Efficient and Accurate 3D Finger Knuckle Matching using Surface Key Points

Kevin H. M. Cheng, Ajay Kumar

***Abstract*—Contactless 3D finger knuckle is a new biometric identifier with a lot of potentials, which can provide an accurate, efficient and convenient alternative for the personal identification. The current 3D finger knuckle recognition methods are limited by computationally complex or inefficient matching algorithms, which attempt to compute the matching scores from all possible translational and rotational parameters for matching a pair of templates. The strength of such approach lies in its simplicity and reliability for accurately matching intra-class samples, but expensive computational time is required. Furthermore, attempting on excessive numbers of translational and rotational parameters can also degrade the overall recognition accuracy because the imposter matches can be increased. In fact, this conventional matching approach is commonly adopted in many biometric studies, but its drawbacks have not received adequate attention. This paper addresses such 3D finger knuckle recognition problem by developing a more efficient matching approach using surface key points extracted from 3D finger knuckle surfaces. Our comparative experimental results with the state-of-the art method on a publicly available 3D finger knuckle database indicates that our approach can offer over 23 times faster with performance improvement on the accuracy. Although the focus of our work is on 3D finger knuckle recognition, we also present the performance of our method on other publicly available databases with similar 3D biometric patterns including 3D palmprint and 3D fingerprint, to validate the effectiveness of the proposed approach.**

***Index Terms*—hand biometrics, 3D finger knuckle recognition, key points extraction, templates matching**

I. Introduction

BIOMETRIC recognition using finger knuckle images [1-4] has gained increasing attention in recent years probably due to its recognition accuracy, reasonable efficiency, and the high convenience of acquiring finger dorsal images. This technology provides a wide range of potential civilian applications such as immigration inspection, unlocking smartphones and online shopping. The recent research trend investigates the 3D information of biometrics in addition to the convention studies on 2D intensity images, because rich information can be extracted from 3D images while such images are usually more robust and illumination invariant. The success of such research direction can be observed in many popular biometrics such as 3D fingerprint [5], 3D palmprint [6-8], 3D face [9-10] and 3D ear [11-12]. Similarly, the use of 3D finger knuckle images for biometric recognition has been investigated in a recent research [13].

TABLE I
COMPARATIVE SUMMARY BETWEEN OUR METHOD AND THE RECENT ADVANCES ON 3D HAND BIOMETRICS

Biometric Trait	Accuracy	Efficiency
3D Finger Knuckle	**More Accurate**	**~23X Faster**
3D Palmprint	More Accurate	~2X Faster
3D Fingerprint	Similar	~2X Faster

The research on 3D finger knuckle recognition [13] introduced the reliable surface gradient derivative features extracted from 3D finger knuckle images as the templates for further matching. During the matching process, a pair of templates are matched by a trial and error approach, which computes the scores from matching a gallery template with all possible combination of translational and rotational shifting of the probe templates. The shifting parameters producing the minimum score is considered to be the best shifting parameters and that matching score is considered as the final matching score. In fact, this matching approach has been widely employed in many biometric researches including finger knuckle recognition [1, 13], palmprint recognition [7-8, 14-16] and iris recognition [17-18] due to its effectiveness and simplicity. However, there are two drawbacks associated with such approach: (1) it is computationally expensive because all possible combinations of translational and rotational shifting are being evaluated; (2) the overall recognition accuracy is reduced because when attempting on many trails, the final scores of matching the imposter pairs will be closer to the final scores of matching the genuine pairs, which results in a situation that the distribution of the imposter matching scores is shifted towards the distribution of the genuine matching scores. These two drawbacks are especially obvious for finger knuckle recognition because of the following reason: since the area of interest of finger knuckle is loosely defined, a sharp boundary does not exist along the finger knuckle regions; therefore, even with a robust detector for segmenting finger knuckle images, a large degree of translational and rotational shifting is required to align a pair of segmented finger knuckle images; a large number of attempts is required for matching the finger knuckle images, therefore the effects of these drawbacks are more severe. These limitations of the matching approach are usually

Authors are with the Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

overlooked by the research community. Therefore, it is motivated to investigate a more efficient matching approach, especially for the 3D finger knuckle recognition problem.

This paper attempts to mitigate the two aforementioned drawbacks by introducing a new matching approach using reliable surface key points. Those surface key points can be extracted from 3D finger knuckle images using the reliable surface gradient derivative features. By simply using the spatial location of key points from a pair of templates, the number of attempts for the matching a pair of templates can be largely reduced. As a consequence, the computational efficiency can be largely enhanced while the overall recognition accuracy can also be slightly improved because the shifting of the distribution of imposter matching scores towards the distribution of genuine matching scores is mitigated. The key contributions of this paper can be summarized as follows:

(i) This paper develops a new matching approach using surface key points extracted from 3D finger knuckle images for more efficient and accurate matching of 3D finger knuckle templates. It can be observed that the conventional matching approach, attempting to compute the matching scores for all possible combinations of translational and rotational shifting, is computationally complex. Such approach also limits the overall recognition accuracy because computing the minimum scores from many attempts results in the shifting of the distribution of imposter matching scores towards the distribution of genuine matching scores. These issues on the finger knuckle recognition problem worth special attentions because of the nature of this biometrics. Since the finger knuckle area of interest is loosely defined, i.e. a sharp boundary for finger knuckle does not exist, almost all segmented finger knuckle images are largely misaligned with other images from the same subject. In order to match those genuine samples, a large degree of translational and rotational shifting is required. Therefore, a large number of matching attempts are required if the conventional matching approach is adopted, which results in severe effects on the degradations of efficiency and accuracy. Our matching approach utilize the discriminative feature points extracted from 3D finger knuckle images using parts of the surface gradient derivative computation introduced in [13]. Those key points provide promising clues for estimating the final shifting parameters, hence largely reducing the shifting parameter space for the trails. The effectiveness of our method is evaluated from the comparative experimental results with the state-of-the-art method on a publicly available 3D finger knuckle database, with a matching time of 23 times faster and more accurate than the conventional approach.

(ii) Since the conventional matching approach is widely employed in other similar biometric problems such as palmprint recognition, we further evaluate the effectiveness of our method on other 3D biometrics including 3D palmprint and 3D fingerprint. The comparative experimental results on two other publicly available databases of 3D palmprint and 3D fingerprints indicates that our method is generalizable and effective on other 3D biometrics. However, such performance improvement is not as significant as the 3D finger knuckle recognition problem, because our surface key points are extracted based on the 3D finger knuckle literature.

This important area investigated in this paper has not yet attracted adequate attention but is crucial to the deployment of biometric systems. Further enhancing the surface key points extraction from respective discriminative features including finger knuckle, palmprint and fingerprint, is also a promising future research direction. Table I shows the comparative summary between our proposed matching approach over the conventional matching approach, with the same state-of-the-art feature descriptor, on various 3D hand biometrics. This summary is supported by the comprehensive experimental results presented in Section IV.

Rest of this paper is organized as follows: Section II discusses the related work of the 3D finger knuckle recognition technology. Section III presents the theoretical detail of our proposed surface key points extraction and matching approach. Section IV presents the comparative experimental results on three publicly available databases of 3D finger knuckle, 3D palmprint and 3D fingerprint. Finally, Section V summarizes this paper and discuss the potential future research directions. The implementation codes for the proposed approach are also provided along with this paper [40].

II. Related Work

Contactless 3D finger knuckle recognition is a new research frontier. It was first briefly studied together with using finger dorsal surfaces for biometric recognition [19]. This research does not arouse much attention for the area of 3D finger knuckle recognition probably because the effectiveness of using 3D finger knuckle information was limited by the low resolution of 3D finger knuckle images being studied and the ineffective feature descriptor employed for extracting 3D finger knuckle features, which is a generic surface descriptor, the Shape Index [20-21]. The attention on 3D finger knuckle recognition has been raised by a recent research [13] which investigates the use of 3D finger knuckle patterns for biometric recognition. This study investigated various important aspects of 3D finger knuckle recognition, such as the discriminative feature description, the possibility of spoofing attacks towards a finger knuckle recognition system, the individuality of finger knuckle, the comparisons between 2D and 3D finger knuckle recognition, and provides a benchmark database for further research and investigation. The effectiveness and potentials of using 3D finger knuckle images for biometric recognition has been validated. However, there are other open questions on using this new biometric trait such as the matching approach addressed in this paper.

