

Face Recognition using 3D Directional Corner Points

Author

Yu, Xun, Gao, Yongsheng, Zhou, Jun

Published

2014

Conference Title

2014 22ND INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR)

DOI

[10.1109/ICPR.2014.483](https://doi.org/10.1109/ICPR.2014.483)

Rights statement

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Downloaded from

<http://hdl.handle.net/10072/66408>

Link to published version

<https://ieeexplore.ieee.org/document/6977196>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Face Recognition using 3D Directional Corner Points

Xun Yu, Yongsheng Gao

School of Engineering
Griffith University

Nathan, QLD, Australia

xun.yu@griffithuni.edu.au, yongsheng.gao@griffith.edu.au

Jun Zhou

School of Information and Communication Technology
Griffith University

Nathan, QLD, Australia

Jun.zhou@griffith.edu.au

Abstract— In this paper, we present a novel face recognition approach using 3D directional corner points (3D DCPs). Traditionally, points and meshes are applied to represent and match 3D shapes. Here we represent 3D surfaces by 3D DCPs derived from ridge and valley curves. Then we develop a 3D DCP matching method to compute the similarity of two different 3D surfaces. This representation, along with the similarity metric can effectively integrate structural and spatial information on 3D surfaces. The added information can provide more and better discriminative power for object recognition. It strengthens and improves the matching process of similar 3D objects such as faces. To evaluate the performance of our method for 3D face recognition, we have performed experiments on Face Recognition Grand Challenge v2.0 database (FRGC v2.0) and resulted in a rank-one recognition rate of 97.1%. This study demonstrates that 3D DCPs provides a new solution for 3D face recognition, which may also find its application in general 3D object representation and recognition.

Keywords—3D directional corner points; 3D directional corner point matching; 3D face recognition

I. INTRODUCTION

Human target recognition has been an active research area in the last decades, with one of the major topics on automatic face detection and matching for the purpose of verification and identification [1]. Significant achievements have been reached on two-dimensional (2D) face recognition, but there are still various problems in handling large amount of facial variances caused by changes in illumination, pose, expression and age. Because the human face is a three-dimensional (3D) object whose 2D projection is sensitive to above changes.

To overcome the inherent limitations associated with 2D face recognition technology, using 3D face information in face recognition has attracted increasing attention, with various techniques being presented in recent years [2]. 3D face recognition is expected to be less sensitive to illumination and pose variances because the 3D shape of a facial surface that is related to the internal face anatomy instead of external appearance and environment. In addition, the geometric information available in 3D data is thought to provide more discriminative features for face recognition.

However, 3D face recognition techniques also have their own drawbacks. The conventional methods [3] use holistic point clouds and meshes on 3D face are computationally expensive and in high storage demand. Therefore, it is crucial to find efficient and meaningful feature descriptors of 3D face

structure to perform recognition. Mahoor and Mohamed [4] encoded the range data of 3D face into a ridge image, which showed the locations of ridge lines around the important facial regions on the face (i.e., eyes, nose, and mouth). Then iterative closet points (ICP) matching method was utilized to match the 3D points lying on ridge image of a given probe to the created ridge images of the subjects in the gallery. In their work, only about $14\% \pm 2\%$ of the total number of points on the range data were used, but achieved 91.8% accuracy in experiments on FRGC v2.0 database. However, without considering the inherent local structural characteristics inside such images, this method uses only the spatial information of ridge image.

In this paper, a novel 3D face description and similarity measuring technique is proposed, which effectively harnesses structural and spatial information on a 3D surface, and reduces the storage requirement. Unlike methods using local operation of isolated points, the proposed approach employs 3D directional corner points (3D DCP) matching in which directional information showing connectivity to its neighbors is utilized in the point correspondence. Moreover, the storage load is further reduced by using salient corner points instead of all the points on those ridge and valley curves. The whole algorithm flowchart is illustrated in Fig. 1.

In the following, Section 2 presents the proposed 3D directional corner points, which incorporates structural information with spatial features. Section 3 presents our 3D directional corner points matching method. Encouraging experimental results on a public database are reported in Section 4. Finally, the paper concludes in Section 5.

