Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Weightiness image Partition in 3D Face Recognition

Guanghui He,Yuanyan Tang,Bin Fang, and Taiping Zhang
Department of Computer Science,Chongqing University, Chongqing 400030, P.R. China
Email: ghhe@cqu.edu.cn

Abstract—**In this paper we present a novel algorithm suitable to improve the accuracy of 3D face recognition. In the proposed algorithm, we represent the 3D points by Point signatures and partition the facial data into fifteen regions according to "three courtyards and five eyes" theory in pencil sketch on facial image in Chinese traditional art. Then in each partition we use ICA getting eigenvalues of feature and structure character and depth information to represent the 3D facial data. We assign different weightiness to each sub-image according to the result of sub-image variety. In order to match incomplete data under structural constraints, we proposed a reformative robust Structural Hausdorff Distance to handle these possible cases. Experiments on FRGC v2.0 data set show that the proposed algorithm is robust and effective to 3D face with expression, lighting and expression variance.**

Index Terms —**3D face recognition, image partition, Structural Hausdorff Distance**

I. Introduction

Face recognition is one of the most important applications of computer vision, particularly with the increasing security concerns in the last decade. In the past years, major advances are focused on 2D intensity images[1][2][3][4][5][6][7],but there are still considerable challenges in uncontrolled environments due to the pose, illumination and expression variations. With the recent availability of accurate 3D sensors , which are capable of sensing both 3D face shape and texture, it is widely considered that some of the inherent problems of 2D intensity-based recognizers can be overcome, because that 3D images can compensate for the depth information which is absent in a 2D image. 3D Facial image includes much information, such as eyes, nose, mouth, depth, the distance between two eyes, the distance between eyes and mouth and their relative positions. We can classify these information into two categories, feature information(FI) including such as eyes, nose, mouth, depth, et al. and structure information(SI) including the distance between two eyes, the distance between eyes and mouth and their relative positions and so on.

Bowyer et al. proposed a thorough survey of 3D face recognition algorithms in[8]. Pan[9] design a pose-invariant recognition system by projecting the pre-registered 3-D point-cloud data to a plane parallel to the face plane. Russ et al.[10]. brought forward an approach for matching 2D depth image(range images), using the original measured data and not their subspace projection. They apply the partial-shape Hausdorff distance metric to range images. The motivation behind using the Hausdorff distance is its partial invariance to inconsistencies such as noise, holes, and occlusions in the 3-D facial data.

Koudelka[11] et al. locate automatically several facial landmarks such as nose tip, sellion, inner eye corners, and mouth center and then sample 150 random points in their neighborhood based on the iterative closest point(ICP) algorithm. The matching of two facial surfaces is then accomplished via a mixture of ICP and Hausdorff algorithms. Irfanoglu et al.[12]use the thin-plate-spline (TPS)-warping algorithm. First, several facial landmarks are located automatically, and a given face image is then warped to an average face model (AFM) using TPS. A similar method is also proposed in[13], where a generic face model is fitted to a given face, and the related displacement information forms a separate deformation image. Finally, the biometric signature is obtained from the wavelet analysis of this deformation image.

Although warping-based registration algorithms may have a potential of establishing better correspondences around the dynamic facial regions, they may have the side effect of suppressing characteristic differences between faces[14]. In order to avoid inefficient effect, Lu and Jain[15] suggested using person-specific deformable models. Approaches on 2D or 3D face recognition may be classified into two main categories. The first relies on the extraction and matching of salient surface features usually based on curvature information[16][17][18], this approach makes facial recognition by FI. In the second approach, the face surface is represented and image classification techniques are applied, the second one [19][20] is absorbed in find the general information mainly including SI to recognize persons. Experiments show that recognition accuracy leads to degraded if only using FI or SI in facial recognition.

In this paper, we propose a novel method to recognition 3D facial images using both FI and SI to improve the recognition accuracy. In the proposed method, we partition the facial images into fifteen subimages based on point signature[21], which are with the same size, then in each subimage we get the all eigenvalues found by LDA to represent the subimage. And we use Structural Hausdorff Distance (SHD) to distinguish each vector representing the facial image. Experimental results show that the proposed method achieves higher facial images recognition accuracy.

This paper is organized as follows. SectionII introduces the partition model of facial image. SectionIII gives facial image similarity with weighted index and we introduce facial feature extraction and classification. SectionIV discuses our experimental results and compares them to other published results on FRGC v2.0 database. Finally, SectionV provides the conclusion.

