A COMPARISON OF DATA REPRESENTATION TYPES, FEATURE TYPES AND FUSION TECHNIQUES FOR 3D FACE BIOMETRY

Helin Dutağacı[(1)], Bülent Sankur[(1)], and Yücel Yemez[(2)]

[(1)] Department of Electrical-Electronics Engineering, Boğaziçi University
Bebek, 34342, Istanbul, Turkey
phone: + (90) 212 359 6414, fax: + (90) 212 287 2465, email: {dutagach, bulent.sankur}@boun.edu.tr
web: http://busim.ee.boun.edu.tr/~sankur/

[(2)] Department of Computer Engineering, Koç University
Rumeli Feneri Yolu, Sarıyer, 34450, Istanbul, Turkey
phone: + (90) 212 338 1585, fax: + (90) 212 338 1548, email: yyemez@ku.edu.tr
web: http://network.ku.edu.tr/~yyemez/

ABSTRACT

This paper focuses on the problems of person identification and authentication using registered 3D face data. The face surface geometry is represented alternately as a point cloud, a depth image or as voxel data. Various local or global feature sets are extracted, such as DFT/DCT coefficients, ICA- and NMF- projections, which results in a rich repertoire of representations/features. The identification and authentication performance of the individual schemes are compared. Fusion schemes are invoked, to improve the performance especially in the case when there are only few samples per subject.

1. INTRODUCTION

Improvements in 3D imaging devices have enabled 3D face data to be a promising new biometric modality. The geometrical information of 3D facial is more robust to such adverse effects as illumination conditions, small pose variations, and skin color cosmetics, all of which are challenging problems for 2D face recognition systems. On the other hand, 3D face data may suffer from lower resolution and the uncertainty at low-reflection points such as facial hair.

In this work, we address identification and authentication problems based solely on 3D face data in a fusion context. Identification corresponds to the person recognition without the user providing any information. The system must assign an identity from among the enrolled faces in the database. In authentication, the system tests the veracity of the claim of the user resulting in "accept" or "reject" decisions.

Research on 3D face biometrics is relatively recent. Lee and Milios [1] modeled the curvature information on the face by Extended Gaussian Image (EGI). Tanaka et al. [2] utilized EGI, by including principal curvatures and their directions and compared EGI's using Fisher's spherical correlation. Chua et al. [3], have used point signatures, a free form surface representation technique. Beumier et al. [4] extracted central and lateral face profiles, and compared curvature values along these profiles. Chang et al. [5] developed a multi-modal face recognition system by applying Principal Component Analysis (PCA) on both 3D range data and intensity images. Another multi-modal scheme based on PCA can be found in [6]. Bronstein et al. [7] focused on 3D face recognition algorithms that are invariant to facial expressions.

The focus of this work is on the fusion of the diverse 3D face recognition techniques described in [8] where we obtained very promising identification results. For example, we achieved 99.8% identification performance provided that the training data have at least four samples of each individual. However, as the training size drops, the performance deteriorates proportionally, hence the need to invoke fusion methods in the more realistic case when there could exist, at most, two or three samples per subject. In addition we also address the authentication (verification) problem.

In section 2 we summarize the representation schemes and feature types. Section 3 introduces the fusion techniques that we use in this work and describes the matching schemes for identification and authentication. In section 4, the experimental results are reported. Finally, we provide concluding remarks in Section 5.

2. REPRESENTATION TYPES AND FEATURES

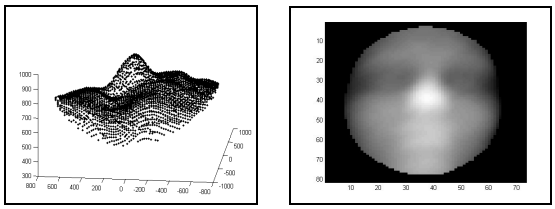
The range images employed in this work are registered and cropped by the algorithm described in [9]. Table 1 gives a summary of the face recognition schemes, which are comparatively assessed in [8]. We summarize in the sequel the 3D data representation types and the features extracted.

2.1 Representation Types

3D Point Cloud: The 3D point cloud is the set of registered 3D coordinates of the range data, where all the correspondences are defined (Figure 1.a).

2D Depth Image: The 2D depth image is obtained by projecting the point cloud onto a regular X-Y grid. The Z-coordinates then constitute the depth image $I(x, y)$, much like an intensity image (Figure 1.b).

3D Voxel Representation: A 3D binary function is defined on a voxel grid using the point cloud. A voxel is assigned a value "1" if it comprises a range point and "0" otherwise. The mapping is then smoothed by applying 3D distance transform, as depicted in Figure 2.