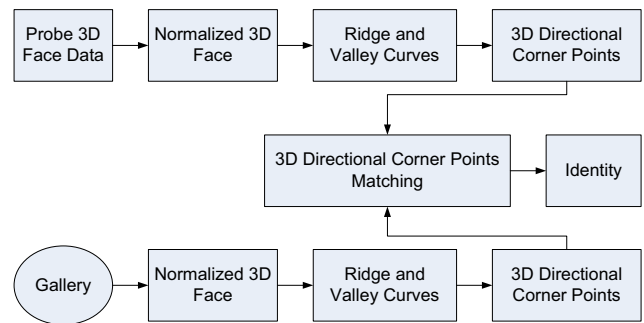


Fig. 1. Overview of the proposed method

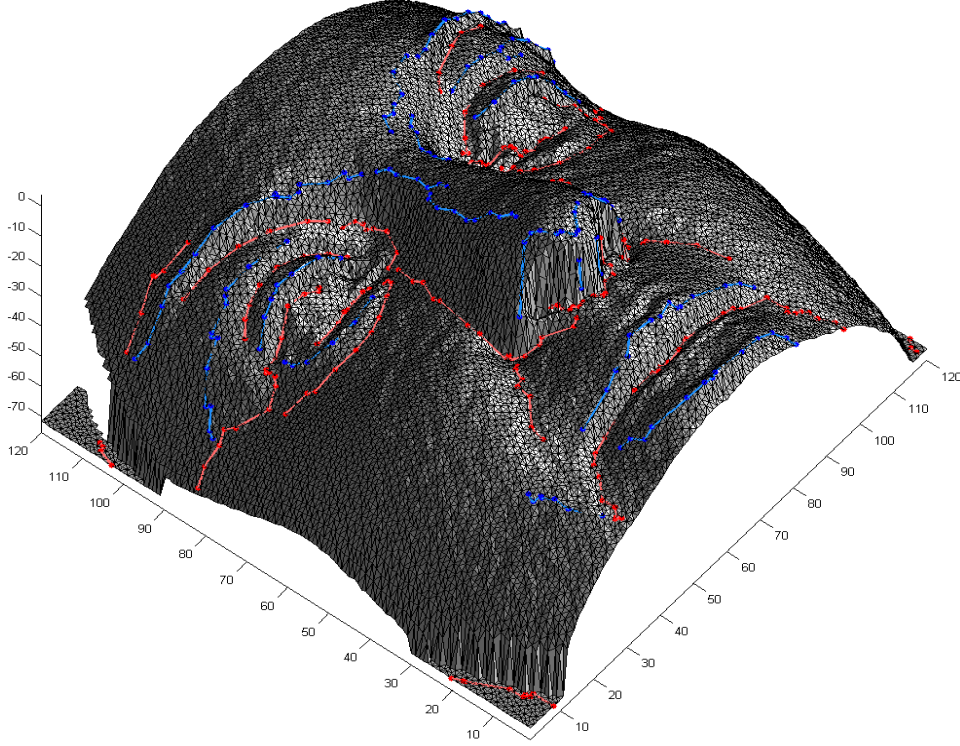


Fig. 2. Describing 3D shapes using 3D directional corner points. The red points are 3D DCPs on ridge curves and the blue points are 3D DCPs on valley curves.

II. 3D DIRECTIONAL CORNER POINTS

In this section, we first give a brief introduction on the concept of principle curvatures, based on which we then propose 3D directional corner points as a descriptor for 3D shapes.

For a given point on 3D face surface S , the maximal and minimal principal are k_{\max} and k_{\min} , and their corresponding principal directions are denoted as t_{\max} and t_{\min} . The definition of points on ridge or valley curves in differential geometry are characterized by

$$\begin{aligned}
 e_{\max} &= \partial k_{\max} / \partial t_{\max}, \quad e_{\min} = \partial k_{\min} / \partial t_{\min} \\
 e_{\max} &= 0, \quad \partial e_{\max} / \partial t_{\max} < 0, \quad k_{\max} > |k_{\min}| \quad (\text{ridges}) \\
 e_{\min} &= 0, \quad \partial e_{\min} / \partial t_{\min} < 0, \quad k_{\min} < -|k_{\max}| \quad (\text{valleys}) \quad (1)
 \end{aligned}$$

As a curvature-based feature descriptor, ridge and valley curves on a 3D face surface along the eyes, the nose, and the mouth where the surface bends sharply are geometrically and perceptually salient surface features, so they are expected to contain enough discriminative information for face recognition. In addition, coding 3D face into ridge and valley can decrease the storage demand and computational expense. However, ridge (or valley) image utilize spatial information of 3D face but lack structural representation.

