



FACE IDENTIFICATION ALGORITHM BASED ON MESH-DERIVED SYNTHETIC LINEAR DESCRIPTORS

R. Kh. Sadykhov ¹⁾, V. A. Samokhval ²⁾

¹⁾ A Computer Department of the Belarusian State University of Informatics and Radioelectronics, 6 P.Brovka st., Minsk, 220027, rsadykhov@gw.bsuir.unibel.by, tel./fax: (+375) 17 231 09 82

²⁾ A United Institute of Informatics Problems National Academy of Science of Belarus, 6 Surganov st., Minsk, 220012, sam@lsi.bas-net.by, tel.: (+375) 17 239 80 39

Abstract: *This paper presents appearance-based face identification algorithm by means of synthetic linear filters. The objective of our research is to construct facial descriptor in the form of linear filter, which should produce high and low outputs for intra- and inter-class recognition problem correspondingly. This filter can be synthesized from 2,5D sparse mesh derived from a given set of images of a person. As ever the filter is created it is then used as facial descriptor, i.e. serves as personal ID for face identification.*

Keywords: *appearance-based approach, face recognition, synthetic linear descriptors, discriminant function.*

1. INTRODUCTION

Face recognition – a well-known research topic in computer vision with a rather long history [1], still remains a very challenging problem today. The problem has been intensively researched in recent years. Researches in face recognition have been motivated by both their scientific values and wide potential applications in public security, law enforcement and commerce. Crowd surveillance, electronic line-up, store security and mug shot matching are some of the security applications. Much progress in face recognition has been made in the past few years [2]. However, face recognition remains difficult, unsolved problem in general [3, 4]. Most of the current systems work under constrained conditions, even requiring the subject to be highly cooperative. Obviously, challenges lie in not only the academic level but also in the level of application system design.

The problem in itself is usually formulated in dependence of real-world application tasks. One of them is face identification, i.e. it is reduced in fact to the matching “one-to-many”. In the identification case, the system has a database that consists of the facial images of the persons that have legitimate access to restricted areas. Identification system tries to match the person's face to the system database and be able to decide whether the person is among the objects listed in the database.

System performance in face recognition depends critically on several key factors. While detecting, the system must locate the face in the image captured by the camera. The next step is face cropping/

alignment. The system must be able to crop the face and remove the unwanted background. Some alignment procedures (rotating, shifting and scaling) are also required as usual.

To minimize light effects special procedures for lighting adjustment should be carried out. Face recognition system must be able to normalize ambient lighting.

In the final stage of classification and matching recognition system must be able to classify the unknown face image for identification/verification task.

The approaches to face recognition have covered sources from 2D intensity or color images up to 4D face data [5]. With that high computational and spatial costs make limitations to use these approaches in real application systems.

In the early stages of research on face recognition, facial geometry was used as the discriminating measure even though it was very difficult to extract precise facial geometry from 2D images. Moreover, human facial geometry has few discriminating factors to distinguish individuals. Therefore, appearance-based approach that uses image of a face itself seems to be preferred. Face recognition researches have focused on the representation of face appearances for efficient matching.

The classical methods used for face representation are the Principal Component Analysis (PCA) [6] and Linear Discriminant Analysis (LDA) [7] with a lot of their modifications. The Independent Component Analysis (ICA) [8] has also been investigated in the context of face representation.

The Linear Discriminant Analysis is a popular pattern recognition method, and some LDA-based face recognition systems [9, 10] have been developed in the last decade and encouraging results have been achieved. This method has suitable and useful modifications obtained by use of different kernels or introducing non-linearity. However, this method suffers from a well-known small size problem.

Appearance-based face recognition algorithms utilize the intensity or intensity-derived features of original images. The dimensionality of feature vector used by these methods is often very high while the training sample size is relatively small. Such training data can lead the classifier to be biased and have a large variance, resulting in a poor performance [11]. To improve classification characteristics a number of approaches have been presented [12, 13].

In many practical tasks it is required to represent the face images in terms of a small number of parameters of poses rather than their own face images due to its effectiveness for data representation, storage and transmission. However, the relationship between the face appearance and the pose parameters is not matched well in the original high-dimensional data space because variations of face images are characterized by a complicated non-linear manifold. The standard way is to project the face image into a lower dimensional space.