(a) (b)

Figure 1 – (a) 3D point cloud. (b) 2D depth image

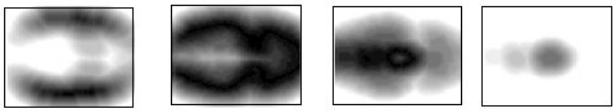


Figure 2 – Slices from the voxel representation

2.2 Features

2.2.1 Independent Component Analysis

We apply independent component analysis (ICA) both to the point cloud and to the depth image. The coordinates of the point cloud or the values of the depth image $I(x, y)$ of each face in the training set, are concatenated to form a single vector which in turn forms one of the columns of the data matrix. The vector dimension is first reduced via PCA and the data matrix $\mathbf{X}$ is then subjected to the FastICA algorithm [10]. ICA yields the mixing matrix, $\mathbf{A}$, and the independent coefficients, $\mathbf{S}$, so that:

$$\mathbf{X} = \mathbf{AS}$$

The columns of the mixing matrix, $\mathbf{A}$, form a basis for the data, and the columns of $\mathbf{S}$ are used as the ICA-based feature vectors of the corresponding faces. The length of the feature vector is determined by the PCA projection. We denote the ICA-based schemes applied to point clouds and depth images as **PC-ICA** and **DI-ICA**, respectively.

2.2.2 Non-negative Matrix Factorization

Non-negative matrix factorization (NMF) is a technique to obtain parts-based representation of images, which is commonly used for 2D face recognition. In this work we extend this approach to 3D face recognition, using 3D point cloud coordinates.

The NMF algorithm results in two non-negative matrices, $\mathbf{W}$ and $\mathbf{H}$ such that,

$$\mathbf{X} \approx \mathbf{WH}$$

where $\mathbf{X}$ is the non-negative data matrix. The columns of $\mathbf{W}$ can be regarded as basis faces, whereas the columns of $\mathbf{H}$ serve as feature vectors of the faces in the training set. In order to obtain the feature vector of an input test face, the pseudo-inverse of $\mathbf{W}$ is multiplied by the data vector representing the input face. We construct the data matrix in a similar manner as in the ICA-based scheme, however, in this case we do not apply PCA to the original data matrix. The data matrix is normalized so as to have non-negative values. We use the multiplicative update rules [11] to obtain the matrices $\mathbf{W}$ and $\mathbf{H}$. We denote the NMF-based scheme applied to a point cloud as **PC-NMF**.

Table 1 - Representation types and features

Representation	Features
3D Point Cloud	▪ ICA (**PC-ICA**) ▪ NMF (**PC-NMF**)
2D Depth Image	▪ Global DFT (**DI-DFT**) ▪ Global DCT (**DI-DCT**) ▪ Block-based DFT (**DI-block-DFT**) ▪ Block-based DCT (**DI-block-DCT**) ▪ ICA (**DI-ICA**)
3D Voxel Representation	▪ 3D DFT (**V-DFT**)

2.2.3 Global DFT/DCT

We apply 2D Discrete Fourier Transform (DFT) to the depth image and extract the first KxK DFT coefficients from the coefficient matrix. The real and imaginary parts of the coefficients are concatenated to construct the feature vector of size $2K^2-1$. Likewise, the DCT-based features are extracted by applying 2D Discrete Cosine Transform to the depth image. The first KxK coefficients are concatenated into the feature vector. In this case the feature vector is of size K^2, since DCT coefficients are real. These two schemes will be referred to as **DI-DFT** and **DI-DCT** in the rest of the paper.

2.2.4 Block-based DFT/DCT

Block-based DFT/DCT-schemes have previously been applied to intensity images for face recognition [12]. In this work, we apply these schemes to depth images. We partition the depth image into blocks and calculate the DFT or DCT on each block separately. Then we extract the first KxK coefficients and form the individual feature vector of each block. The fusion is conducted at feature level. The first KxK DFT/DCT coefficients of the blocks of a face are concatenated to form one single feature vector. These block-based schemes will be referred to as **DI-block-DFT** and **DI-block-DCT**.

2.2.5 Three-dimensional DFT

The last face representation that we consider is constructed based on the 3D distance transform of the surface voxel data. The real and imaginary parts of the first KxKxK coefficients obtained from the 3D-DFT of the distance transform are concatenated to form a feature vector. The resulting feature vector of size $2K^3$-1 is referred to as **V-DFT**.

3. FEATURE MATCHING AND CLASSIFIER FUSION

We use linear discriminant analysis for matching features. Each class (person identity) in the training set is modeled by

a multivariate normal density. The covariance matrix is estimated using the whole data in the training set, since we have very small number of training samples per class. The distance of a test feature to the class means are calculated using the discriminant functions. The class with the smallest distance is selected as the identity of the input face.