In this study, we propose a 3D face feature descriptor, 3D directional corner points to solve this problem, which combines the structural connectivity information with spatial information of 3D faces. After detecting ridge and valley curves on 3D surface, a corner point detection process, which is based on Douglas-Peucker algorithm [5], is then applied to generate 3D DCPs of the face. A 3D DCP, represented as $P(x, y, z, \hat{n}_1, \hat{n}_2)$, consist of Cartesian coordinates (x, y, z) and two unit directional vectors \hat{n}_1 and \hat{n}_2 . \hat{n}_1 is the unit vector that points to its front neighboring corner point. Similarly, \hat{n}_2 is the unit vector that points to its rear neighboring corner point. If a 3D DCP is a start point of a curve, a null is assigned to \hat{n}_1 . If a 3D DCP is an end point of a curve, a null is assigned to \hat{n}_2 . A 3D DCP is a two-directional corner point with two unit directional vectors pointing to its two neighboring 3D DCPs or a one-directional corner point (start/end point) of the valley (or ridge) curves with a single unit directional vector pointing to its neighboring 3D DCP. These directional vectors provide isolated feature points with additional structural information about the connectivity to their neighbors, which can enhance the discriminative power of the descriptor. Moreover, the 3D DCP descriptor, using sparse points, further reduces the storage demand of a 3D shape representation. Fig. 2 shows the 3D DCPs of an example 3D face in FRGC v2.0 database.

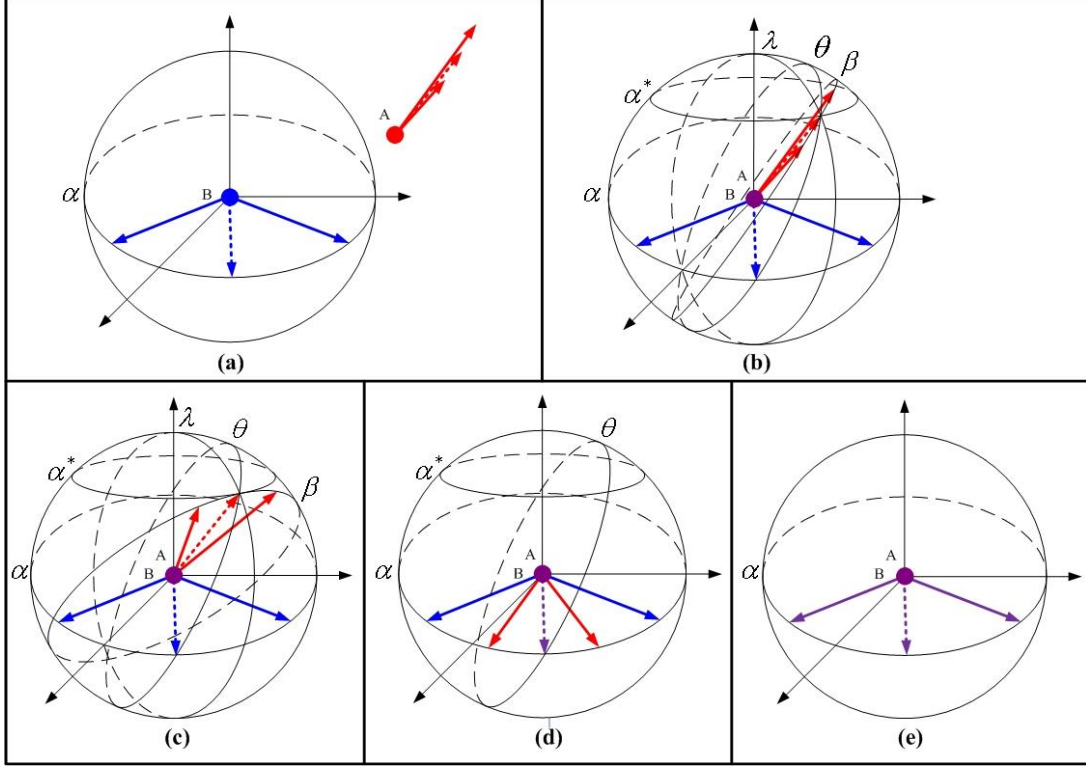


Fig. 3. An illustration of 3D DCP conversion: (a) two 3D DCPs before converting, (b) after translation operation, (c) after rotation 1 operation, (d) after rotation 2 operation, (e) after open/close operation

III. 3D DCP MATCHING

After utilizing the 3D DCP detecting process above, a 3D face is encoded into a set of 3D DCPs along ridge and valley curves, which contain both position and direction features. A point-to-point converting process is developed to calculate the difference between two 3D DCPs from a face in probe database and a face in gallery database. The dissimilarity between the two faces is then calculated through a global converting process between the two 3D DCP sets.