Reduction of dimensionality can be performed on the base of different approaches. Common used techniques are principal component analysis, multidimensional scaling (MDS). Classical MDS finds an embedding that preserves the interpoint distances, equivalent to PCA when those distances are Euclidean. The major algorithmic features of PCA and MDS are computational efficiency, global optimality, and asymptotic convergence with the flexibility to learn a broad class of manifolds.

Being consistent with appearance-based approach which does not require selection of specific facial features, we build recognition model directly from the image data. In our approach a number of facial images of given person (object class) are considered as a cluster in high-dimensional space and the separation between classes is unknown. The idea is contained in the following: given set of feature vectors (members of some object class) find vector-function which produce equal output for each of these vectors. This set of vectors is called the training data. All the other class exemplars are test set of data. The more training exemplars are available for given object class the better final result of intra-class recognition. As to inter-class separability, high value for it is not always ensured and the problem is our ongoing investigation. The basis of proposed approach is to model human face as approximate parameterized 2,5D sparse mesh

which could fill up training set well enough to achieve reliable identification.

2. ALGORITHM OVERVIEW

The most valuable variations in face images are induced by different lightning, pose and expression. Therefore the proposed algorithm requires a number of training facial images for given person obtained under different conditions. In order to eliminate/minimize noise introduced by different background, illumination and scale, special procedures were carried out at the preprocessing stage. At first, face area is limited by elliptic mask and rescaled to standard size (Fig. 1). Geometrical standardization concerns the between-eye distance and its position in the area of elliptic mask. One can choose different ways both how do these and what facial parameters have to be preserved.

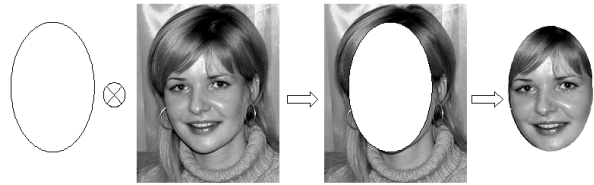


Fig.1 – Selection region of interest. Procedure is aimed to eliminate background noise providing essentially better identification results.

To minimize illumination effects on final classification results all images are transformed into standard dynamic range of intensity. Conditional transform for image point $I(x,y)$ should yield to new intensity value $I'(x,y)$ and may be written in simple form as

$$I'(x,y) = A (I(x,y) - I_{min}) / (I_{max} - I_{min}) \quad (1)$$

with constraint $I_{max} \neq I_{min}$ and amplitude A defines the dynamic range of illumination (intensity) and usually equals to 255 in correspondence with gray scale. Obtained image $I'(x,y)$ is then converted into parameterized 2,5D sparse mesh.

Here should be pointed out that sparseness property of the proposed image representation by these meshes is of great importance for synthetic discriminant function algorithm. It provides the ability to vary the dimensionality of feature space used allowing both to reduce and to extend it if necessary.

Preprocessed images of given person p form class-specific set of feature vectors $\{x_1, x_2, \dots, x_m\}^p$ which are then considered as functions of a basis on which required filter has to be linear decomposed on. Formalization comes to matrix equation for

linear filter, specifying conditions needed the task to be resolved (Fig.2).

3. FORMALIZATION

Approaches with using synthetic discriminant functions for pattern recognition are good practice and suitable technique. There are many types of pattern recognition problems and there exists many ways to construct synthetic discriminant functions. Comprehensive methodology of the subject for many different cases of classification task is given