Since the studies on 3D finger knuckle recognition is new, it is mainly supported by the literatures of finger knuckle recognition [1-4] and palmprint recognition [6-8, 14-16, 33-34]. A survey paper [3] summarized many previous researches on finger knuckle recognition. However, it is shown in [13] that a state-of-the-art method on 2D palmprint recognition, *Difference of Normal* (*DoN*) [14], outperforms other methods such as *Fast-RLOC* and *Fast-Comcode* [8] for the 2D finger knuckle recognition evaluation. On the other hand, *Surface Code* [7] and *Binary Shape* [8] are the two baseline methods for

the performance comparison of 3D finger knuckle recognition in [13]. Both methods are originally developed for 3D palmprint recognition, which also compute binary matching templates from 3D depth images. Unlike these two methods, the surface gradient derivative (*SGD*) feature descriptor [13] computes binary matching templates from 3D surface normal images. It is currently the state-of-the-art method for 3D finger knuckle recognition.

Besides the conventional hand crafted feature approaches, deep learning based methods have been actively developed for many applications in computer vision related tasks such as object recognition [22-24], instance segmentation [25-26], action recognition [37-38] as well as biometric recognition [18, 39]. Although deep learning based approaches can offer promising performance, its application on specific biometric problem requires customized development. For example, the success of a recent iris recognition research [18] requires the considerations of biometric aspects including the use of binary templates and the bit-shifting strategy for matching the templates. The incorporation of deep learning based approaches for the biometric problem addressed in this paper will be a promising future extension.

Lastly but not least, it is also worth mentioning that, the conventional matching approach, which takes the minimum score for matching the gallery template with the translationally and rotationally shifted versions of the probe templates, is widely adopted in many biometric studies including finger knuckle recognition [1, 13], palmprint recognition [7-8, 14-16] and iris recognition [17-18]. The drawbacks of this approach are usually overlooked, probably because they are not quite severe. Furthermore, although this matching problem seems to be like the classical point correspondence problem, the objectives of these two problems are quite different. The matching problem addressed in this paper is less complex than the point correspondence problem which requires the consideration of pixel-to-pixel alignment due to the deformation of images. Therefore, we attempt to introduce a simple and efficient solution for this less complicated problem. In the best of our knowledge, there is no further advancement on this 'bit-shifting' approach for matching binary templates. This paper focuses on mitigating these drawbacks by introducing a new matching approach using surface key points.

III. 3D Finger Knuckle Matching using Key Points

This section presents theoretical details of our proposed matching approach. We begin the discussion from theoretically formulating the matching problem in Section III.A. The technical details of the detection of finger knuckle surface key points will be presented in Section III.B. The estimation of shifting parameters consisting of translation and rotation will be discussed in Section III.C and Section III.D.

A. Problem Formulation

This paper attempts to address a sub-problem, matching a pair of finger knuckle templates, from the 3D finger knuckle recognition problem. To begin our discussion, we first formulate the matching problem as follows. The objective of this problem is to compute both the shifting parameters and the matching score resulted from the best shifting parameters. The matching score s results from matching a pair of feature templates $\boldsymbol{A}$ and $\boldsymbol{B}$ with a dimension of $m_i \times n_i \times c$, where m_i and n_i are the spatial dimension of the image templates, c is the channel dimension. In order to match these two templates, one effective way is to extract a smaller window with dimension $m_w \times n_w$ from both original image templates, such that the matching window $\boldsymbol{A'}$ is the center rectangular region of $\boldsymbol{A}$, while the matching window $\boldsymbol{B'}$ is obtained from the center rectangular region of $\boldsymbol{B}$ with translational and rotational shifts. The matching score s between the pair of templates can be simply represented by a general matching function f:

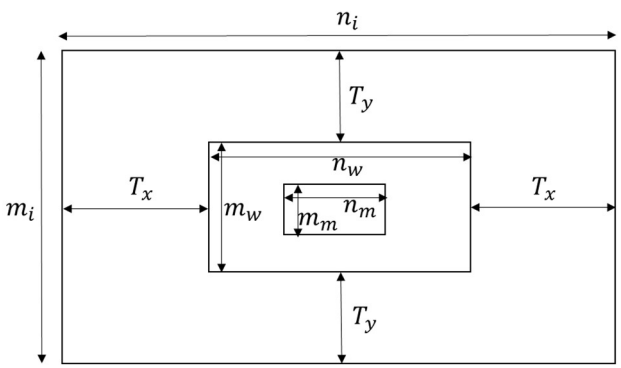


Fig. 1. A schematic diagram with the dimension of the image templates, matching windows and the mask for finger knuckle key points detection.

$$s = \min_{t_x,t_y,r_d} f(\boldsymbol{A'}, \boldsymbol{B'}(t_x, t_y, r_d)),$$
$$t_x \in [-T_x, T_x], t_y \in [-T_y, T_y], r_d \in [-R_d, R_d] \quad (1)$$

where t_x is the translational shift parameter in horizontal direction; t_y is the translational shift parameter in vertical direction; r_d is the rotational shift parameter; T_x is the maximum number of pixels for the translational shift in horizontal direction; T_y is the maximum number of pixels for the translational shift in vertical direction; R_d is the maximum degrees for the rotational shift. The dimension of the matching window is constrained by the maximum number of pixels for the translational shift:

$$m_w = m_i - 2T_y \ , n_w = n_i - 2T_x \quad (2)$$

There are two implications from this equation. Firstly, if larger translational shifts T_x and T_y are required to tolerate the misalignment from the segmented finger knuckle images, the dimension of the matching window must be smaller. Secondly, if a larger matching window is required to accommodate more finger knuckle features and information, the maximum number of pixels for the translational shifting must be smaller. This constraint equation can be visualized in Figure 1 showing a schematic diagram with the dimension of the image templates and matching windows.

The conventional approach simply computes the matching scores using all the possible combination from the parameter space with a dimension of $(2T_x + 1) \times (2T_y + 1) \times (2R_d + 1)$. The parameter set $\{t_x, t_y, r_d\}$ which produces the minimum dissimilarity score is considered as the final shifting parameter set and that score is considered as the final matching score.

While increasing the number of matching attempts, this trial-and-error approach may produce smaller genuine matching scores on one side, but it may also produce smaller imposter matching scores on the other side. When the number of attempts is adequate for matching the genuine pairs, further increasing the number of attempts do not decrease the genuine matching scores but only decrease the imposter matching scores. As a result, the separation between the distribution of genuine matching scores and that of imposter matching scores decreases and the overall recognition accuracy drops. Therefore, excessive number of trial-and-error attempts decreases the recognition performance. More importantly, excessive number of trial-and-error attempts reduces the computational efficiency, i.e. the matching time is proportional to the number of trial-and-error attempts.

In this paper, we attempt to reduce the number of attempts, i.e. the parameter space, by wisely guessing the more likely shifting parameters from the clues of reliable finger knuckle surface key points. The reduction of the number of attempts is expected to reduce the computational time as well as the overall recognition errors.

B. Detection of Finger Knuckle Surface Key Points

In order to effectively reduce the number of trial-and-error attempts, it is crucial to have a reliable finger knuckle key point detector. It is well known that key points detected from fingerprint, also known as fingerprint minutiae, are the key component for the success of fingerprint recognition methods [28]. Research on the investigation of using palmprint minutiae for improved recognition [29-30] are also emerging. It is motivated to investigate the use of finger knuckle minutiae for improved recognition. A study [31] attempted to derive finger knuckle minutiae for finger knuckle recognition. However, the effectiveness of this method has not been demonstrated from the recognition performance. Unlike fingerprint and palmprint, finger knuckle pattern is located right above the joint connecting the middle and proximal phalanges of fingers, which is in motion frequently. The finger knuckle pattern consists of irregular ridges and valley regions with varying thickness. Those effective minutiae features defined for fingerprint such as ridge ending and bifurcation may not be suitable for finger knuckle pattern. It is challenging to define robust finger knuckle minutiae for improved finger knuckle recognition.

Since an effective finger knuckle minutiae definition has not yet been developed, we attempt to investigate the detection of reliable finger knuckle key points from the perspective of reliable feature representation. A recent 3D finger knuckle research [13] introduced a reliable feature descriptor, *SGD*, for 3D finger knuckle recognition. This feature descriptor employs the surface gradient derivative features and offers outperforming recognition performance. Therefore, it is reasonable to believe that the surface gradient derivative features are robust and discriminative for the context of 3D finger knuckle. We attempted to define reliable finger knuckle key points from employing the surface gradient derivative features.