A. Point-to-point Correspondence

Let $A(x^A, y^A, z^A, \hat{n}_1^A, \hat{n}_2^A)$ and $B(x^B, y^B, z^B, \hat{n}_1^B, \hat{n}_2^B)$ be two 3D DCPs. The cost of converting A to B (vice verse) is calculated through a four-step process that consists of translation operation related to location feature, and rotation and open/close operations related to direction feature. Fig. 3 illustrates the whole process where unit vector $\hat{n}^A = (\hat{n}_1^A + \hat{n}_2^A) / |\hat{n}_1^A + \hat{n}_2^A|$ and $\hat{n}^B = (\hat{n}_1^B + \hat{n}_2^B) / |\hat{n}_1^B + \hat{n}_2^B|$ are considered as the principal directional vector of 3D DCPs A and B , which are plotted in red and blue dashed lines. To help demonstrate the whole converting process, there are several planes plotted in Fig. 2: $\alpha = \hat{n}_1^B \cap \hat{n}_2^B$, $\beta = \hat{n}_1^A \cap \hat{n}_2^A$, $\theta = \hat{n}^A \cap \hat{n}^B$, $\lambda = \hat{n}^A \cap (\hat{n}_1^B \times \hat{n}_2^B)$, $\alpha \parallel \alpha^*$.

1) Translation operation: A translation operation from A to B , denoted as $T(A \rightarrow B)$, moves A to the location of B , then $x^A = x^B$, $y^A = y^B$ and $z^A = z^B$ (Fig. 3(a) and (b)). The cost function for a translation operation from A to B is defined as

$$C[T(A \rightarrow B)] = \sqrt{(x^A - x^B)^2 + (y^A - y^B)^2 + (z^A - z^B)^2} \quad (2)$$

2) Rotation 1 operation: A rotation 1 operation from A to B , denoted as $R_1(A \rightarrow B)$, rotates w.r.t to \hat{n}^A till $\beta \perp \lambda$ (Fig. 3(b) and (c)). The cost function for a rotation 1 operation from A to B is defined as:

$$C[R_1(A \rightarrow B)] = 90 - \arccos(|\hat{n}^\beta \cdot \hat{n}^\lambda|) = 90 - \arccos(|(\hat{n}_1^A \times \hat{n}_2^A) \cdot (\hat{n}^A \times (\hat{n}_1^B \times \hat{n}_2^B))|) \quad (3)$$

where \hat{n}^β and \hat{n}^λ are the normal vectors of planes β and λ respectively.

3) Rotation 2 operation: A rotation 2 operation from A to B , denoted as $R_2(A \rightarrow B)$, rotates w.r.t to \hat{n}^A till $\beta \equiv \alpha$ (Fig. 3(c) and (d)). The cost function for a rotation 2 operation from A to B is defined as:

$$C[R_2(A \rightarrow B)] = \arccos(\hat{n}^A \cdot \hat{n}^B) \quad (4)$$

4) Open/Close operation: An open (or close) operation from A to B, denoted as $O/C(A \rightarrow B)$, opens (or closes) the two directional vectors of A until the two directional vectors coincide with the corresponding directional vectors of B (Fig. 3(d) and (e)). The cost function for an open/close operation from A to B is defined as:

$$C[O/C(A \rightarrow B)] = \left| \arccos(\hat{n}_1^B \cdot \hat{n}_2^B) - \arccos(\hat{n}_1^A \cdot \hat{n}_2^A) \right| \quad (5)$$

Let $A \rightarrow B$ denote a converting operation from A to B. Equation (6) defines the cost function for converting A to B as an integrated cost of above four operations.

$$C[A \rightarrow B] = \sqrt{C^2[T(A \rightarrow B)] + f^2 \left\{ \begin{array}{l} C[R_1(A \rightarrow B)] \\ + C[R_2(A \rightarrow B)] \\ + C[O/C(A \rightarrow B)] \end{array} \right\}} \quad (6)$$

where $f(x)$ is a non-linear function to penalize large angle deviation resulted from inter-class difference, but ignore small variation derived from segmentation error or intra-class difference. In this paper, a quadratic function

$$f(x) = \frac{x^2}{W} \quad (7)$$

is used, where W is the weight to be determined by a training process and the delimit of $f(x)$ is $[0, 450^\circ)$ as illustrated in Fig. 2.