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

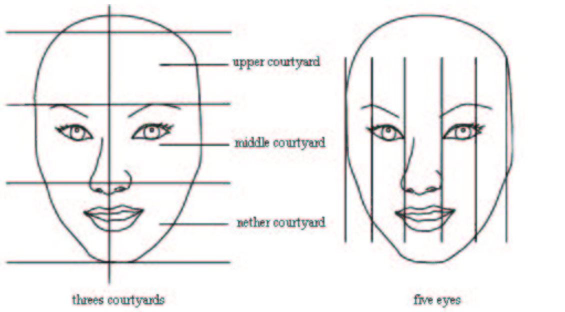

Fig. 1. The geometry description of " three courtyards and five eyes" theory

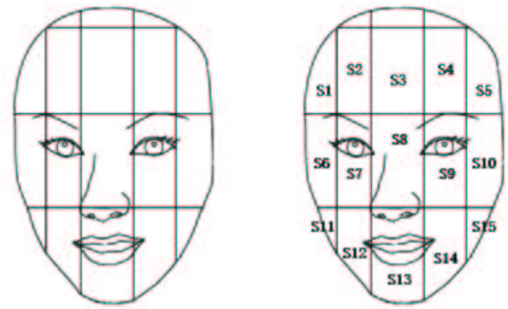

Fig. 2. Partition the facial image into fifteen regions and the corresponding encodings

II. Partition for Facial Image

There is a "three courtyards and five eyes" theory in pencil sketch on facial image in Chinese traditional art[22]. We draw a vertical line from forehead to chin and pass through nasoscope and philtrum firstly, and then draw a horizontal line pass through glabella, draw another horizontal line pass through the bottom of nosewing. Thus the two horizontal lines will divide the face image into three parts, which are called "three courtyards" and the height of each courtyards are uniform. The "five eyes" are that the distance from the outboard of the canthus to the verge of the same side bob is a "eye", and the distance of the two eyes is a "eye". Because of symmetry of face image, we will get "five eyes". This theory defines that the facial image of normal person satisfy " three courtyards and five eyes", where "courtyard" means a region of face and "eye" means a length of one eye. So the "three courtyards and five eyes" shows that the person facial image can be partition into three courtyards with same height in vertical direction, and partition five eyes-length in horizontal direction, as shown in fig.1. From Fig.1. we can draw a conclusion that " three courtyards and five eyes" theory partition the facial image into three and five same distance regions in vertical and horizontal respectively. These partitions include the FI and SI, in others words, these partitions include the objects information such as eyes, mouth, nose, eyebrow etc, and the information of object relative position namely topological information. So we partition the facial image into fifteen regions, as shown in Fig.2. Let S be facial image, and the sub-images are encoded as S^i from left to right and upper to down,$i = 1, 2, \cdots, 15$, and $Sim(S_i^k, S_j^k)$denote the similar of two sub-images S_i^k and S_j^k of two different global facial images S_i and S_j.

Lemma 1 Let $S_i^k \in S_i$, and $S_j^k \in S_j$ $(i \neq j)$, then $m \leq Sim(S_i, S_j) \leq M$, where $m = \ min\left\{Sim(S_i^k, S_j^k)\right\}$, $M = \max\left\{Sim(S_i^k, S_j^k)\right\}$,$k = 1, 2, \cdots, 15$.

In practical, we usually confirm two facial images as the same person's or not with some distinct features such as eyes, nose, although some others features are not so similar. That is to say that the corresponding features on the two facial images have not the same similar degree. The similar degree of two facial image $S(S_i, S_j)$ is less than the similar degree of distinct features and larger than the other features.

Lemma 2 Let $S_i^k \in S_i$, and $S_j^k, S_j^l \in S_j$,$(k \neq l, i \neq j)$, then $Sim(S_i^k, S_j^k) \geq Sim(S_i^k, S_j^l)$. S_i^k and S_j^l $(k \neq l)$ are the different partition of two facial images.

In our partition, each partition usually has different feature of face. So $Sim(S_i^k, S_j^k) \geq Sim(S_i^k, S_j^l)$. From the above lemmas we can deduce that $Sim(S_i^k, S_j^k)$ of the two sub-images S_i^k and S_j^k in the same position of the arbitrary two facial images should be larger than $Sim(S_i^k, S_j^l)$ $(k \neq l)$. Because the corresponding position sub-images include the same objects in two facial images, which reflects the topological information namely structure information(SI). Now, we have partition a facial image into fifteen sub-images in sequence,and we can draw a conclusion that our partition method preserves the SI comparing to other partition methods.