Theoretically, surface gradient derivative features encode the concavity of a finger knuckle surface. For a pair of finger knuckle feature templates belonging to the same person, the feature templates are expected to be similar due to its similar surface pattern. However, it is difficult to locate a single pair of corresponding points from the feature templates because the imaging of finger knuckle pattern is usually distorted by various types of noises. Therefore, we attempt to locate a set of points from the region containing the most discriminative finger knuckle patterns, which is expected to be associated with the sharpest finger knuckle lines. Those lines may correspond to the most concave region, which can be obtained from the minimums from the surface gradient derivative computation.

The surface gradients p and q can be computed from either the 3D surface normal images or the 3D depth images. If the unit surface normal vector for a pixel in the form of $\hat{\boldsymbol{n}} = [n_x, n_y, n_z]^T$ is known, the surface gradients in horizonal and vertical directions p, q can be computed as:

$$p = \frac{n_x}{n_z} \quad , \quad q = \frac{n_y}{n_z} \tag{3}$$

On the other hand, if the finger knuckle depth surface z, described in terms of a function along 2D coordinates x and y, is known:

$$z = g(x, y) \tag{4}$$

The surface gradients p, q can also be computed as the gradient along horizontal and vertical directions respectively:

$$p = \frac{\partial g(x,y)}{\partial x} \quad , \quad q = \frac{\partial g(x,y)}{\partial y} \tag{5}$$

After obtaining the surface gradients, the derivatives of surface gradient variables p, q are then computed as in the following:

$$\frac{\partial p}{\partial x} = \frac{\partial^2 g(x,y)}{\partial x^2} \quad , \quad \frac{\partial q}{\partial y} = \frac{\partial^2 g(x,y)}{\partial y^2} \tag{6}$$

We present a general response function h for the detection of finger knuckle key point with coordinate x, y. The finger knuckle key points detection problem can be formulated as follows:

$$\arg\min_{x,y} h(x, y) \tag{7}$$

Since we wish to locate a set of points from the region containing the most discriminative finger knuckle patterns, which correspond to the minimums from the surface gradient derivative features, we can assign the response function h to be $\frac{\partial p}{\partial x}$ and $\frac{\partial q}{\partial y}$ separately. Ideally, when assigning h to be $\frac{\partial p}{\partial x}$, the optimal solution is a point with coordinate (u_x, v_x), so that this point can effectively correspond to the other point detected from another templates belonging to the same person. However, due to the presence of various noise in the reality, it is judicious to consider more key points for the proceeding to the matching attempts. When also assigning the response function h to be $\frac{\partial q}{\partial y}$ separately, the optimal solution produces another point with coordinate (u_y, v_y). Similarly, we can also extract more key points from the second, third and fourth minimums and so on, to produce k_p number of key points using $\frac{\partial p}{\partial x}$ and $\frac{\partial q}{\partial y}$ separately. The set of finger knuckle key points can be represented as:

$$\left\{ \left(u_x^j, v_x^j\right), \left(u_y^j, v_y^j\right) \mid j \in [1, k_p] \right\} \tag{8}$$

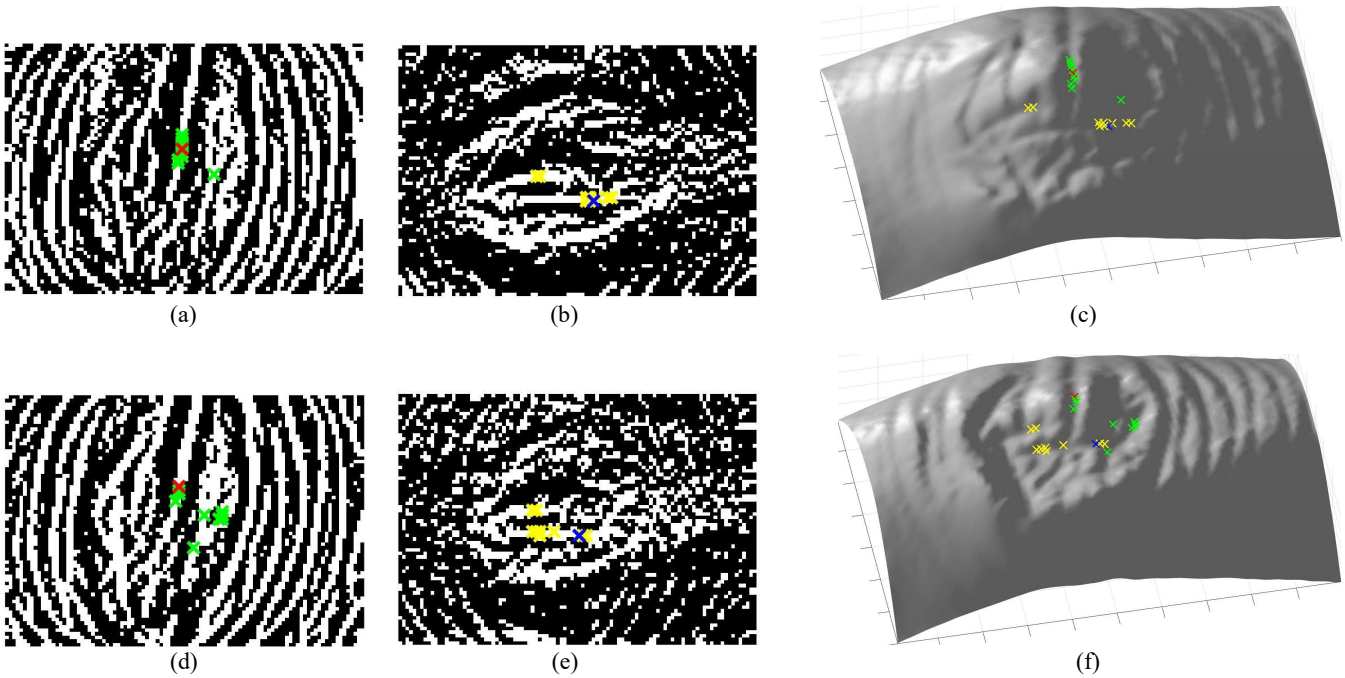


Fig. 2. Pairs of genuine images showing detected finger knuckle surface key points on: (a)/(d) *SGD* feature templates of the horizontal direction; (b)/(e) *SGD* feature templates of the vertical direction; (c)/(f) 3D depth images.

From equation (2), since there is a limit for the maximum number of pixels for translational shifting, we can also constraint the region of detecting finger knuckle key points. For example, if a key point is detected at the left most pixel in template $\boldsymbol{A}$ while another key point is detected at the right most pixel in template $\boldsymbol{B}$, the estimated translational shift to the right is approximately the width of the template, which must exceed the maximum number of pixels for translational shifting and is unnecessary to include in the parameter space. Therefore, we attempt to constraint the key points falling in the center regions defined by a rectangular mask. Furthermore, the center region of the segmented finger knuckle images usually contains rich discriminative information. It is reasonable to detect key points from that region. The dimension of the rectangular mask is constrained by:

$$|m_m - 1| \leq T_y \quad , \quad |n_m - 1| \leq T_x \tag{9}$$

The relationship between the mask for finger knuckle key points detection and the maximum number of pixels for the translational shift can also be observed from Figure 1. This constraint will be further explained in the following subsection after the presentation of how to match a pair of templates using these key points. In order to illustrate the key points detection step, we present a pair of genuine image templates showing detected finger knuckle surface key points on *SGD* templates and 3D depth images in Figure 2. The red crosses represent the first minimum resulted from the response function $h = \frac{\partial p}{\partial x}$, while the green crosses represent the second to the tenth minimums. Similarly, the blue crosses represent the first minimum resulted from the response function $h = \frac{\partial q}{\partial y}$, while the yellow crosses represent the second to the tenth minimums. It can be observed that the red/blue points between these pair of templates corresponds to each other quite accurately, while the green/yellow points provide more tolerance for the inaccurate detection situations.