For conversion between two one-directional 3D DCPs $A(x^A, y^A, z^A, \hat{n}_i^A)$ and $B(x^B, y^B, z^B, \hat{n}_j^B)$ ($i = 1$ or $2, j = 1$ or 2), the translation operation is the same as two-directional 3D DCPs, but the rotation and open/close operations are different. A rotation operation from A to B rotates \hat{n}_i^A to \hat{n}_j^B . The cost function for a rotation operation between two single-directional 3D DCP is defined as

$$C[R(A \rightarrow B)] = \arccos(\hat{n}_i^A \cdot \hat{n}_j^B) \quad (8)$$

To arrange the direction related operation cost in the same range $[0, 450^\circ)$ as that of two-direction 3D DCPs. The cost function between two one-directional 3D DCPs is defined as

$$C[A \rightarrow B] = \sqrt{C^2[T(A \rightarrow B)] + f^2 \{2.5 \times C[R(A \rightarrow B)]\}} \quad (9)$$

For converting between a one-directional 3D DCP and a two-directional 3D DCP, the translation operation is the same, but the maximum value 450° is set as the direction related operation cost. Because it is desirable to prohibit converting between two 3D DCPs of different types. The cost function between two one-directional 3D DCPs is defined as

$$C[A \rightarrow B] = \sqrt{C^2[T(A \rightarrow B)] + f^2(450)} \quad (10)$$

B. Set-to-set Correspondence

We start from the definition of two finite 3D DCP sets $G = G^r \cup G^v = G\{A_1^r, A_2^r, \dots, A_p^r, A_1^v, A_2^v, \dots, A_q^v\}$ and $P = P^r \cup P^v = P\{B_1^r, B_2^r, \dots, B_s^r, B_1^v, B_2^v, \dots, B_t^v\}$ that represent a gallery and a probe in the 3D face database respectively, where superscripts r and v stand for ridge and valley. Therefore, G consists of two subsets G^r and G^v that correspond to 3D DCPs along ridge and valley curves, and similar setting applies to P.

A 3D DCP set to set converting process is proposed to establish every 3D DCP correspondence between the two 3D DCP sets by minimizing the global conversion cost. For each 3D DCP A_i^k in G^k , its corresponding 3D DCP B_j^k in P^k is identified as the one with minimum converting cost from A_i^k to B_j^k among all $B_j^k \in P^k$ (k represents r or v). The cost for establishing the correspondence for A_i^k can be calculated by

$$C(A_i^k) = \min_{B_j^k \in P^k} C(A_i^k \rightarrow B_j^k) \quad (11)$$

Similar to least trimmed square - HD (LTS - HD) [6] to handle the outlier noise derived from the 3D scanning and corner detecting process, the cost for converting the whole set G^k and P^k , denoted as $G^k \Rightarrow P^k$, is defined as

$$C(G^k \Rightarrow P^k) = \frac{1}{H} \sum_{i=1}^H C(A_i^k) = \frac{1}{H} \sum_{i=1}^H \min_{B_j^k \in P^k} C(A_i^k \rightarrow B_j^k) \quad (12)$$

where $H = h \times N_{G^k}$ ($0 \leq h \leq 1$), N_{G^k} is the number of 3D DCP A_i^k in G^k , and $C(A_i^k)$ are sorted in sequence $C(A_1^k) \leq C(A_2^k) \leq \dots \leq C(A_{N_{G^k}}^k)$. The measure is calculated by eliminating the large conversion cost values and only keeping the h fraction of the smallest converting cost. In this study, the value of h is set as 0.8, which results the best recognition performance.