For authentication, the distance to the claimed class mean is calculated. If the distance is below a threshold, the input face is accepted. Since the linear discriminant analysis minimizes the within-class scatter and maximizes the between-class scatter in the training data, we need an additional "training" set, to determine the threshold. The false rejection and false acceptance rates are calculated on several threshold values, and the one that gives the equal error rate is determined as the threshold for the authentication system. The false acceptance and rejection errors are then evaluated on a separate test set.

We have considered three fusion techniques to improve the performance of the individual schemes used for the identification and authentication tasks. All fusion techniques that we employ are performed at decision level as described in the sequel.

3.1. Consensus Voting

In consensus voting, the identity with the highest number of votes is declared as the identified subject. For authentication, each method independently accepts or rejects the claimed identity, and the consensus is formed on the most favorable mode.

3.2. Borda Count

Each method assigns its own rank to all the faces in the database based on their distances to the input face. The ranks from individual schemes are summed up to obtain a final rank for each face in the database. Then the identity of the face is declared to be the one with the highest rank. Borda count is not applicable to the authentication problem, since it involves comparison with all faces in the database.

3.3. Sum Rule

Each method produces a distance value from the test face to the classes. The distance values are first mapped to the range [0, 1] and then summed up to obtain the final distance.

4. EXPERIMENTAL RESULTS

4.1. Results for individual schemes

We have used the 3DRMA data set [13] in our experiments. The 3DRMA data set contains face scans of 106 subjects with 5-6 sessions per subject. The total number of faces is 617. The faces are registered using the ICP algorithm described in [9].

The point cloud representation contains 3389 points. The depth images are of size 81x73, and the size for the voxel representation is selected as 64x64x64.

Table 2 gives the feature vector dimension utilized for each scheme. The size of the feature vector varies with respect to the number of training samples per subject, since the feature dimension that LDA can handle is constrained with the size of the training set. The same number of features is used for identification and authentication.

We have altered the training and test sets, and conducted N-fold experiments. For the cases with two and three training samples per subject, we have conducted 10 experiments. For the case with four training samples, the fold number was 5. Table 3 gives the identification results averaged over the number of experiments. The best values are obtained with ICA and NMF-based features obtained from point cloud representation. When applied to depth images ICA-based method yields poorer results as compared to the other methods. This relatively poor performance is due to the high energy variability at the borders of the faces among the depth images. This variability is modeled by ICA but has little contribution to recognition.

Table 2 – Feature vector dimension

Schemes	2 training samples	3 training samples	4 training samples
PC-ICA	30	40	50
PC-NMF	30	30	50
DI-DFT	49	71	127
DI-DCT	49	81	121
DI-block-DFT	84	84	84
DI-block-DCT	108	108	108
DI-ICA	60	80	50
V-DFT	53	127	127

Table 3 – Identification results for individual schemes

Schemes	2 training samples	3 training samples	4 training samples
PC-ICA	96.8	98.8	99.8
PC-NMF	96.7	98.7	99.8
DI-DFT	90.5	96.1	98.2
DI-DCT	88.3	94.7	96.7
DI-block-DFT	93.8	97.3	98.8
DI-block-DCT	93.5	97.9	98.2
DI-ICA	84.0	93.9	96.8
V-DFT	88.4	94.1	98.3

For authentication, two or three samples per subject are used for training the linear discriminants. One sample per subject is reserved for determining the threshold and the rest are used for testing. Table 4 and 5 display the authentication performances for the cases with (2+1) and (3+1) training samples per subject, respectively. The second column gives the equal error rates over the training set used for finding the optimal threshold. The third and fourth columns give false rejection and false acceptance rates, respectively, when this threshold is applied to the test set.

We observe that the DFT/DCT based schemes yield authentication errors above 5%. The intra-distance measures for these schemes contain outliers (Figures 3 and 4), which affect the threshold computation in an unfavorable way, whereas ICA-based features extracted from point clouds yield compact intra-distance distributions (Figure 5).