C. *Estimation of Translational Shifting Parameters*

The finger knuckle surface key points are detected from the method described in the last sub-section. This sub-section presents the process of estimating translational shifting parameters for matching a pair of finger knuckle templates from the key points. Suppose there are k_p points extracted from using the two response functions $\frac{\partial p}{\partial x}$ and $\frac{\partial q}{\partial y}$ separately, we attempt to match all k_p points with each other for each response function separately, resulting in $2{k_p}^2$ combinations. The set of possible combination for the translational shifting parameters are represented as follows:

$$\left(t_x^j, t_y^j\right) = \{\left(u_x^{j_A} - u_x^{j_B}, v_x^{j_A} - v_x^{j_B}\right), \left(u_y^{j_A} - u_y^{j_B}, v_y^{j_A} - v_y^{j_B}\right)\} \tag{10}$$

where $j \in \left[1, {k_p}^2\right]$, $j_A, j_B \in \left[1, k_p\right]$. Since some combinations from the key points may refer to the same translational shifting parameters, those duplicate parameters are automatically removed. The final number of matching attempts for translational shifting is generally much less than $2{k_p}^2$.

In order to further explain the constraint of the mask for the detection of finger knuckle surface key point in equation (9), we can consider the extreme cases. For example, if a key point is detected at the right most pixel within the mask in template $\boldsymbol{A}$, while another key point is detected at the left most pixel within the mask in template $\boldsymbol{B}$, the translational shifting parameter in horizontal direction will be $n_m - 1$. Inversely, if a key point is detected at the left most pixel within the mask in template $\boldsymbol{A}$, while another key point is detected at the right most pixel within the mask in template $\boldsymbol{B}$, the shifting parameter will

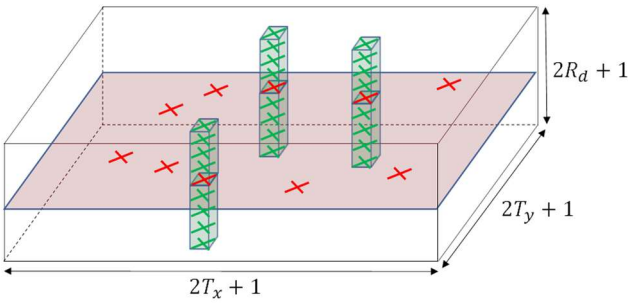


Fig. 3. A schematic diagram showing the parameter space in the domain of translational and rotational shifting.

be $1 - n_m$. From the constraint described in equation (9), any combination of translational shifting parameters in horizontal direction resulting from the key points within the mask must be within the respective parameter space, i.e. $[-T_x, T_x]$. Similarly, any combination of translational shifting parameters in vertical direction resulting from the key points within the mask must be within the respective parameter space, i.e. $[-T_y, T_y]$.

D. Two-stage Translational and Rotational Matching

The finger knuckle key points are helpful for estimating the possible translational shifting parameters. However, it is difficult to estimate an accurate rotational shifting parameter. In order to address the two drawbacks described in the introduction section, we attempt to develop a two-stage translational and rotational matching approach. Since rotational misalignment is usually less severe than translational misalignment in segmented finger knuckle images, we can address the matching problem with translational attempts first, followed by rotational attempts.

In the first stage, we assume there is no rotational misalignment between a pair of feature templates. These feature templates are matched by the parameter sets introduced in Section III.C, which is generally much less than $2{k_p}^2$. In the second stage, we select k_s number of the translational shifting parameter sets (those obtaining minimum scores from 1st stage) from less than $2{k_p}^2$ as candidates for the trials of rotational matching. Theses k_s parameters are combined with all possible combinations of rotational shifting parameter, i.e. $[-R_d, R_d]$. The estimation of translational shifting parameter reduce the number of trial-and-error attempts from $(2T_x + 1) \times (2T_y + 1) \times (2R_d + 1)$ to less than $2{k_p}^2 \times (2R_d + 1)$ while the two-stage translational and rotational matching approach further reduce the number of trial-and-error attempts to less than $2{k_p}^2 + k_s \times (2R_d + 1)$ where k_s is much less than $2{k_p}^2$. Figure 3 shows a schematic diagram of the parameter space of the number of trial-and-error attempts in the dimension of translational and rotational domain. The red crosses refer to the trials of translational matching in stage 1, e.g. 10 attempts, while the green crosses/cuboid refer to the trials of rotational matching in stage 2, e.g. 3×8 attempts. Algorithm 1 summarizes the procedure to match a pair of feature templates. Line 2-5 computes k_p key points for both feature templates $\boldsymbol{A}$ and $\boldsymbol{B}$. Line 6-13 computes the translational shifting parameter set. Line 14-15 refers to the first stage (translational) matching. Line 16-18 refers to the second stage (rotational) matching.

Algorithm 1. Matching a Pair of Feature Templates

Input: $\boldsymbol{A}, \boldsymbol{B}$: a pair of feature templates;
k_p: number of key points;
k_s: number of the selected translational shifting parameters.
Output: s: final matching score.

1: **procedure** MATCH ($\boldsymbol{A}$, $\boldsymbol{B}$, k_p, k_s)
2: **for** j = 1 → k_p **do**
3: $(u_x^j, v_x^j) \leftarrow \arg\min_{x,y} h_x(x,y)$;
4: $(u_y^j, v_y^j) \leftarrow \arg\min_{x,y} h_y(x,y)$;
5: **end for**
6: **for** j_A = 1 → k_p **do**
7: **for** j_B = 1 → k_p **do**
8: $(t_x^j, t_y^j) \leftarrow \{(u_x^{j_A} - u_x^{j_B}, v_x^{j_A} - v_x^{j_B}), (u_y^{j_A} - u_y^{j_B}, v_y^{j_A} - v_y^{j_B})\}$;
9: **end for**
10: **end for**
11: **if** any parameters (t_x^j, t_y^j) are duplicated **then**
12: remove those parameters from the set;
13: **end if**
14: $r_d \leftarrow 0$; *(r_d: rotational shift parameter)*
15: match $\boldsymbol{A}$ and $\boldsymbol{B}$ using the translational shifting parameter set;
16: **for all** k_s **do**
17: match $\boldsymbol{A}$ and $\boldsymbol{B}$ using the selected translational shifting parameters with all possible r_d;
18: **end for**
19: $s \leftarrow$ the minimum score among all matching attempts;
20: **end procedure**

TABLE II
EFFECT OF k_p ON THE RECOGNITION ACCURACY AND EFFICIENCY

k_p	1	5	**10**	15	20	*SGD (TPAMI20)*
EER (%)	16.01	3.53	**2.71**	3.04	3.10	3.29
Number of Attempts	≤42	≤1050	**≤4200**	≤9450	≤16800	39627
Computational Time (s)	0.07	0.14	**0.32**	0.42	0.59	2.85

TABLE III
EFFECT OF k_s ON THE RECOGNITION ACCURACY AND EFFICIENCY

k_s	1	5	10	15	**20**	25	30
EER (%)	3.34	2.91	2.76	2.70	**2.67**	2.69	2.70
Number of Attempts	≤221	≤305	≤410	≤515	**≤620**	≤725	≤830
Computational Time (s)	0.08	0.09	0.10	0.11	**0.12**	0.13	0.13

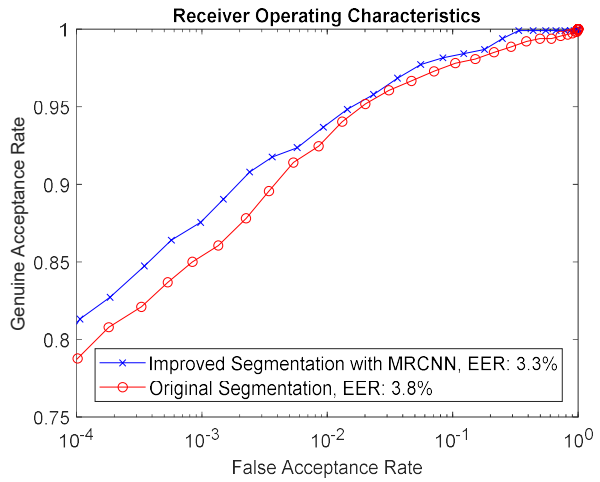
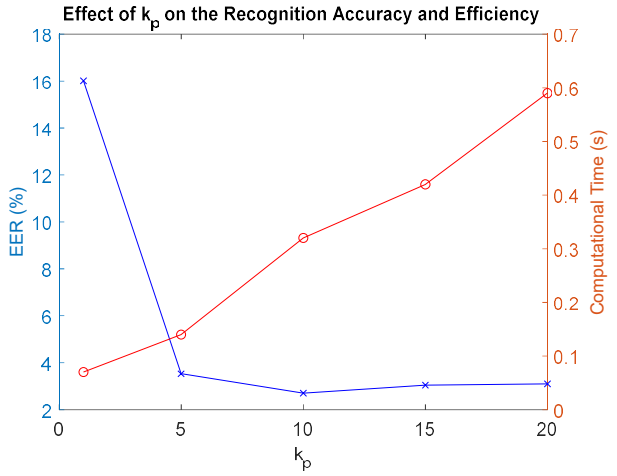
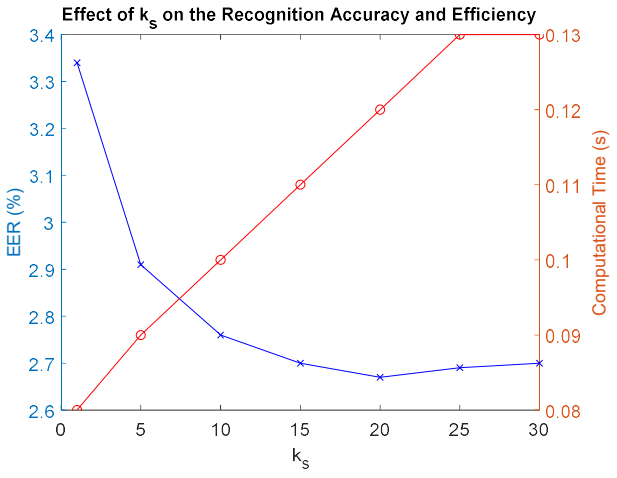


