3D FACE RECOGNITION USING AFFINE INTEGRAL INVARIANTS

S. Feng, H. Krim, I. Gu†and M. Viberg†

ECE Dept., NCSU, Raleigh, NC.

†Signals and Systems Dept. Chalmers University, Goteborg, SW

ABSTRACT

A new 3D face representation and recognition approach is presented in this paper. Two sets of facial curves are extracted from a face range image, and a novel facial feature representation, the affine integral invariant, is introduced to mitigate the effect of pose on the facial curves. A human face is shown to be representable by a small subset of those affine integral invariant curves. A recognition procedure based on the Discriminant Analysis and Jensen-Shannon Divergence analysis is proposed. Substantiating examples are provided with an achieved classification accuracy of 92.57%

1. INTRODUCTION

Face recognition has been extensively studied for over 30 years, and its complexity together with its timely relevance in security and surveillance problems have recently led to a renewed research interest. Techniques in 2D face recognition abound, and Zhao and Chellappa [1] provide a fairly comprehensive account of the state of the art. Lighting and pose in 2D face recognition are widely recognized to be major impediments to the deployment of robustly performing algorithms. Specifically, the performance of many existing algorithms greatly deteriorates when the training and testing sets do not share a significant number of common views and lighting conditions. The recently developed 3D scanning techniques are believed to provide a potential to alleviate the limitation due to lighting and pose, and their rapid deployment would go far in paving the way for a viable recognition system. Exploiting such data amounts to extracting the intrinsic geometric information and utilizing it as the basis for characterizing individual faces.

Recent research activity in 3D face recognition techniques is reviewed in Chang *et al.*[2]. We can distinguish two main classes of data driven techniques:

- A nonparametric and statistical approach includes an Extended Gaussian Image (EGI) model by Lee and Milios[4], the Gaussian mixture-based Iterative Close Point algorithm of Cook et al. [8] as well as a PCA-based model proposed by Chang et al. [3].
- A more geometric approach includes a curvature feature-driven technique by Gordon [5], a curve and profile-based description of a face by Nagamine *et al.*[6] and Baumier [10], and a volumetric approximative representation of a face proposed by Irganoglu *et al.* [9]. Also related is the Multidimensional scaling (MDS)-based technique proposed by Bronstein, Bronstein and Kimmel [7] to construct an isometric transformation for 3D face analysis. Lu and Jain [11] jointly exploit range and texture information with a Linear Discriminant Analysis to construct a hierarchical system.

In this paper, we propose in Section 2 a novel technique which bases an accurate description of a face range image on a set of curves intrinsic to the 3D surface. To better contend with face pose variability, we propose to derive in Section 3 integral invariants as associated features to descriptive curves of a face . The affine invariance provide robustness of curves to rotation, scaling and shearing. With a set of such invariants in hand, we demonstrate that a relatively small set of such features for a face is sufficient for its accurate representation. Section 4 discusses the use of these features in a face comparison task by way of Discrimination Analysis and Jensen-Shanon divergence measure. We provide in Section 5 illustrating examples.

2. FACIAL CURVE FEATURE EXTRACTION

2.1. Data Set

The experimental data we are using in this paper consists of two data subsets from the University of Notre Dame Biometrics database. One subset henceforth referred to as S_1, comprising 10 subjects with 2 range images per subject, is selected as the experimental set to be analyzed by way of its extracted features. Another subset S_2 of 175 range images from 35 subjects is used as the testing set to validate the feature selection procedure in S_1 and to test its performance in a recognition task.

2.2. Facial Curve Extraction

The 3D range image data from the database are in a form of a point cloud as shown in Fig. 1-a. Using the 3D coordinates for all points in a given cloud (or face), we process the raw data to obtain a 3D triangular mesh as shown in Fig. 1-b.