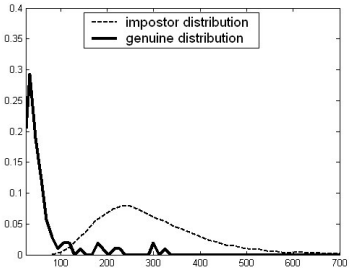


Figure 3 – Impostor and genuine distributions for **DI-block-DFT** scheme (obtained from a single experiment)

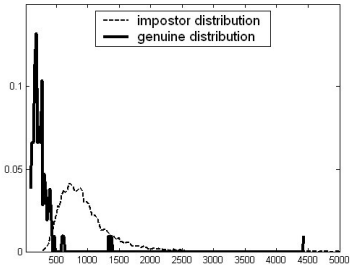


Figure 4 – Impostor and genuine distributions for **V-DFT** scheme (obtained from a single experiment)

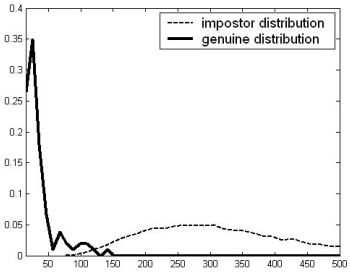


Figure 5 – Impostor and genuine distributions for **PC-ICA** scheme (obtained from a single experiment)

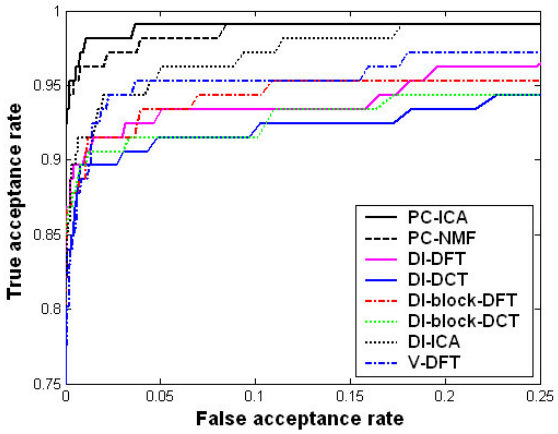


Figure 6 – ROC curves for individual schemes

Table 4 – Authentication results for 2+1 training samples per subject

Schemes	**Equal Error Rate (%) (Training)**	**False Reject. Rate (%) (Test)**	**False Accept. Rate (%) (Test)**
PC-ICA	2.3	2.4	2.3
PC-NMF	2.7	2.9	2.7
DI-DFT	8.5	9.8	8.0
DI-DCT	8.4	11.8	8.0
DI-block-DFT	6.8	6.1	7.1
DI-block-DCT	7.6	6.5	7.8
DI-ICA	5.8	9.2	5.6
V-DFT	6.5	9.7	6.8

Table 5 – Authentication results for 3+1 training samples per subject

Schemes	**Equal Error Rate (%) (Training)**	**False Reject. Rate (%) (Test)**	**False Accept. Rate (%) (Test)**
PC-ICA	1.8	1.7	1.6
PC-NMF	2.2	1.7	2.3
DI-DFT	7.5	8.9	7.3
DI-DCT	9.5	10.5	9.3
DI-block-DFT	6.1	4.7	6.1
DI-block-DCT	6.9	4.8	6.9
DI-ICA	3.9	7.1	3.5
V-DFT	7.1	8.5	7.5

4.2. Results with fusion algorithms

Tables 6 to 9 display the identification results obtained with fusion algorithms. In each case, a different combination of schemes is used. When all the schemes are fused, the results do not further improve, except for a slight increase obtained with the consensus voting technique.

Fusing the best two and three methods (Table 7 and 8) outperforms the individual techniques. We obtain the highest performance (99.2%) with consensus voting of **PC-ICA**, **PC-NMF** and **DI-block-DCT** methods.

Table 6 – Identification results with fusion of all schemes

Fusion techniques (All schemes are fused)	**2 training samples**	**3 training samples**
Max. individual perfor.	96.8	98.8
Consensus Voting	96.7	99.0
Borda Count	94.5	97.7
Sum Rule	96.7	98.8

Table 7 – Identification results with fusion of **PC-ICA** and **PC-NMF**

Fusion techniques (PC-ICA+PC-NMF)	**2 training samples**	**3 training samples**
Max. individual perfor.	96.8	98.8
Borda Count	96.9	98.9
Sum Rule	97.2	99.1

Table 8 – Identification results with fusion of **PC-ICA**, **PC-NMF** and **DI-block-DCT**

Fusion techniques (PC-ICA+PC-NMF+ DI-block-DCT)	**2 training samples**	**3 training samples**
Max. individual perfor.	96.8	98.8
Consensus Voting	97.1	99.2
Borda Count	96.7	98.8
Sum Rule	97.3	99.0

Table 9 – Identification results with fusion of weak methods

Fusion techniques (6 weakest methods)	**2 training samples**	**3 training samples**
Max. individual perfor.	93.8	97.9
Consensus Voting	95.5	98.6
Borda Count	93.5	97.7
Sum Rule	95.6	98.3

For authentication, we have fused PC-ICA and PC-NMF techniques using the sum rule. With fusion, the average of false acceptance and false rejection errors reduces from 2.4 to 1.1% for the case with three training samples per subject, and from 1.7 to 1.0 for the case with four training samples per subject. Figure 7 shows the ROC curves obtained by fusion of these best two techniques. The ROC curve of the individual schemes are also provided on the same plot.