Finally, the dissimilarity between G^k and P^k is defined as the maximum of the two minimum costs to establish correspondences from G^k to P^k and vice versa. In (13), the dissimilarity costs between ridge and valley subsets are calculated first, then the final result is merged from the two subset dissimilarities. This fusion process can further improve the discriminative power.

$$\begin{aligned} D(G, P) &= D(G^r, P^r) + D(G^v, P^v) \\ &= \max[C(G^r \Rightarrow P^r), C(P^r \Rightarrow G^r)] \\ &\quad + \max[C(G^v \Rightarrow P^v), C(P^v \Rightarrow G^v)] \end{aligned} \quad (13)$$

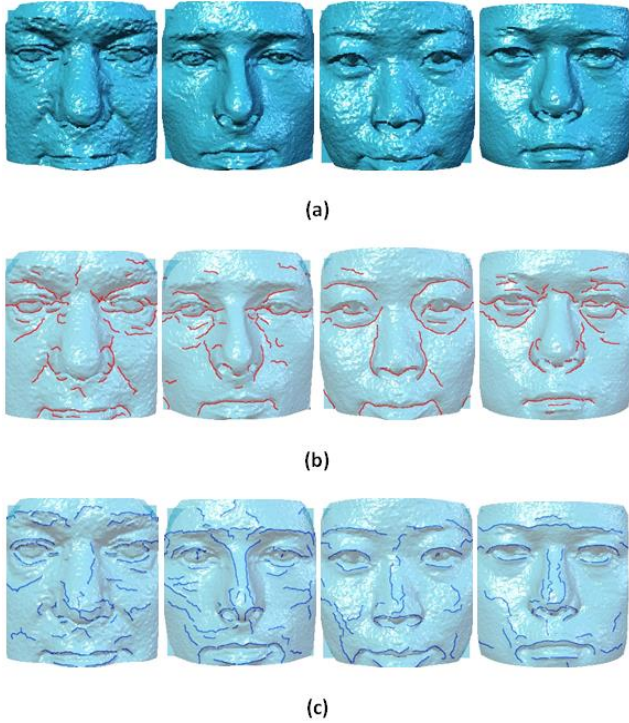


Fig. 4. Sample 3D faces used in our experiments. (a) Normalized faces, (b) Ridge data, (c) Valley data.

IV. EXPERIMENTS AND RESULTS

The FRGC v2.0 [7] 3D face database was used in our experiments to evaluate the feasibility and effectiveness of the proposed approach. The FRGC v2.0 database contains 4950 face texture and range images, divided into three sets, namely Spring2003, Fall2004 and Spring2004. In line with FRGC v2.0 protocol, in our experiments, the training set was generated from Spring2003 and the test images were generated from the other two sets. The face images are extracted, normalized and cropped in the same manner as in [8]. Spikes in the range maps are removed and holes are filled. Fig. 4 illustrates several samples of normalized and cropped face images in the FRGC v2.0 dataset.

A. Determination of Parameter W

In this section, we investigate the effect of parameter W in (7) on the recognition accuracy. In this experiment, 100 people with 2 neutral 3D faces per person from the FRGC v2.0 Spring2003 were used to create a training dataset. The neutral 3D face in session one were used to construct the gallery database, and the neutral 3D face in session two were used as probe images. The recognition rate is plotted against the values of W in Fig. 5. It is observed that the algorithm performed badly with a low value of W and only achieved 7.2% when $W = 10$. The performance increased quickly and reached the optimal value of 98% when W was 600 and remained stable till 1200. In the rest of experiments in this study, W was set as 1000.

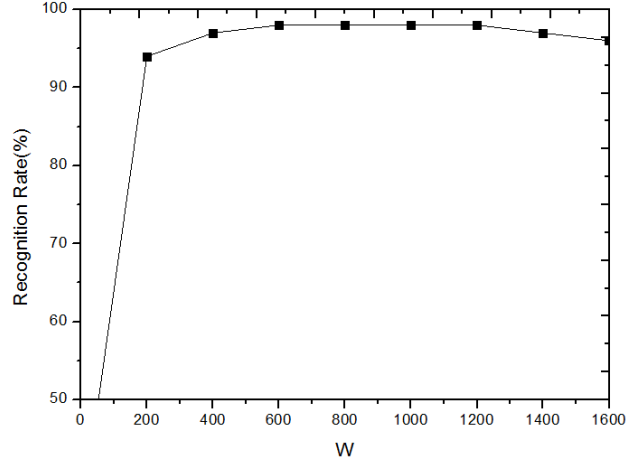


Fig. 5. The effect of W on recognition rate.