Fig. 4. Comparative Experimental Results on the 3D Finger Knuckle Dataset for the Improved Segmentation Method.

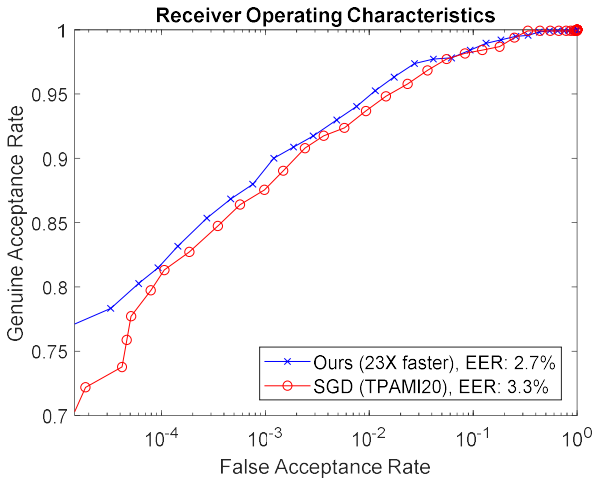


(a)

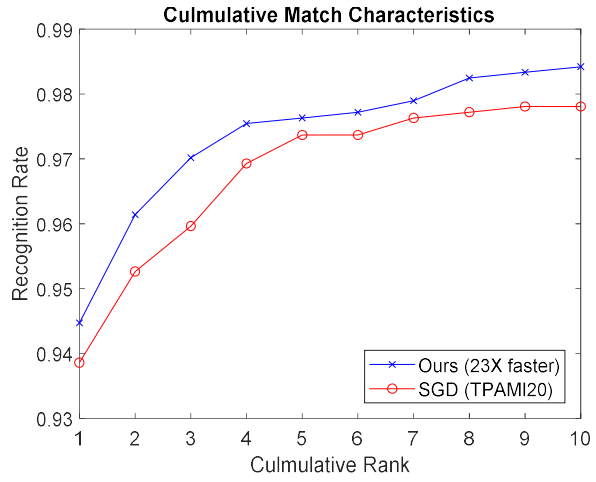


(b)

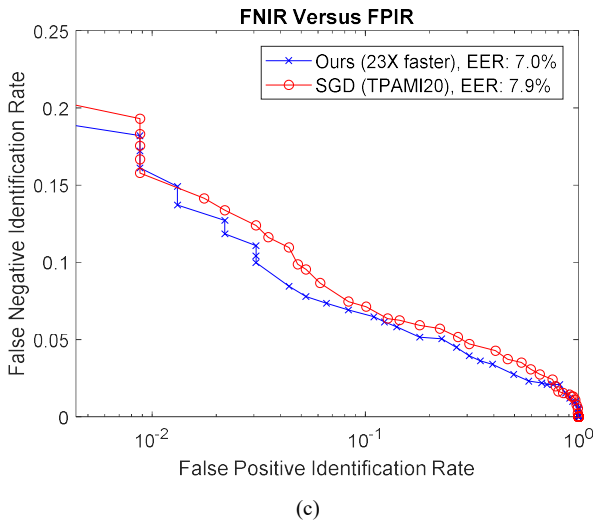
Fig. 5. Effects of k_p and k_s on the Recognition Accuracy and Efficiency.



(a)



(b)



(c)

Fig. 6. Comparative Experimental Results on the 3D Finger Knuckle Dataset: (a) ROC; (b) CMC; (c) FNIR versus FPIR.

IV. Experiments and Results

This section presents comparative experimental results from ablation studies and with state-of-the-art methods using publicly available databases. Although the focus of this paper is on 3D finger knuckle recognition, we also present additional experimental results using publicly available databases of similar 3D hand biometric patterns including 3D palmprint and 3D fingerprint, for demonstrating the potentials of our approach. We performed comprehensive experiments to

ascertain the effectiveness for the verification and identification problems. These experimental results are presented using the receiver operating characteristics (ROC) curve with equal error rates (EER), and cumulative match characteristics (CMC) curve. Furthermore, since unregistered user may be identified as the enrolled users in deployed biometric systems, such open-set identification is widely considered as the more challenging problem and therefore we also performed such evaluation in this work. These results are presented using False Negative Identification Rate (FNIR) versus False Positive Identification Rate (FPIR) curves. The experimental results presented in this paper are reproducible and indicate diversities of applications with the proposed approach.

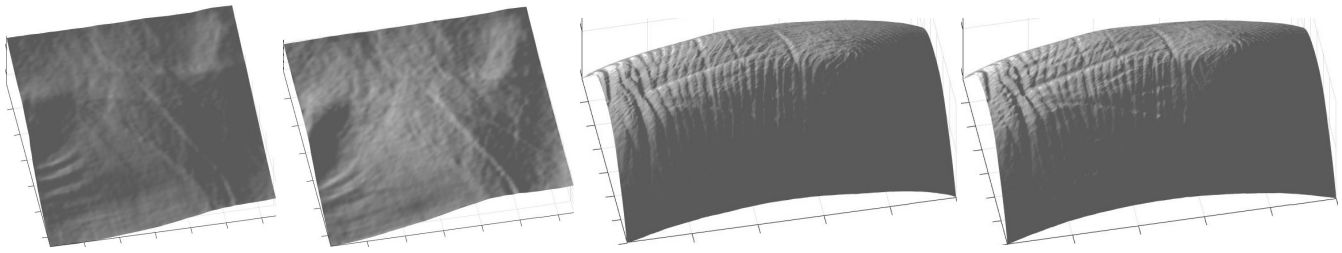
Since the proposed approach is developed for matching feature templates, we incorporate this approach with the state-of-the-art 3D finger knuckle feature descriptor [13]. This descriptor extracts the discriminative information from 3D finger knuckle surface normal images with the considerations of potential convex and concave regions corresponding to the irregular ridges and valley regions respectively.

A. Evaluation with 3D Finger Knuckle Images

The HKPolyU 3D finger knuckle images database [13] is currently the only publicly available dataset providing 3D finger knuckle images. This recently released database can be considered as a benchmark dataset for the evaluation of the performance of 3D finger knuckle recognition. This dataset provides 1410 forefinger images and 1410 middle finger images from 130 subjects, while 105 subjects contain two-session images. Since this dataset is quite small, we acquire more images from another 98 subjects. The combined dataset contains 2508 forefinger images and 2508 middle finger images from 228 subjects, while 190 subjects contain two-session images. Six forefinger images and six middle finger images are available for each subject per session. For the evaluation in this paper, we employ the forefinger knuckle images from the 190 subjects containing two-sessions images. A standard two-session evaluation protocol, which uses the first session images for the training and the second session images for the testing is adopted. This protocol generates 215460 (190×189×6) imposter matching scores and 1140 (190×6) genuine matching scores. As for the open-set identification experiments, 152 subjects (80%) are considered as enrolled users while the remaining 38 subjects (20%) are considered as unenrolled users.

