results using our method and the four baseline methods. Our method generally outperforms all baseline methods in verification (Figure 11(a)), closed-set identification (Figure 11(b)), and open-set identification (Figure 11(c)) experiments. These results indicate that our proposed approach using surface normal images is effective for both 3D finger knuckle and 3D palmprint recognition. Additional experimental results using statistical significant tests for the area under ROC curve with the method described in a reference



(a) (b) (c) (c) Figure 9: Experimental results on the contactless 3D palmprint database: (a) ROC; (b) CMC; (c) FNIR versus FPIR.

[31] are presented in the supplementary file.

5. Conclusions and Further Work

This paper introduces a 3D feature extraction approach which utilizes both texture and direction information using surface normal vectors for more accurate 3D hand biometric recognition. By considering whether the neighbouring surface normal vectors are pointing inwards to or outwards from the central vector, the irregular local 3D surface characteristics, i.e. ridge and valley patterns, which constitute to the distinctive features for biometric recognition, can be effectively encoded (the core idea is demonstrated in figure 1). The experimental results presented in this paper using verification, closed-set and open-set identification scenarios have validated the effectiveness of the proposed method for both 3D finger knuckle and 3D palmprint recognition. The learning based approaches does not offer promising performance, probably because of the challenges induced from the cross-lenses and two-session imaging condition of the dataset. Such imaging condition can simulate the realworld applications that acquire challenging images using different camera setups for the same subject in more than one-session, which generates significant variations between the distribution of gallery and probe samples. Despite deep learning based approaches are popular for various computer vision problems as well as biometric problems, a certain level of customized development is required on specific biometric problems. It is highly motivated to investigate the application of deep learning approaches into the problems addressed in this paper.

The key limitations of the proposed approach lie in its relatively bulky feature extraction procedures and thus a longer computational time. However, such a drawback can be justified by the more effective recognition performance as demonstrated in the experimental results presented in this paper. This paper provides two important insights that surface normal vectors can provide a reliable source of information for 3D biometrics, while the curvature information in diagonal directions in addition to horizontal and vertical directions is also helpful. Our future studies will investigate the necessity of each procedure and develop more advanced feature representation and matching methods. The representation of 3D surfaces can also be extended to the study of textured like patterns from human body of those acquired during medical imaging.

References

- J. Blasco, T.M. Chen, J. Tapiador and P. Peris-Lopez. A survey of wearable biometric recognition systems. *ACM Computing Surveys (CSUR)*, 49(3), p.43. 2016.
- [2] N. Suki, N. Poh, F.M. Senan, N.A. Zamani and M.Z.A. Darus. On the reproducibility and repeatability of likelihood ratio in forensics: A case study using face biometrics. In *International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 2016.
- [3] F. Schroff, D. Kalenichenko and J. Philbin. Facenet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015.
- [4] L. N. Darlow and B. Rosman. Fingerprint minutiae extraction using deep learning. In 2017 IEEE International Joint Conference on Biometrics (IJCB). 2017
- [5] K. H. M. Cheng and A. Kumar. Revisiting Outlier Rejection Approach for Non-Lambertian Photometric Stereo. *IEEE Transactions on Image Processing*, 28(3), pp.1544-1555, 2019.
- [6] J. Daugman. Evolving methods in iris recognition. IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS), 2016.
- [7] L. Ghiani, G. L. Marcialis and F. Roli. Fingerprint presentation attacks detection based on the user-specific effect. In 2017 IEEE International Joint Conference on Biometrics (IJCB). 2017.
- [8] Weblink for downloading implementation codes. http://www4.comp.polyu.edu.hk/~csajaykr/knuckle_ijcb20.zip
- [9] W. Jia, B. Zhang, J. Lu, Y. Zhu, Y. Zhao, W. Zuo and H. Ling. Palmprint recognition based on complete direction representation. *IEEE Transactions on Image Processing*, 26(9), pp.4483-4498. 2017.
- [10] L. Fei, B. Zhang, Y. Xu and L. Yan. Palmprint recognition using neighboring direction indicator. *IEEE Transactions on*

Human-Machine Systems, 46(6), pp.787-798. 2016.