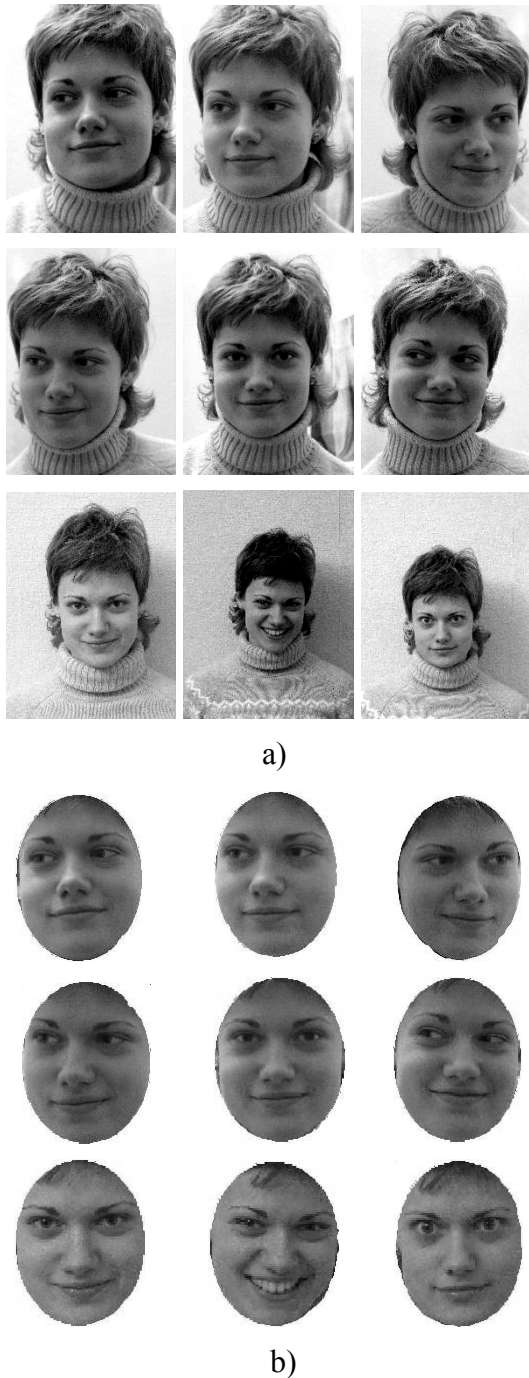


Fig. 2. Feature vectors for given person K are created from the training set of images (a) by selecting elliptical areas of interest, normalizing in scale and brightness (b) and constructing parameterized 2,5D sparse mesh (c).

by David Casasent in [14]. We will concentrate our attention on such types of SDF, which produce equal output for class members. In [14] these filters are called ECP SDF or Equal Correlation Peak Synthetic Discriminant Function.

We will now consider general approach of how the synthetic discriminant function (in the form of linear filter) can be constructed.

Consider image set $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ of some object class, where all \mathbf{x}_i $|_{i=1, \dots, m}$ are $n \times 1$ column vectors. To synthesize filter \mathbf{f} we form linear equation

$$W \mathbf{f} = \mathbf{u}, \quad (2)$$

where $m \times n$ matrix $W = \{\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_m^T\}$ is constructed from feature vectors of the training set

of a given object class, filter \mathbf{f} is $n \times 1$ vector and \mathbf{u} is $m \times 1$ vector of desired outputs u_i . The main idea is that of the filter \mathbf{f} produces equal values for all samples from the training set, because they are members of the same object class. Thus, $u_i = u_j = u$. It can be shown, that if the filter \mathbf{f} is linear combination of m training vectors, the decision of (2) is given by pseudo-inverse of W as follows

$$\mathbf{f} = W^T (W W^T)^{-1} \mathbf{u} \quad (3)$$

Define matrix $R_{im} = W W^T$ as square $m \times m$ matrix and which is correlation matrix of images. Now equation (3) can be rewritten as

$$\mathbf{f} = W^T R_{im}^{-1} \mathbf{u}. \quad (4)$$

In general case m different vectors (regardless of they are samples of the same object class) may be expected linear independent, and if $m \leq n$ matrix R_{im} is of full rank, the inversion R_{im}^{-1} exists and delivers the single decision of (4).

4. EXPERIMENTAL RESULTS

For computer simulation special database for 17 persons was prepared. It contains over 230 images with $11 \div 13$ samples per class. Additional tests were also made with ORL Face Database of 400 facial images, 10 images for every 40 person. The differences between ORL classes are mainly in pose and expression. To evaluate recognition rate of proposed method we construct facial descriptors for 10 randomly chosen object classes.