III. Facial Image Similarity computing

From the Lemma 1, it is obvious that the facial image recognition rely on the similarity of sub-images. That is to say, if each similarity of the two facial sub-images is small, then the similarity of the two global facial images will be small. Considering that the facial images we get perhaps imply pose invariance, lighting condition and expression, so some similarities of the sub-image pairs are more than the global similarity and the others are less than the global similarity. But is the sub-image has the same contribution to the global similarity?

There are fifteen sub-images in Fig.2, but each sub-image contains different objects, such as S_7 and S_9 containing left eye and eyebrow, right eye and eyebrow respectively, S_8 containing mainly part of the nose, and S_{12}, S_{13}, S_{14} containing part of the mouth. And the other contain nearly no features. As we known that the eye, eye brow, nose, mouth are most important objects to facial recognition. So a novel idea comes into being, that is the sub-images have different contribution to the similarity of facial images. In order to enhance the accuracy, We should evaluate different weightiness to similarity of sub-images. On the other hand, as we know, the same person's face images in different expressions are widely different. So when we assign weightiness to each sub-image, we should take expression variety into account.

It is obvious that the pixel variabilities of the same face in different expression are distinct. In order to denote the difference, we computer the variance V_k of each sub-image S_k and assign weightiness W_k to each sub-image S_k according to its variance W_k In this way, we can roughly distinguish the face expression for our experiment. Thus we can present the face image with weightiness W_k the sub-matrix S_k and we will have the following computing formula:

$$Sim(S_i, S_j) = \frac{\sum_{k=1}^{15} W_k \times Sim(S_i^k, S_j^k)}{\sum_{k=1}^{15} W_k} \quad (1)$$

Based on the 2D depth image(range images), 3D features for a test face are to be extracted. Let $\alpha = [\lambda^1, \cdots, \lambda^{15}, d^1, \cdots, d^{15}, \mu^1, \cdots, \mu^{15}]$ be the features in the depth image, where $\lambda^i (1 \leq i \leq 15)$ is the max eigenvalues of sub-image S^i matrix according to PCA algorithm ,d^i and μ^i are the average depth and the depth variance of nodes in sub-image S^i respectively. With α, a robust Structural Hausdorff Distance (SHD) measure is proposed to compare a test face with all models in 3D domains. The aim of using the SHD measure is to constrain the Hausdorff Distance matching with the structure of the image. In this paper, the structure information is encoded in a confidence matrix, which is computed based on the intra-set 3D distances using the Eigenvector approach [23] . The Eigenvector approach based on the intra-set distances was suggested by Shapiro and Brady. Its main idea is discussed as follows. Denote two vectors as $p_1 = [p_{11}, \cdots, p_{1m}]$ and $p_2 = [p_{21}, \cdots, p_{2n}]$. The traditional Eigenvector method is briefly introduced as follows. For the set p_1, create the Gaussian proximity matrix $M_1 = \{M_{1ij}\}$, where

$$M_{1ij} = \exp(-r_{ij}^2 / 2\sigma^2) \quad (2)$$

and $r_{ij}^2 = \|p_{i1} - p_{j1}\|^2$. The parameter σ controls the intersection between features in the set p_1. Next step, we compute the eigenvalues λ_i and eigenvectors E_i(termed as modes here) of M_1, then

$$M_1 = V_1 D_1 V_1^T \quad (3)$$

The diagonal matrix D_1, contains the eigenvalues along its diagonal in descending order. The modal matrix $V_1 = [E_1, \cdots, E_m]$is orthogonal. Each row of V_1 can then be thought of as a feature vector Y_{1i}, i.e. $V_1 = [Y_{11}, \cdots, Y_{1m}]^T$. Similarly,this computation is applied to the data set p_2(n features), and we can obtain $M_2 = V_2 D_2 V_2^T$, The associated feature vectors are written as $Y_{1i}(1 \leq i \leq m)$ and $Y_{2j}(1 \leq j \leq m)$. The final stage is to correlate these two sets of features, yielding the confidence matrix Z. The confidence matrix,$Z = \{Z_{i,j}\}(1 \leq i, j \leq k)$ is derived with

$$Z_{ij} = \|Y_{1i} - Y_{2j}\|^2 \quad (4)$$

It is solvable to scale the eigenvectors by their corresponding eigenvalues. This effectively give more weights to the dominant modes. Carcassoni and Hancock [24] investigated the performances when defining different proximity matrices. In this paper, we adopt the suggested proximity matrix which exhibit good performances. Let r_{ij} be defined as the 3D distance between feature point i and feature point j. Specifically, the proximity matrix is described as follows:

$$M_{ij} = [1 + r_{ij}/\sigma]^{-1}, \sigma = \frac{1}{m(m-1)} \sum_{i,j,i \neq j}^{m} r_{ij} \quad (5)$$