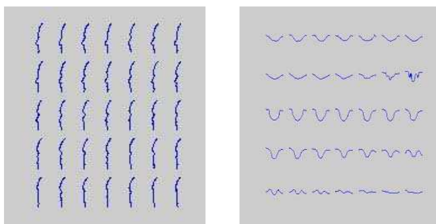
(a) (b) (c)

Fig. 1. *(a)Range image in database, (b) face mesh and (c) feature points location.*

The curve feature extraction proceeds in the vertical and horizontal directions for each representative mesh image embedded in a

Euclidean reference frame. The region of interest is delineated by five feature points, namely two outer corners of the eyes, the tip of the nose, the upper end of the eyebrows, and the lower end of the lower lip as illustrated in Fig. 1-c. While the localization of these feature points may be achieved automatically, we carry it out manually in this paper.

A feature curve is defined as the intersection of a plane with the surface of a face. A total of 35 vertical and 35 horizontal planes are used. The vertical planes are perpendicular to the straight line joining the two eye corners A and B as illustrated in Fig. 1-c. These are uniformly placed at equal distance from the the left corner of the face to the center of the face. The two eye corners A, B and the nose tip C define a reference plane. The Horizontal planes are perpendicular to the intersection line of the reference plane and the vertical plane, which range from the eyebrows D to the mouth E. The 70 vertical and horizontal curves which are collected for each face are shown in Fig. 2.



(a) (b)

Fig. 2. *vertical(a) and horizontal(b) curves.*

3. INTEGRAL INVARIANT

Curves may be subjected to transformations as a result of a variation in pose of a subject. The transformations include translation, scaling, rotation and shearing. This clearly impacts the performance of any face recognition procedure, whenever a different set of reference templates are invoked. It is hence important to rather describe these feature curves by invariants which are insensitive to any potential transformation.

The feature curves of interest here are plane curve, hence necessitating affine transformations. Affine Integral invariants were proposed by Hann and Hickman [12]and [13]. Such a transformation in 2D space may be defined as

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{1}$$

where $\begin{vmatrix} a & b \\ d & e \end{vmatrix} \neq 0$.

The corresponding affine integral invariant for a given curve $y = y(x)$ is defined in terms of the integrals of coordinate variables:

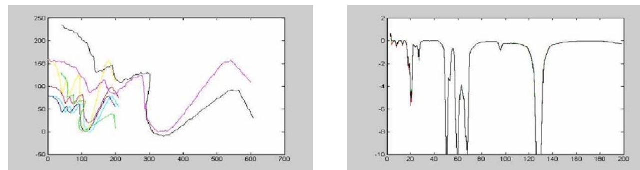
$$\begin{aligned} \kappa(x,y) &= \frac{1}{3}(-2(x-x_0)w + 6(y-y_0)v + 4z^2 \qquad (2) \\ &- 2(2xy - 2x_0y_0 + yx_0 - xy_0)z \\ &+ yy_0(x-x_0)^2)/((x-x_0)(y-y_0) - 2z^2) \end{aligned}$$

(x, y) is point on the facial curve, (x_0, y_0) is the starting point of the curve, where

$$z = \int_{x_0}^{x} y\,dx, \;\; v = \int_{x_0}^{x} xy\,dx, \;\; w = \int_{x_0}^{x} y^2\,dx.$$

The transformed coordinates (x', y') obtained by applying Eq. 1 to (x, y) thus lead to

$$\kappa(x, y) = \kappa(x', y').$$



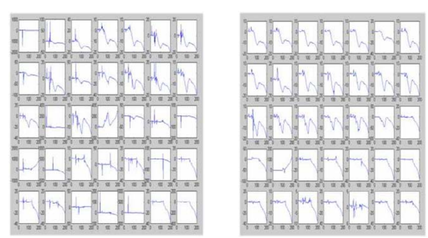
(a) (b)

Fig. 3. *(a)Various affine transformations of a curve (b)Corresponding and conciding affine integral invariants*

An example is shown in Fig. 3-a. The solid line curve is the original face profile curve whereas all other styled line curves are its affine transforms. The associated affine integral invariants are computed and displayed in Fig. 3-b and shown to coincide.