- [11] C. Ravikanth and A. Kumar. Biometric authentication using finger-back surface. *IEEE Conference on Computer Vision* and Pattern Recognition (CVPR), 2007.
- [12] D. L. Woodard and P. J. Flynn. Finger surface as a biometric identifier. *Computer Vision and Image Understanding*, vol. 100, no. 3, pp. 357-384, Dec 2005.
- [13] L. Zhang, L. Zhang, and D. Zhang. Finger-knuckle-print: a new biometric identifier. *IEEE International Conference on Image Processing (ICIP)*, 2009.
- [14] A. Kumar and Y. B. Zhou. Human identification using KnuckleCodes. *IEEE International Conference on Biomet*rics: Theory, Applications and Systems (BTAS), 2009.
- [15] P. Yan and K. W. Bowyer. Biometric recognition using 3D ear shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1297-1308, Aug 2007.
- [16] I.I. Ganapathi and S. Prakash. 3D ear recognition using global and local features. *IET Biometrics*, 7(3), pp.232-241. 2018.
- [17] A. S. Mian, M. Bennamoun, and R. Owens. An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1927-1943, Nov 2007.
- [18] K. I. Chang, K. W. Bowyer, and P. J. Flynn. An evaluation of multimodal 2D+3D face biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 619-624, Apr 2005.
- [19] X. Zhu, Z. Lei, X. Liu, H. Shi and S.Z. Li. Face alignment across large poses: A 3d solution. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
- [20] Q. Zheng, A. Kumar, and G. Pan. Suspecting less and doing better: new insights on palmprint identification for faster and more accurate matching. *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 633-641, Mar 2016.
- [21] A. Kumar and C. Kwong. Towards contactless, low-cost and accurate 3D fingerprint identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 681-696, Mar 2015.
- [22] P. Grother and M. Ngan. The IJB-A Face Identification Challenge Performance Report. In *Technical report. National Institute of Standards and Technology*, 2017.
- [23] V. Kanhangad, A. Kumar, and D. Zhang. A unified framework for contactless hand verification. *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1014-1027, Sep 2011.
- [24] C. Dorai and A. K. Jain. COSMOS A representation scheme for 3D free-form objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 10, pp. 1115-1130, Oct 1997.
- [25] J. J. Koenderink and A. J. Vandoorn. Surface shape and curvature scales. *Image and Vision Computing*, vol. 10, no. 8, pp. 557-564, Oct 1992.
- [26] K. H. M. Cheng and A. Kumar. Advancing Surface Feature Encoding and Matching for More Accurate 3D Biometric Recognition. In *International Conference on Pattern Recognition (ICPR)*, 2018.
- [27] L. Zhang, Y. Shen, H. Li, and J. Lu. 3D Palmprint Identification Using Block-Wise Features and Collaborative Represen-

tation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp. 1730-1736, Mar 2015.

- [28] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 439-451, Jul 1988.
- [29] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-scale Image Recognition. *International Conference on Learning Representations (ICLR)*, 2015.
- [30] K. He and X. Zhang and S. Ren and J. Sun. Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [31] E. R. DeLong, D. M. DeLong and D. L. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, pp. 837-845, 1988.
- [32] L. Fei, B. Zhang, Y. Xu, W. Jia, J. Wen and J. Wu. Precision direction and compact surface type representation for 3D palmprint identification. *Pattern Recognition*, 87, pp.237-247. 2019.
- [33] K. H. M. Cheng and A. Kumar. Contactless biometric identification using 3D finger knuckle patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(8), pp. 1868-1883, 2020.
- [34] Z. Zhao and A. Kumar. Towards More Accurate Iris Recognition Using Deeply Learned Spatially Corresponding Features. In *International Conference on Computer Vision* (*ICCV*), pp. 1-10, Venice, Italy, Italy, October 2017.
- [35] J. J. Engelsma, K. Cao and A. K. Jain. Learning a Fixed-Length Fingerprint Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2019.