After obtaining the proximity matrices for the model face and test face, the confidence matrix Z can be derived based on the above introduced EA approach. In this paper, Z_{ij} is proposed to be integrated in the Hausdorff Distance measure as a weight for the distance between the features extracted from point p_{i1} and point p_{j2}. This is the essence of the Structural Hausdorff Distance measure presented in this paper. The Hausdorff Distance was originally proposed by Huttenlocher et al.[25] for binary image comparison and computer vision. Unlike most shape comparison methods, the Hausdorff Distance can be calculated without the explicit pairing of points in their respective data sets, and can be extended to allow partial matching. It is because of this desirable property that we use the HD measure in our method.

Conventionally, given two finite points sets $A = \{a_1, \cdots, a_p\}$ and $B = \{b_1, \cdots, b_q\}$, the Hausdorff Distance is defined as

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (6)$$

where

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (7)$$

is called the directed Hausdorff Distance from A to B. Since the Hausdorff Distance measure is very sensitive to degradations caused by noise and occlusions, improved methods such as partial Hausdorff Distance[25], M2HD [26] and spatially weighted Hausdorff Distance[25] have been proposed. Here, we use a directed Structural Hausdorff Distance (SHD) function by considering the correspondence probability when comparing two features

$$h(A, B) = \frac{1}{p} \sum_{i=1}^{p} \rho_i \min_{b_j \in N_B^{a_i}} (Z_{i,j} \|a_i - b_j\|) \quad (8)$$

where Z_{ij} is the element of the correspondence matrix computed based on Eq. (4), $\rho_i = 1$ iff $\min_{b \in N_B^{a_i}} Z_{ij} \|a_i - b\| < \tau$, otherwise, $\rho_i = 0$, and τ is a threshold to eliminate outliers and is carefully chosen through experiments. In order to avoid the full search in the conventional HD method, $N_B^{a_i}$ is defined as the neighborhood of point $a_i \in A$ in set B as in[26], and is constrained by the pre-localized fiducially points. Since the structural information is encoded in the correspondence matrix, compared with the general HD-based approach, the matching results of the SHD based approach are expected to be more robust. This will be verified in the next section.

TABLE I
PERFORMANCE COMPARISON ON THE FRGC V2.0 DATABASE

Method	Recognition Rate	Variation
HD	92.7667	2.9122
M2HD	92.8103	3.1498
SHD	93.0452	4.5187
WPHD	95.1380	2.5071

IV. Face Recognition

The proposed method is tested on neutral expression faces of the FRGC v2.0 data set. The FRGC v2.0 data set has 400 gallery faces and about 1500 neutral probe faces under variations in pose, varying ages and lighting. In this section, experiments have been conducted to show the effectiveness of our proposed method for face recognition on FRGC v2.0 databases. For a face to be tested, after determining its view label and localizing the 2D partition in Section 3, its represent vectors, α is determined. Employing the robust SHD measure, the best matched model face is identified based on the linear integration of the normalized comparison results in 2D depth image.

The experiments on 2D depth image (range image) have been systematically performed. A number of conclusions and interesting points were revealed. The derivation of our method, theoretical analysis and face representation (Fig. 2) in SectionIII show our method is able to use both the FI and SI, especially depth information. The method is more effective to describe 2D depth face image representation by preserving local topology structure. The method consistently performs better than HD, M2HD and SHD, as show in TableI.These experiments show that our proposed method is robust to changes in lighting, facial expression. The present algorithm is not able to resolve face recognition with lange-angle pose. Effective algorithm design for face recognition with lange-angle pose would be a topic of our future work.

V. Conclusion

In this paper, we propose a novel method to recognition 2D depth facial images using both FI and SI including the depth information to improve the recognition accuracy. Our method partitions a facial image into fifteen sub-images and presents each sub-image with a forty-five dimension vector. Our experiments show that the proposed method is robust to expression variance and lighting.