To demonstrate the changes of recognition rate we investigate two series of filters. In the first, training set consists of 3 images, while in the second it has 5 images of each class. Test sets consisted of 7 and 5 images for these two series of experiment respectively. All images both in training and testing sets were preprocessed as described in section 2. Then facial descriptors were synthesized for each of 10 object classes using available training data. Once the system was trained, it was ready to be used for face recognition. To identify input facial image we must obtain its projection onto class-derived synthetic descriptor. Mathematically it is the inner product of the descriptor and the image under the test. The closer the value of inner product to '1' (identified as a 'member') the better classification rate obtained. Otherwise we obtain 'nonmember' and the case of discrimination. To further formalize the classification procedure we have to define decision boundaries. Since we are free to choose outputs u_k in (2), here k is class number, we have to assign $u_k = k$ as it is shown by Z-axis in Fig. 2. Thus, decision boundaries for k -th class are defined simply as low and high values equal to $k-1/2$ and $k+1/2$ respectively. This definition leads the testing image be accepted as 'member' if the filter output falls into interval $]k-1/2; k+1/2[$, and as 'nonmember' otherwise. Series of experiment in Fig. 3 represent classification performance of descriptors in the term of inner product value (z-axis) obtained with all members of given class (y-axis, 1-10).

First series (top) represent use of the descriptors constructed on 3-member training set. The second series (bottom) represent descriptors, trained on 5 images. The improvement in classification performance appears from top to bottom.

Experimental results show high recognition rate for intra-class identification and low true rejection for inter-class recognition as well. Furthermore, one can see significant improvement of classification performance when more data are included in the training set. This fact confirms experimentally the initial idea that the more information on the object we have the better the final result.

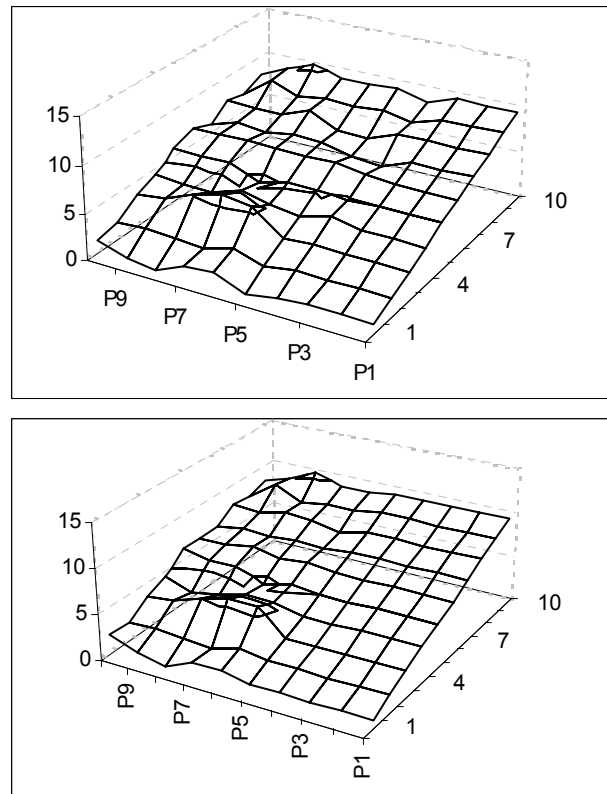


Fig.3. Two series of experiment with synthetic discriminant filters, constructed for ten randomly chosen object classes (x-axis, P1-P10).