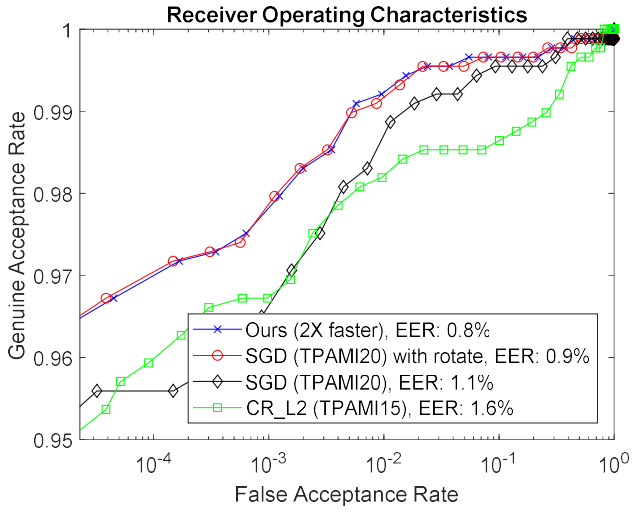
This work also attempts to improve the segmentation of finger knuckle from the earlier work. One of the reliable finger knuckle segmentation method described in [13] employed a simple edge pixel counting mechanism. Despite this method produce acceptable segmentation performance, it fails for some challenging samples which limits the finger knuckle recognition performance. Therefore, we attempted to incorporate a popular deep learning method, Mask R-CNN [25] for improving the finger knuckle segmentation. We first prepare a training set using the remaining 38 subjects containing only one-session images to ensure no overlapping for the training and testing sets for the Mask R-CNN. The ground truth masks are prepared by applying the original finger knuckle segmentation method. This method firstly computes the edge image from the original finger knuckle image, followed by counting the number of edge pixels within a fixed size sliding window. The sliding window is shifted vertically and horizontally along the image. The location where the maximum number of edge pixels within the sliding window is considered as the area of interest. Since this method may fail for some challenging samples, we inspect and remove those unsuccessful samples and utilize the remaining samples for training the Mask R-CNN with fine tuning from the COCO dataset [32]. This human inspection is only needed for producing the ground truth mask for finger knuckle segmentation while other steps are completely automatic. The segmented 3D finger knuckle images are employed for all the experiments in this sub section. Such experimental results of this ablation study are presented in figure 4.

We begin our ablation studies by evaluating the effects of the two parameters, the number of key points k_p and the number of the translational shifting parameter candidate k_s, on both the recognition accuracy and efficiency. The recognition accuracy is represented by the EER from the ROC curves, while the efficiency is represented by the number of trial-and-error attempts as well as the *average* computational time required for matching a pair of templates. For the evaluation of the computational time, we perform our experiments on a machine with CPU Intel Core i7-6700HQ (2.60GHz) using MATLAB 2017b with Image Processing Toolbox, Windows 10. For our experiments, we adopt the following configurations: the maximum number of pixels for the translational shift in horizontal direction T_x is 25; the maximum number of pixels for the translational shift in vertical direction T_y is 18; the maximum degrees for the rotational shift R_d is 10. When evaluating the effect of k_p, we attempt to try all possible combination of rotations, results in $\leq 42{k_p}^2$ trail-and-error attempts. When evaluating the effect of k_s, we select the best k_p, i.e. 10, results in $\leq 200 + 21k_s$ trail-and-error attempts.

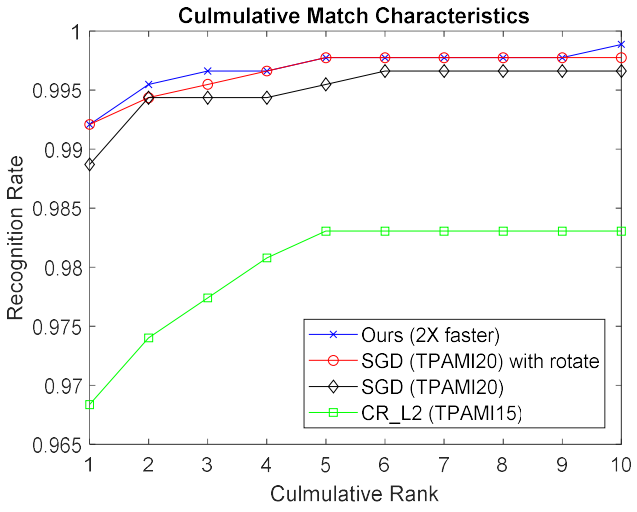


(a) (b) (c) (d)

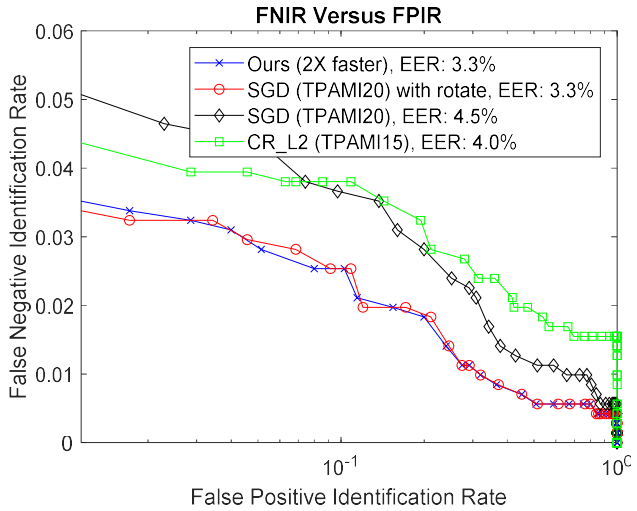
Fig. 7. Pairs of genuine 3D depth images from: (a)/(b) the HKPolyU Contact-free 3D/2D Hand Images Database; (c)/(d) the HKPolyU 3D Fingerprint Images Database.



(a)



(b)



(c)

Fig. 8. Comparative Experimental Results on the HKPolyU Contact-free 3D/2D Hand Images Database: (a) ROC; (b) CMC; (c) FNIR versus FPIR.

TABLE IV
COMPARATIVE EXPERIMENTAL RESULTS ON THE HKPOLYU CONTACT-FREE 3D/2D HAND IMAGES DATABASE

	SGD (*TPAMI20*) with rotate	Ours
EER (%)	0.88	0.82
Number of Attempts	1701	≤1115
Computational Time (s)	0.27	0.12

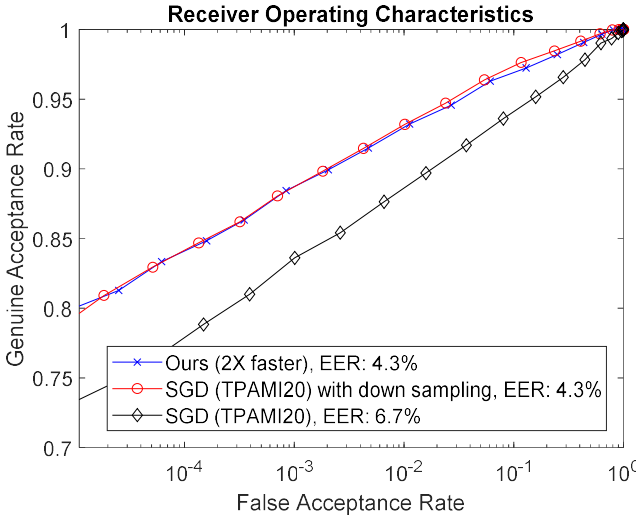
Table II and Table III shows the comparative experimental results. The columns with the final selected parameters (k_p: 10; k_s: 20) are in bold, i.e. 10 surface key points are selected for the 3D finger knuckle recognition experiments. These results can also be visualized in figure 5.

In the best of our knowledge, there are no further advancement on the conventional 'bit-shifting' approach for matching binary templates, which was widely employed in [1, 7-8, 13-18]. Therefore, we fairly compare our approach with the currently best performing state-of-the-art 3D finger knuckle recognition approach, *SGD* [13], with employing the conventional matching algorithm. It is worth noting that the *SGD* feature descriptor has been reported to be outperforming other two 3D hand biometric feature descriptors *Surface Code* [7] and *Binary Shape* [8]. Therefore, these two less competing methods are not selected as baselines for the systematic comparision in this paper. We present the experimental results using the new dataset using two-session images from 190 subject, while the finger knuckle images are segmented by Mask R-CNN method. k_p is set to be 10 and k_s is set to be 20 for obtaining the best possible performance.