VI. Acknowledgment

This work is supported by the Program for New Century Excellent Talents of Educational Ministry of China (NCET-06-0762), Natural Science Foundation of China (60873092) and Natural Science Foundation of Chongqing (CSTC2007BA2003).

References

[1] W. Y. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face recognition a literature survey," ACM Comput. Surv., vol. 35, no. 4, pp. 399-458, 2003.

[2] M. Turk and A. P. Pentland, "Eigenfaces for recognition," J. Cogn.Neurosci., vol. 3, no. 1, pp. 71-86, 1991.

[3] S. Z. Li, X. W. Hou, H. J. Zhang, and Q. Cheng, "Learning spatially localized, parts-based representation," in Proc. IEEE Comput. Vis. Pattern Recognit., 2001, pp. 207-212.

[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces : Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711-720, Jul.1997.

[5] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Generative models for recognition under variable pose and illumination," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 6,pp. 643-660, Jun. 2001.

[6] I. Dagher and R. Nachar, "Face recognition using IPCA-ICA algorithm," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 6, pp.996-1000, Jun. 2006.

[7] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces" IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 3, pp. 328-340, Mar. 2005.

[8] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition,"Comput. Vis. Image Underst., vol. 101, no.1,pp.1-15, Jan. 2006.

[9] G. Pan, S. Han, Z. Wu, and Y. Wang, "3D face recognition using mapped depth images," in Proc. IEEE Workshop Face Recog. Grand Challenge Experiments, 2005, p. 175.

[10] Russ, T.D.; Koch, M.W.; Little, C.Q.A 2D Range Hausdorff Approach for 3D Face Recognition,Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on 25-25 June 2005 Page(s):169 - 169

[11] Koudelka, M.L.; Koch, M.W.; Russ, T.D. A Prescreener for 3D Face Recognition Using Radial Symmetry and the Hausdorff Fraction, Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on 25-25 June 2005 Page(s):168 - 168

[12] M. O. BIrfanoglu, B. Gokberk, and L. Akarun, "3D shape-based face recognition using automatically registered facial surfaces," in Proc. Int. Conf. Pattern Recog., 2004, pp. 183-186.

[13] G. Passalis, I. Kakadiaris, T. Theoharis, G. Toderici, and N. Murtuza, "Evaluation of 3D face recognition in the presence of facial expressions: An annotated deformable model approach," in Proc. IEEEWorkshop Face Recog. Grand Challenge Experiments, 2005, p. 171.

[14] B. Gokberk, M. O. Irfanoglu, and L. Akarun, "3D shape-based face representation and facial feature extraction for face recognition," Image Vis. Comput., vol. 24, no. 8, pp. 857-869, 2006.

[15] X. Lu and A. K. Jain, "Deformation modeling for robust 3D face matching," in Proc. IEEE Comput. Soc. Conf. CVPR, 2006, pp. 1377-1383.

[16] S. Zafeiriou, A. Tefas, I. Buciu, and I. Pitas, "Exploiting discriminant information in nonnegative matrix factorization with application to frontal face veri?cation," IEEE Trans. Neural Netw., vol. 17, no. 3,pp. 683-695, Mar. 2006.

[17] J. C. Lee and E. Milios, "Matching range images of human faces," in Proc. IEEE Int. Conf. Image Processing, 1990, pp. 722-726.

[18] T. Russ, M. Koch, and C. Little, "A 2D range Hausdorff approach for 3D face recognition," in Proc. IEEE Workshop Face Recog. Grand Challenge Experiments, 2005, p. 169.

[19] S. Chen and Y. Zhu, "Subpattern-based principal component analysis,"Pattern Recognit., vol. 37, no. 5, pp. 1081-1083, May 2004.

[20] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Survey, vol. 35, no. 4, pp.

[21] C.S. Chua, F Han, K, Ho,3d human face recognition using point signature, in: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000,pp.233-238.

[22] http://www.umedm.com/Article/2008/20080109231904-1294.html

[23] L. Shapiro, J.M. Brady, Feature-based correspondence: an eigenvector approach, Image and Vision Computing 10 (5) (1992) 283-288.

[24] M. Carcassoni, R. Hancock, An improved point proximity matrix for modal matching, in: Proceedings of the Fifteenth IEEE International Conference on Pattern Recognition, 2000, pp.34-37.

[25] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge, Comparing images using the Hausdorff distance, IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (9) (1993) 850-863.

[26] B. Takacs, Comparing face images using the modified hausdorff distance, Pattern Recognition 31 (12) (1998) 1873-1881.