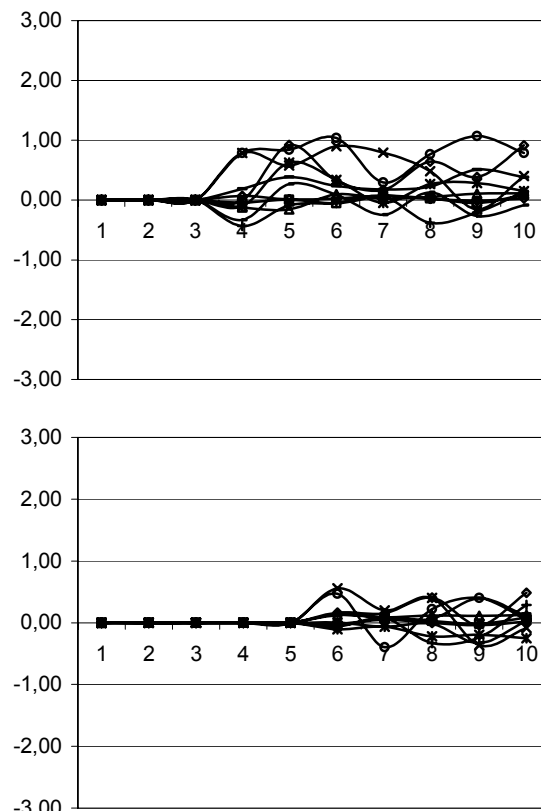


Fig.4 – Intra-class recognition. Personal synthetic discriminant filters, constructed for ten randomly chosen object classes (x-axis, P1-P10), using three (top) and five (bottom) training images. One can find classification improvement from top to bottom while using more representative training set.

In order to evaluate intra- and inter-class recognition performance it is convenient to reduce results to common scale. Zero value of inner product will indicate a 'member', and decision boundaries falls into value interval $]-0.5; +0.5[$. Classification results for intra- and inter-class recognition are shown in Fig. 4 and Fig. 5 respectively. Intra-class identification demonstrates almost excellent results especially in the case of extended training set (Fig. 4 bottom), while the results of inter-class problem look as less convincing example. Investigation of the robustness of proposed algorithm for multi-class recognition and searches of ways to improve classification performance are the subjects of our future works.

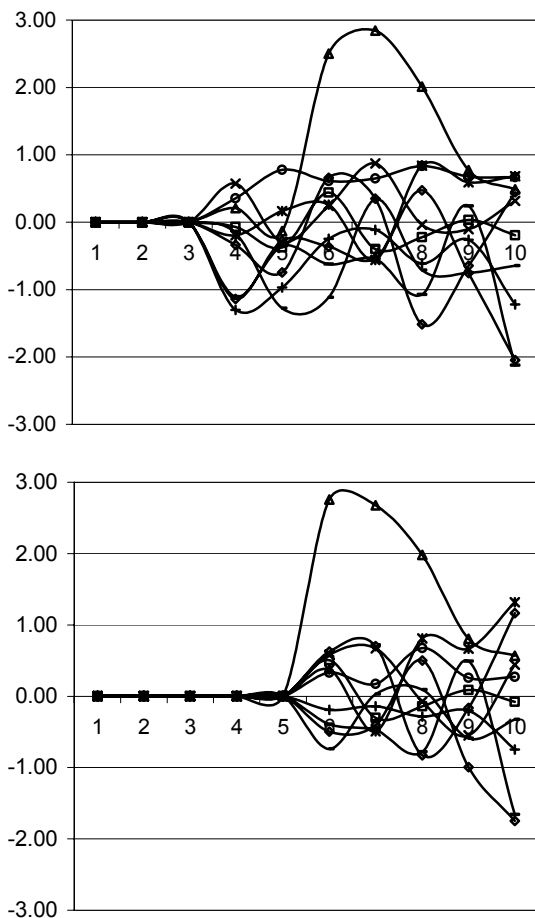


Fig.5 – Inter-class recognition. Personal synthetic discriminant filters are applied to test objects of different classes. To be compared with Fig.4.

5. CONCLUSION

We have described the algorithm for face identification with maintaining high level intra-class recognition rather than inter-class one. The recognition algorithm first uses an area of interest localization procedure to provide rough face regions under different poses, followed by a procedure which minimizes illumination effects.

Synthetic discriminant function approach in the stage of classification can potentially deliver the very low False Acceptance Rate (FAR) for intra-class problem if the power of training set is enough. In our research use of 5-member training set instead of 3-member one improves the performance of the system significantly.

In order to utilize SDF technique the 2,5-dimensional sparse mesh for face image representation was offered. Some obvious advantages of proposed method for data coding are the following. At first, it enables to vary the dimensionality of the problem and to fit it to obtain non-zero solution of matrix equation (4). In other words, we can guarantee correct solution until the number of training samples is less than the dimensionality of vectors included in matrix W . Furthermore, in pure geometric sense the 2,5-D face model enables to perform (pseudo)rotations of the face not only in the image plane, but also in depth.