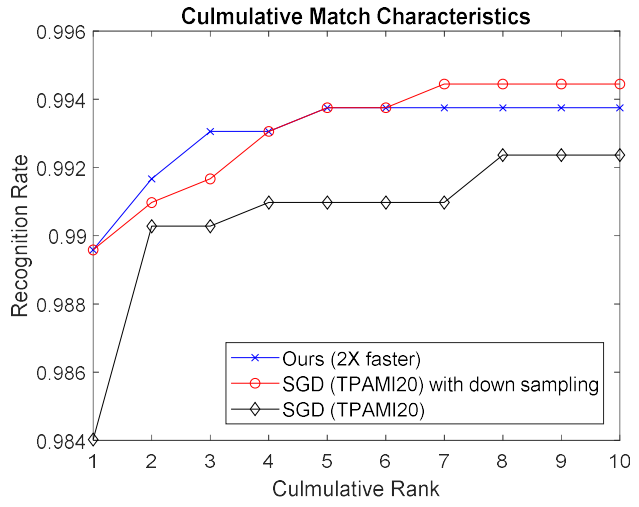
Figure 6 shows the comparative experimental results using ROC curves, CMC curves and FNIR versus FPIR curves. It can be observed that our improvement on the matching approach enables slight improvement on the recognition accuracy, which validates the theoretical arguments presented in Section I. More importantly, our method enables a huge improvement on the efficiency, with the number of trail-and-error attempts is ≤620 while the original method requires 39627 attempts, and the computational time is 0.12 seconds while the original method requires 2.85 seconds.

B. Evaluation with 3D Palmprint Images

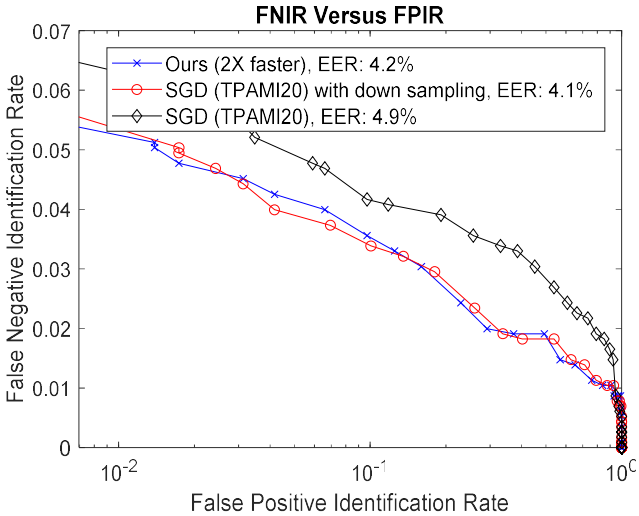
Although the focus of this paper is on 3D finger knuckle recognition, we also evaluate the performance of our proposed approach on another similar 3D biometrics, i.e. 3D palmprint, for demonstrating the potential of our approach. We employ the same 3D palmprint dataset as in [13], the HKPolyU Contact-free 3D/2D Hand Images Database Version 1.0 [7], containing two-sessions images from 177 subjects (each with five images per session), for the performance evaluation. Figure 7 (a)/(b) shows a pair of sample 3D depth images from the same subject. We also adopt the same evaluation protocol using the first session images for the training and the second session images for the testing, which generates 885 (177 × 5) genuine and 155760 (177 × 176 ×5) imposter matching scores. For the open-set identification experiments, 142 subjects (80%) are



(a)



(b)



(c)

Fig. 9. Comparative Experimental Results on the HKPolyU 3D Fingerprint Images Database: (a) ROC; (b) CMC; (c) FNIR versus FPIR.

TABLE V
COMPARATIVE EXPERIMENTAL RESULTS ON THE HKPOLYU 3D FINGERPRINT IMAGES DATABASE

	SGD (*TPAMI20*) with down sampling	Ours
EER (%)	4.25	4.25
Number of Attempts	2925	≤4096
Computational Time (s)	0.27	0.11

considered as enrolled users while the remaining 35 subjects (20%) are considered as unenrolled users.

For these 3D palmprint experiments, we compare our method with *SGD* [13], which is also reported to be outperforming another 3D hand biometric feature descriptors *Binary Shape* [8]. We further improve the reported experimental results in [13] by also considering the rotational shifting. For our approach and the *SGD* method, we adopt the following configurations: the maximum number of pixels for the translational shift in horizontal direction T_x is 4; the maximum number of pixels for the translational shift in vertical direction T_y is 4; the maximum degrees for the rotational shift R_d is 10. k_p is set to be 20 and k_s is set to be 15 for obtaining the best possible performance.

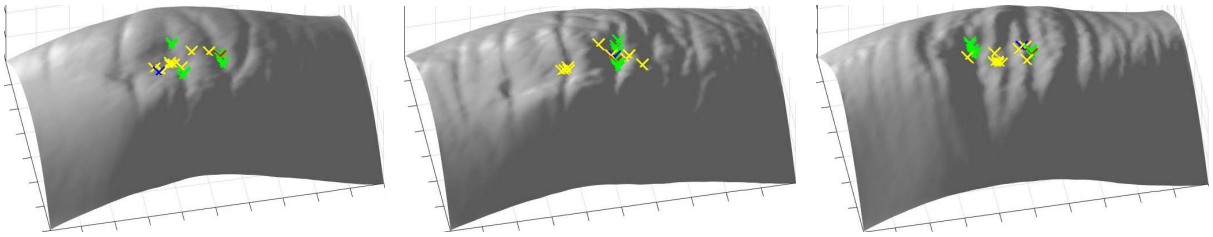
Besides, we also compare our approach with another state-of-the-art 3D palmprint recognition method, using collaborative representation based framework with L2-norm regularizations (*CR_L2*) [6], which has reported superior performance. In order to fairly compare with this method, we first investigate the variations between the reported database [27] and the selected database. Since both databases provide 3D depth images with the same resolution, it is reasonable to employ the same parameters provided along with *CR_L2* [6], which is already optimized for the reported database [27].

Figure 8 shows the comparative experimental results using ROC curves, CMC curves and FNIR versus FPIR curves. It can be observed that our improvement on the matching approach also enables slight improvement on the recognition accuracy for 3D palmprint recognition, reflected by the EER of ROC curve for verification experiments and CMC for close-set identification experiments, which again validates the theoretical arguments presented in Section III.

Furthermore, we compare the efficiency of our approach with the best performing approach, i.e. *SGD* with rotate. Table IV shows such comparative experimental results using EER from the ROC curves, the number of trial-and-error attempts and the average computational time required for matching a pair of templates. Our method outperforms the baseline method with the number of trail-and-error attempts is ≤ 1115 while the baseline method requires 1701 attempts, and the computational time is 0.12 seconds while the baseline method requires 0.27 seconds. Since the 3D palmprint images from this dataset do not contain large translational variations as in the 3D finger knuckle images, the effectiveness of our method over the baseline method is less significant.

C. *Evaluation with 3D Fingerprint Images*

Similar to Section IV.B, we also attempt to evaluate the performance of our proposed approach on another similar 3D biometric datasets, i.e. 3D fingerprint. We also employ the



(a) (b) (c)

Fig. 10. Sample Images with Degraded Accuracy for the Detection of Surface Key Points.

same 3D fingerprint dataset as in [13], the HKPolyU 3D Fingerprint Images Database [5], with one-session images from 240 subjects (each with six images), for the performance evaluation. Figure 7 (c)/(d) shows a pair of sample 3D depth images from the same subject. We adopt the same standard evaluation protocol for one-session datasets, which is an all-to-all protocol. This protocol matches all images with all other images, which generates 3600 ($240 \times C_2^6$) genuine and 1032480 ($C_2^{240} \times 6 \times 6$) imposter matching scores. For the open-set identification experiments, 192 subjects (80%) are considered as enrolled users while the remaining 48 subjects (20%) are considered as unenrolled users.

For these 3D fingerprint experiments, we compare our method with *SGD* [13], which is also reported to be outperforming another pixelwise 3D fingerprint feature descriptors *Finger Surface Code* [5]. We further improve the reported experimental results in [13] by also considering down sampling by half. The original configurations are as follows: the maximum number of pixels for the translational shift in horizontal direction T_x is 63; the maximum number of pixels for the translational shift in vertical direction T_y is 44. For the down sampled version, we adopt the following configurations: the maximum number of pixels for the translational shift in horizontal direction T_x is 32; the maximum number of pixels for the translational shift in vertical direction T_y is 22. k_p is set to be 64 for obtaining the best possible performance.

Figure 9 shows the comparative experimental results using ROC curves, CMC curves and FNIR versus FPIR curves. It can be observed that the down sampling is helpful for the recognition accuracy. Besides, it can be observed that our proposed approach produces comparable recognition accuracy for this 3D fingerprint recognition experiment.