4. FEATURE CURVE SELECTION AND ANALYSIS

As previously noted, a total of 70 curves are extracted for each face in set $S1$ and 70 corresponding invariants are computed. The vertical set of invariants is shown in Fig. 4. Due to regions of missing data on the face range image, large (noise) spikes may appear as may be seen in Fig. 4(a). Smoothing of these spikes may be effected in one of two ways: interpolating in the raw data domain to smooth over the missing data regions, or by thresholding the invariants and thereby eliminating the spikes directly as demonstrated in Fig. 4(b). Upon obtaining the set of invariants, a natural question which arises



(a) (b)

Fig. 4. *integral invariant for one vertical curve set(a) and spike removal results(b).*

is about the discrimination power of such features among different faces as well as how to quantitatively evaluate such a measure. To that end, we propose a statistical robust approach which effectively evaluates the clustering of features for a given subject face. Two measures will be specifically exploited, one being that of a scatter ratio in Discriminant Analysis and the other is Jensen Shannon Divergence measure of dissimilarity between invariant curves [15, 14].

4.1. Discriminant Analysis

Discriminant analysis unveils directions that are efficient for discrimination. The entire invariant curve set is

$$\{\vec{\kappa}_{i,j}^k | i \in [1, N], j \in [1, M], k \in [1, K]\}$$

where i is the subject index set, k is the observation index for each subject, and j is the index of extracted curves of each face. For set S_1, N is 10, K is 2 and M is 70. The vertical invariant set is $\{\vec{\kappa}_{i,j}^k | k \in [1, 35]\}$ and the horizontal invariant set is $\{\vec{\kappa}_{i,j}^k | k \in [36, 70]\}$.

The overall mean of each feature invariant curve set is defined as:

$$\vec{m_j} = \frac{\sum_{i=1}^N \sum_{k=1}^K (\vec{\kappa}_{i,j}^k)}{KN}.$$

The mean of a feature invariant curve set for each subject is:

$$\vec{m_{i,j}} = \frac{\sum_{k=1}^K \vec{\kappa}_{i,j}^k}{K}.$$

The inter-class scatter $\tilde{S}_{Bj}$ and the intra-class scatter $\tilde{S}_{Wj}$ are defined as

$$\tilde{S}_{Bj} = \sum_{i=1}^N K(\vec{m}_{i,j} - \vec{m}_j)^T(\vec{m}_{i,j} - \vec{m}_j) \tag{3}$$

$$\tilde{S}_{Wj} = \sum_{i=1}^N \sum_{k=1}^K (\vec{\kappa}_{i,j}^k - \vec{m}_{i,j})^T(\vec{\kappa}_{i,j}^k - \vec{m}_{i,j}) \tag{4}$$

A good characteristic measure will exhibit a large inter-class scatter to better distinguish subjects, and a small intra-class scatter to reflect the similarity/coherence of feature for the same subject. The ratio J_j is a good candidate measurement for the quality of each feature.

$$J_j = \frac{\tilde{S}_{Bj}}{\tilde{S}_{Wj}} \tag{5}$$

4.2. Jensen-Shannon Divergence Analysis

As a special case of Jensen-Renyi Divergence[15], Jensen-Shannon Divergence is a powerful tool to measure the similarity of Probability Density Functions(pdf). Invariant curves may easily be normalized (e.g., by translation and scaling) to satisfy the non-negativity and the integrability properties of a pdf. Such processing simplifies the comparison of invariant curves to subsequently yield a quantitative measure of clustering of intra-class features relative to inter-class features. Thus identifying each $\vec{\kappa}_{i,j}^k$ with $p_{i,j}^k$, we may proceed to compute such a measure which we would expect to be large for two different subjects and small for a positive recognition. An inter-class divergence JSD_{Bj}, and an intra-class divergence $JSD_{Wi,j}$ yield a divergence ratio JSD_j. It is an alternative measurement for feature quality.

$$JSD_{Bj} = S\left(\sum_{i=1}^N \sum_{k=1}^2 \pi_i p_{i,j}^k\right) - \sum_{i=1}^N \sum_{k=1}^2 \pi_i S(p_{i,j}^k) \tag{6}$$

$$JSD_{Wi,j} = S\left(\sum_{k=1}^2 \pi_i p_{i,j}^k\right) - \sum_{k=1}^2 \pi_i S(p_{i,j}^k) \tag{7}$$