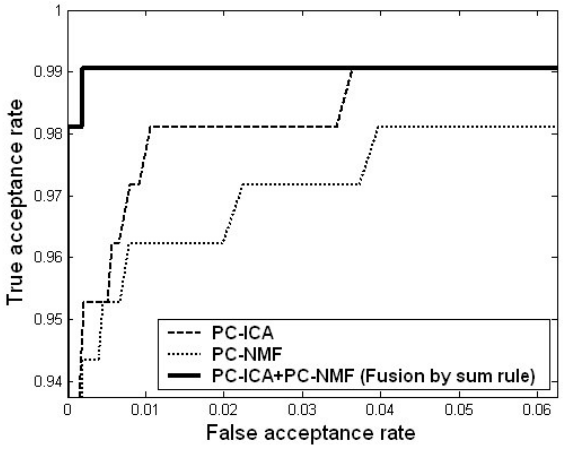


Figure 7 – ROC curves for **PC-ICA** , **PC-NMF** and their fusion

5. CONCLUSION

We have tested various representation types and feature extraction methods, both for recognition and authentication. The experimental results show that the best features are NMF and ICA-based features extracted from the point clouds. We have also conducted fusion experiments. The two best methods that individually result in the highest performances, namely PC-ICA and PC-NMF, give further improved results when fused using borda count and sum rule. When relatively weak methods are fused, the individual identification performances are also outperformed (Table 9). The methods we have proposed for registered 3D face data are very promising in that they achieve 99 % correct identification and with 1.8 % equal error rate at their best. There is evidence of further room for improvement using fusion schemes [14].

REFERENCES

[1] J. C. Lee and E. Milios, "Matching range images of human faces," in *Proc. IEEE International Conference on Computer Vision*, 1990, pp. 722-726.

[2] H. T. Tanaka, M. Ikeda, and H. Chiaki. "Curvature based face surface recognition using spherical correlation - principal directions for curved object recognition," in *Proc. Int. Conf. on Automatic Faceand Gesture Recognition*, 1998, pp. 372–377.

[3] C. S. Chua, F. Han, and Y. K. Ho, "3D face recognition using point signature", In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, March 2000, pp. 233-238.

[4] C. Beumier, M. Acheroy, "Face verification from 3D and grey-level clues", *Pattern Recognition Letters*, Vol. 22, 2001, pp 1321-1329.

[5] K. Chang, K. Bowyer, and P. Flynn, "Face recognition using 2D and 3D facial fata," in *IEEE International Workshop on Analysis and Modeling of Faces and Gestures. Nice*, France. October 2003.

[6] F. Tsalakanidou, D. Tzocaras, and M. Strintzis, "Use of depth and colour eigenfaces for face recognition," *Pattern Recognition Letters*, 24:1427-1435, 2003.

[7] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, „Expression-invariant 3D face recognition," *Audioand Video-Based Person Authentication (AVBPA 2003)*, LCNS 2688, J. Kittler and M.S. Nixon, eds, 2003, pp. 62-70.

[8] H. Dutağacı, B. Sankur and Y. Yemez, "3D Face Recognition by Projection Based Methods," *SPIE Electronic Imaging Conference*, San Jose, January 2006.

[9] M. O. İrfanoğlu, B. Gökberk, and L. Akarun, "3D shape-based face recognition using automatically registered facial surfaces," *Proceedings of the 17th International Conference on Pattern Recognition (ICPR2004)*, 2004, Cambridge.

[10] A. Hyvarinen and E. Oja, "Independent Component Analysis: Algorithms and Applications," *Neural Netw*orks 13 (4-5), 2000.

[11] D. D. Lee, and H. S. Seung., "Algorithms for nonnegative matrix factorization", Advances in Neural Information Processing Systems 13. 2001.

[12] H.K.Ekenel, R. Stiefelhagen, "Local appearance based face recognition using discrete cosine transform," *13th European Signal Processing Conference (EUSIPCO 2005)*, Antalya Turkey, September 2005.

[13] http://www.sic.rma.ac.be/~beumier/DB/3d_rma.html

[14] B. Gokberk, H. Dutagaci, B. Sankur, L. Akarun, "Fusion of Face Experts at the Decision-level for 3D", submitted to IEEE Trans. on System, Man and Cybernetics, 2006.