Furthermore, we compare the efficiency of our approach with the best performing approach, i.e. *SGD* with down sampling. Table V shows such comparative experimental results using EER from the ROC curves, the number of trial-and-error attempts and the average computational time required for matching a pair of templates. It seems that a greater number of attempts are required for our method. In fact, the actual number of attempts is much less than 4096, and even much less than 2925 when comparing to the baseline condition with all possible trial-and-error attempts, because most translational shifting parameter sets are duplicated. Therefore, less computational time can be observed from the experimental results (0.11 versus 0.27 seconds). Since the surface gradient derivative features employed is most effective for 3D finger knuckle patterns, while minutiae features are expected to be the most effective for 3D fingerprint, it is not surprising that the effectiveness of our method over the state-of-the-art baseline is less significant.

V. Conclusions and Further Work

This paper introduces a new matching approach for improving 3D finger knuckle recognition. Experimental results presented in Section IV of this paper, on publicly available 3D hand biometric databases, indicate that the proposed approach can significantly improve the matching time, also the matching accuracy to a varying degree. This method is simple, theoretically justified, generalizable and can be easily integrated with other existing methods. Our approach firstly detects several key points using the reliable surface gradient derivative features extracted from 3D finger knuckle surfaces. Since those features have been shown to be discriminative for 3D finger knuckle recognition, it is judicious to employ those features for the detection of reliable key points. We further utilize the finger knuckle key points and develop a two-stage matching approach to match the templates accurately and efficiently. The effectiveness of our method is validated from our experimental results using three publicly available 3D biometric datasets including 3D finger knuckle pattern, 3D palmprint and 3D fingerprint. While comparing with the current state-of-the-art method, our approach outperforms the original methods by largely enhancing the computational efficiency and slightly improving the recognition accuracy. These diversified experimental results also suggest the generalizability of our proposed method. Our improvement on the matching problem can also be easily integrated with many existing methods using the trivial trial-and error matching approach such as many studies mentioned in this paper [1, 7-8, 13-18]. In the best of our knowledge, it is the first time to study on the improvement of matching binary templates by using key points, we employ the objective function using surface gradient derivative features. Those 3D images without sharp valley positions may result in less accurate detection of surface key points (shown in figure 10). Reference [35] introduces an interesting approach on finger knuckle recognition and incorporated a deep matching approach [36] which detect points correspondence between a pair of images. While the focus of this paper is to develop an interpretable, training free and more efficient approach, the development of more advanced methods for detecting more reliable key points, possibly with interpretable deep learning models, is also a part of further work in this area.

References

[1] L. Zhang, L. Zhang, D. Zhang and H. Zhu. Online Finger Knuckle-Print Verification for Personal Authentication. *Pattern Recognition*, vol. 43, no. 7, pp. 2560-2571, Jul. 2010.

[2] A. Nigam, K. Tiwari, and P. Gupta. Multiple texture information fusion for finger-knuckle-print authentication system. *Neurocomputing*, 188, pp.190-205, 2016.

[3] G. Jaswal, A. Kaul, and R. Nath. Knuckle Print Biometrics and Fusion Schemes – Overview, Challenges, and Solutions. *ACM Computing Surveys*, Vol. 49, No. 2, Article 34, Nov. 2016.

[4] J. Kim, K. Oh, B. S. Oh, Z. Lin and K. A. Toh. A Line Feature Extraction Method for Finger-Knuckle-Print Verification. *Cognitive Computation*, pp.1-21, 2018.

[5] A. Kumar and C. Kwong. Towards Contactless, Low-cost and Accurate 3D Fngerprint Identification. In *Computer Vision and Pattern Recognition (CVPR)*, 2013.

[6] L. Zhang, Y. Shen, H. Li, and J. Lu. 3D Palmprint Identification Using Block-Wise Features and Collaborative Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp. 1730-1736, Mar 2015.

[7] 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.

[8] 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. 633641, Mar 2016.

[9] P. Liu, Y. Wang, D. Huang, Z. Zhang, and L. Chen. Learning the Spherical Harmonic Features for 3-D Face Recognition. *IEEE Transactions on Image Processing*. 22(3), pp.914-925, 2013.

[10] S. Z. Gilani and A. Mian. Learning from Millions of 3D Scans for Large-Scale 3D Face Recognition. In *Computer Vision and Pattern Recognition (CVPR)*, 2018.

[11] 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.

[12] I. I. Ganapathi and S. Prakash. 3D Ear Recognition using Global and Local Features. *IET Biometrics*, 7(3), pp.232-241, 2018.

[13] 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.

[14] Q. Zheng, A. Kumar, and G. Pan. A 3D Feature Descriptor Recovered from a Single 2D Palmprint Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 6, pp. 1272-1279, Jun 2016.

[15] A. W. K. Kong and D. Zhang. Competitive Coding Scheme for Palmprint Verification. In *International Conference on Pattern Recognition (ICPR)*, 2004.

[16] W. Jia, D. S. Huang, and D. Zhang. Palmprint Verification Based on Robust Line Orientation Code. *Pattern Recognition*, vol. 41, no. 5, pp. 1504-1513, May 2008.

[17] J. Daugman. How Iris Recognition Works. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21-30, 2004.

[18] 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.

[19] 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.

[20] J. J. Koenderink and A. J. Vandoorn. Surface Shape and Curvature Scales. *Image and Vision Computing*, no. 8, pp. 557-564, Oct 1992.

[21] 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.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *Neural Information Processing Systems (NIPS)*, 2012.

[23] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-scale Image Recognition. In *International Conference on Learning Representations (ICLR)*, 2015.

[24] K. He, X. Zhang, S. Ren and J. Sun. Deep Residual Learning for Image Recognition. In *Computer Vision and Pattern Recognition (CVPR)*, 2016.

[25] K. He, G. Gkioxari, P. Dollár and R. Girshick. Mask RCNN. In *International Conference on Computer Vision (ICCV)*, 2017.

[26] R. Girshick. Fast R-CNN. In *International Conference on Computer Vision (ICCV)*, 2015.

[27] (Jun. 2014). [Online]. Available: PolyU 2D and 3D Palmprint Database, www.comp.polyu.edu.hk/~biometrics/.

[28] A. Jain, L. Hong, and R. Bolle. On-Line Fingerprint Verification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, April 1997.

[29] F. Chen, X. Huang, and J. Zhou. Hierarchical Minutiae Matching for Fingerprint and Palmprint Identification. *IEEE Transactions on Image Processing*, 22(12), pp.4964-4971, 2013.

[30] R. Cappelli, M. Ferrara, and D. Maio. A Fast and Accurate Palmprint Recognition System Based on Minutiae. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(3), pp.956-962, 2012.

[31] A. Kumar and B. Wang. Recovering and Matching Minutiae Patterns from Finger Knuckle Images. *Pattern Recognition Letters*, 68, pp.361-367, 2015.

[32] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C.L. Zitnick. Microsoft COCO: Common Objects in Context. In *European Conference on Computer Vision (ECCV)*, 2014.

[33] L. Fei, B. Zhang, Y. Xu, Z. Guo, J. Wen, and W. Jia. Learning Discriminant Direction Binary Palmprint Descriptor. *IEEE Transactions on Image Processing*, 28(8), pp.3808-3820, 2019.

[34] 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.

[35] G. Jaswal, A. Nigam, and R. Nath. DeepKnuckle: revealing the human identity. *Multimedia Tools and Applications*, 76(18), pp.18955-18984, 2017.

[36] J. Revaud, P. Weinzaepfel, Z. Harchaoui and C. Schmid. Deepmatching: Hierarchical deformable dense matching. *International Journal of Computer Vision*, 120(3), pp.300-323, 2016.

[37] X. Wang, L. Gao, J. Song and H. T. Shen. Beyond frame-level CNN: saliency-aware 3-D CNN with LSTM for video action recognition. *IEEE Signal Processing Letters*, 24(4), pp.510-514, 2016.

[38] L. Gao, X. Li, J. Song and H. T. Shen. Hierarchical LSTMs with adaptive attention for visual captioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[39] K. Sundararajan and D. L. Woodard. Deep Learning for Biometrics: A Survey. *ACM Computing Surveys (CSUR)*, 51(3), pp.1-34, 2018.

[40] Weblink for downloading the implementation codes. http://www4.comp.polyu.edu.hk/~csajaykr/knuckle_key.zip