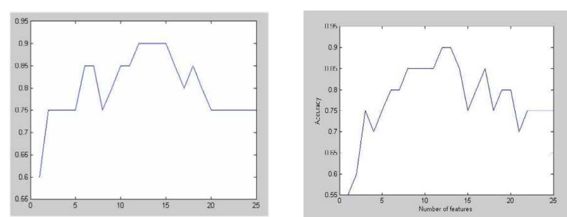
$$\text{where } S(p) = -\sum_{l=1}^L p_l log p_l \tag{8}$$

$$JSD_j = \frac{JSD_{Bj}}{\sum_{i=1}^N JSD_{Wi,j}} \tag{9}$$

4.3. Feature Invariant Selection

Scatter and divergence ratios reflect the importance of each feature. The number of relevant and required curve invariants to fully represent a face is also an important parameter to determine. The classification performance of a face recognizer may for instance be used to unravel this parameter. Such a classifier is implemented as a Nearest Neighbor (NN) Classifier in Euclidean Space using a L2 distance as the metric. On account of the small size of the data set, Leave-One-Out Cross (LOOC)-validation is adopted as the procedure of choice.

Starting with an invariant corresponding to the largest scatter ratio J_j, we progressively increase the set of accounted features by one in correspondence to a decreasing J, and simultaneously monitor the performance of the NN classifier with Leave-One-Out Cross-validation. We note that this performance dramatically changes with the first few features and subsequently reaches a state of small fluctuations as illustrated in Fig. 5-a.



(a) (b)

Fig. 5. *(a) Classification performance as a function of feature selection by Scatter ratio and (b) by JSD ratio.*

Carrying out the same procedure using Jensen Shannon Divergence ratio yields similar results shown in Fig. 5-b.

In comparing Figs. 5-a and Figs. 5-b, it is interesting to see that the graphs, albeit different both indicate that a *best* performance is achieved with a total of 12 curve invariants. The 12 curves selected by both approaches coincide albeit chosen in a different order. 10 out of the 12 feature curves are vertical curves, namely curve 10, 12, 15, 25, 27, 29, 30, 32, 33 and 35. The two horizontal curves are curves 15 and 30. The feature curve locations are depicted in Fig. 6.

That the vertical curves contribute to the classification more significantly as indicated by the above results, is an interesting point. Most of the distinguishable vertical feature are located near the face center profile (32, 33, 35), center (10, 12, 15) and corner (27, 29, 30) of eye regions. The two horizontal curves (15,30) also characterize the eye and nose. While the psychophysical interpretation of this result is certainly significant, we defer this discussion to a future paper.

4.4. Dimension Reduction

The dimension of each vertical and horizontal feature vector is 200, making each of the 12 feature curve invariants a 1x2400 long fea-

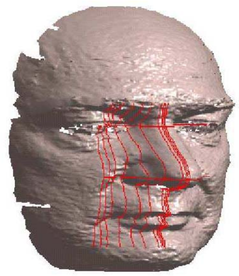


Fig. 6. *feature curve locations.*

ture vector. Fully realizing the presence of redundant information in such a vector, and towards reducing the computational load of our approach, we proceed to project our feature data onto a lower dimensional space with no loss in classification performance. To that end, we first construct a covariance matrix of feature vectors K_j for each curve set (at a location),

$$K_j = \left[\vec{\kappa}^1_{1,j},\ \cdots,\ \vec{\kappa}^K_{1,j},\ \cdots,\ \vec{\kappa}^1_{N,j},\ \cdots,\ \vec{\kappa}^K_{N,j}\right].$$

Performing a PCA analysis of the covariance matrix and identifying the most significant Eigen values defines our non-redundant space of interest. Our Experiments show that the first 8 eigenvectors may be used to characterize every feature invariant curve set to achieve the same performance as that using the full feature space. This procedure effectively reduces the complexity from a 2400 dimension to a 96 dimension.

5. FACE RECOGNITION EXPERIMENT

The analysis of set S_1 indicates that the 3D face may be represented by a 96 dimensional feature vector. The performance is to be evaluated using the testing set $S2$.