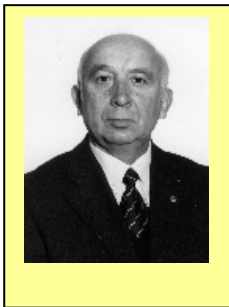
The question on the efficiency of proposed approach for inter-class recognition tasks remains open. The problem seems to be considered under some additional constraints while the discriminant filters construction is in progress.

6. REFERENCES

- [1] R. Chellappa, C.L. Wilson, S. Sirohey. Human and machine recognition of faces: A survey, *Proc. IEEE*, 5 (83) (1995), pp. 705-740.
- [2] W.Y. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips, "Face recognition: A literature survey", *UMD CfAR Technical Report CAR-TR-948*, 2000
- [3] P. J. Phillips, H. Moon, D.M. Blackburn, E. Tabassi, and J.M. Bone. The FERET evaluation methodology for face recognition algorithms, *IEEE Trans. on PAMI*, 22 (10) (2000), pp.1090-1104.
- [4] P. J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, and J.M. Bone. *FRVT 2002: Evaluation Report*, March 2003.
- [5] T. Papatheodorou, D. Rueckert. Evaluation of automatic 4D face recognition using surface and texture registration. *Proc. of the Sixth IEEE International Conference on Face and gesture Recognition, FG'2004*, Seoul, Korea 17-19 May 2004, pp.321-326.
- [6] M. Turk, A. Pentland. Eigenfaces for recognition, *Jour. Cognitive Neuroscience*, 1 (3) 1991.
- [7] P. Belhumeur, J. Hespanha, D. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Trans. on Pattern Recognition and Machine Intelligence*, 7 (19) 1997, pp. 711-720.
- [8] C. Liu, H. Wechsler. Comparative assessment of independent component analysis, *Proc. the 2nd*

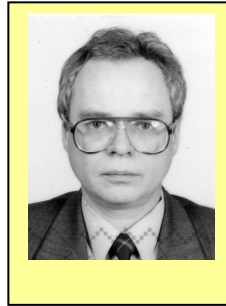
Int. Conf. Audio- and Video-based Biometric Person Authentication, Washington D.C., 22-24 March 1999.

- [9] D. Dai, P. Yuen. Regularized discriminant analysis and its application to face recognition, *Pattern Recognition*, (36) (2003), pp. 845-847.
- [10] K. Etemad, R. Chellappa. Discriminant analysis for recognition of human face images, *J. Opt. Soc. Am. A*, (14) (1997), pp. 1724-1733.
- [11] M. Skurichina, R. Duin. Bagging, boosting and the random subspace method for linear classifiers, *Pattern Analysis and Applications*, 2 (5) (2002), pp. 121-135.
- [12] L. Breiman. Bagging predictors, *Machine Learning*, 24 (2) (1996), pp. 123-140.
- [13] T. Ho. The random subspace method for constructing decision forests, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8 (20) (1998), pp. 832-844.
- [14] D. Casasent. Unified synthetic discriminant function computational formulation, *Appl. Opt.*, 23 (10) (1984), pp. 1620-1627.



Rauf Kh. Sadykhov received his BS in Computer engineering in 1967 from Azerbaijan Technical University, Baku and PhD in Computer Science in 1975 from Institute of Engineering Cybernetics, Belarus, Minsk. He is currently the head of

Computer Systems Department in the Belarusian State University of Informatics and Radioelectronics, professor in computer Science. He was the Chairman of International Conferences "Pattern Recognition and Information Processing" (1999, 2005), "Neural Networks and Artificial Intelligence" (2001, 2003). His research interests include digital signal and image processing, pattern recognition and computer vision systems, parallel computer systems architectures, artificial intelligence and multiagent systems. He got his D.Sci. (engineering) degree in 1991.



Vladimir A. Samokhval graduated from the Physics Department of the Byelorussian State University in 1982. He became software developer for specialized scientific equipment, i.e. Texture Analyzing System. Providing software support for quantitative microscopy he permanently entered into the field of image analysis and pattern recognition. In 1998 he received PhD degree for handwritten character recognition (symbols and signatures). Now he is senior researcher of the Laboratory of system identification in the United Institute of Informatics Problems (UIIP), National Academy of Sciences of Belarus. His research interests lies in pattern recognition and image processing for intelligent system development in computer vision.