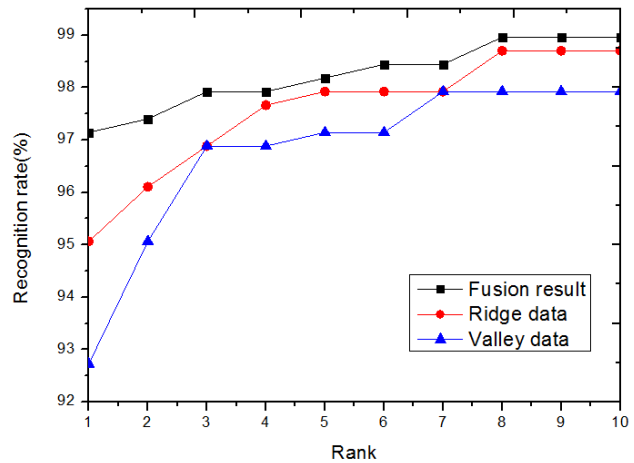


Fig. 6. CMC curves of the proposed approach

B. Face Recognition

In this experiment, we investigated the effectiveness of the proposed approach on the neutral 3D face images of the FRGC v2.0 database. There are 388 subjects which have at least two neutral images captured in different sessions. However, some images were found lost or corrupted after downloading through Internet. 385 subjects are complete and can be used. For each subject, one neutral 3D face was used as a gallery while the other was used as a probe.

The performance is measured in terms of the Cumulative Match Characteristics (CMC) [9] and the rank-1 recognition rate. In order to demonstrate the performance improvement resulted from the fusion process, we also presented the final fusion result, as well as the result based on ridge (or valley) data only, as illustrated in Fig. 6. It is encouraging to find that the fusion process improve the recognition performance greatly. The rank-1 recognition rate of the methods given in Fig. 6 above are tabled in Table 1 together with the reported results of Mahoor and Mohamed [4]. Note our methods perform consistently superior to the benchmark method based

TABLE 1. RECOGNITION ACCURACIES ON 3D FACE SCANS

Method	Recognition Result	Point Number*
3D DCP (fusion result)	97.1%	5.9%
3D DCP (ridge data)	95.1%	3.7%
3D DCP (valley data)	92.7%	2.2%
Mahoor and Mohamed [4] (ridge image)	91.8%	14%
Mahoor and Mohamed [4] (entire surface)	93.7%	100%

* Point number is calculated through the percentage of the entire 3D face scans.

on ridge image, which demonstrates that the added structural information makes 3D DCP a more discriminative measure than traditional feature point based method. The fusion result of our method is also higher than Mahoor and Mohamed [4] on the entire surface by 3.4%. In addition, compared with the benchmark that using entire surface, our 3D DCP methods based on ridge data, valley data and fusion process, require only about 3.7% , 2.2% and 5.9 % storage space, which further decreases the storage demand.

V. CONCLUSIONS

This paper presents a new 3D face recognition method using 3D directional corner points (3D DCPs), which employs both spatial and structural information of ridge and valley curves on 3D surface. In order to represent the 3D shape efficiently, we extract 3D directional corner points from ridge and valley curves. Such sparse point representation can further reduce the storage demand and the directional attributes can effectively enhance the discriminative power.

The proposed method has been evaluated on the FRGV v2 dataset and been compared with a benchmark approach based on ridge image. It is very encouraging to find that the 3D DCP method performed superior to the benchmark approach in terms of higher recognition accuracy and less storage space demand. This study reveals that 3D DCPs provides a new solution for 3D face recognition, which may also find its application in general 3D object representation and recognition.

REFERENCES

- [1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, pp. 399-459, 2003.
- [2] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Computer Vision and Image Understanding*, vol. 101, pp. 1-15, 2006.
- [3] G. Medioni and R. Waupotitsch, "Face modeling and recognition in 3-D," in *Proceedings of IEEE International Workshop on Analysis and Modeling of Face and Gestures*, pp. 232-233, 2003.
- [4] M. H. Mahoor and M. Abdel-Mottaleb, "Face recognition based on 3D ridge images obtained from range data," *Pattern Recognition*, vol. 42, pp. 445-451, 2009.

- [5] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, pp. 112-122, 1973.
- [6] D. G. Sim, O. K. Kwon, and R. H. Park, "Object matching algorithms using robust Hausdorff distance measures," *Ieee Transactions on Image Processing*, vol. 8, pp. 425-429, 1999.
- [7] P. J. Phillips, *et al.*, "Overview of the face recognition grand challenge," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 947-954, 2005.
- [8] 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, pp. 1927-1943, 2007.
- [9] S. A. Rizvi, P. J. Phillips, and H. Moon, "The FERET verification testing protocol for face recognition algorithms," in *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, pp. 48-53, 1998.