Out of the 70 feature curves for each face in set $S2$, the invariants of 12 feature curves are calculated and projected onto the eigenvectors in Sec. 4.4 to generate a 96 dimensional feature vector.

Proceeding as in Sec. 4.3, a 3-NN classifier with LOOC-validation is implemented. Euclidean distance is used to measure the dissimilarity and the class label is assigned by majority voting rule. Leave-One-Out cross-validation is a special case of k-fold cross validation when k is equal to the number of data in the testing set. Each of the the data is tested again all others and the accuracy is the average of the overall performance for every face.

Out of the 175 faces, 13 errors occur. The accuracy is 92.57%. Nearest Neighbor is a very basic classifier. The performance may be further improved by e.g., a Support Vector Machine Classifier. Our results are numerically comparable with recently proposed and geometry-driven techniques. The testing is unfortunately carried out on different databases [4]-[11] making a strict comparison so far difficult.

6. CONCLUSIONS

In this paper, we presented a new curve invariant based method for 3D face recognition. An affine integral invariant of each facial curve is not affected by translation, scaling, rotation and shearing distortion of 3D faces. Our experiments indicates that the human face can be characterized by 12 affine invariant curves, which are located near the face center profile, center and corner of eye regions. Each curve is projected onto a 8 dimensional space to construct a feature vector with a resulting performance by a 3-NN classifier of 92.57%.

7. ACKNOWLEDGEMENT

This work is in part supported by NSF and AFOSR. The authors would also like to thank the Swedish Foundation for International Cooperation in Research and Education. Furthermore, Prof. Kevin Bowyer and Prof. Patrick Flynn from University of Notre Dame are acknowledged for providing us the UND Biometric Database.

8. REFERENCES

[1] W. Zhao, R. Chellappa, P.J Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey", *ACM Computing Survey*, December Issue, pp. 399-458, 2003.

[2] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "An Evaluation of Multi-modal 2D+3D Face Biometrics", *IEEE Transactions on PAMI*, vol. 27, pp. 619-624, April 2005.

[3] K. Chang, K. Bowyer, and P. Flynn. "Face recognition using 2D and 3D facial data", *2003 Multimodal User Authentication Workshop*, 2003, pp. 25-32.

[4] J. C. Lee and E. Milios. "Matching range images of human faces", *ICCV*, 1990, pp. 722-726.

[5] G. Gordon. "Face recognition based on depth and curvature features", *CVPR* , 1992, pp. 108-110.

[6] T. Nagamine, T. Uemura, and I. Masuda. "3D facial image analysis for human identification", *International Conference on Pattern Recognition*, 1992, pp. 324-327.

[7] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. "Expression-invariant 3D face recognition", *Audioand Video-Based Person Authentication*, J. Kittler and M.S. Nixon, 2003.

[8] J. Cook, V. Chandran, S. Sridharan, C. Fookes. "Face recognition from 3D data using Iterative Closest Point algorithm and Gaussian mixture models", *3D Data Processing, Visualization and Transmission*, 2004, pp. 502-509

[9] M. O. Irfanoglu, B. Gokberk, L. Akarun. "3D shape-based face recognition using automatically registered facial surfaces", *ICPR*, 2004, pp. 183-186

[10] C. Beumier . "3D face recognition", *Computational Intelligence for Homeland Security and Safety*, 2004, pp. 93-96

[11] X. Lu and A. K. Jain, "Integrating Range and Texture Information for 3D Face Recognition," *Proc. 7th IEEE Workshop on Applications of Computer Vision*, pp. 156-163,

[12] C. Hann, C. E. Hickerman, "Projective curvature and integral invariants", *Acta applicandae mathematicae*, vol.74, pp. 177-193, 2002.

[13] J. Sato and R. Cipolla, "Affine integral invariants and matching of curves", *ICPR*, 1996, pp. 25-29.

[14] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Trans. Inform. Theory*, vol. 37, pp.145-151, 1991.

[15] Y. He, A. Ben Hamza, H. Krim, "A Generalized Divergence Measure for Robust Image Registration", *IEEE Transactions on Signal Processing*, vol. 51, pp. 1211-1220, 2003