A GEOMETRIC APPROACH TO FACE RECOGNITION

V. Starovoitov and D. Samal

Institute of Engineering Cybernetics,

Surganov Street 6, 220012 Minsk, BELARUS

E-mail: {ValeryS,samal}@newman.bas-net.by

ABSTRACT

We present our study of the face recognition problem. Due to difference in human pose, face expression, hairstyle, image style and lighting conditions, the problem is very difficult. To solve it we have to test different image processing tools and heuristics for robust recognition. The main features of our approach are detection of fiducial points, calculation of geometric features and application of nonlinear image dissimilarity function at the final recognition stage. We demonstrate the power of the approach by experiments.

1. INTRODUCTION

Since the beginning of the 1990s, the race recognition problem has become a major issue in computer vision, mainly due to the important real world applications of face recognition: smart surveillance, secure access, telecommunications, digital libraries, medicine and so on.

On the theoretical side, face recognition is a specific and hard case of object recognition. Faces are very specific objects whose most common appearance (frontal faces) roughly looks alike. Subtle changes make the faces different. Therefore, in a traditional feature space, frontal faces will form a dense cluster, and standard pattern recognition techniques will generally fail to discriminate between them [1].

There are two main types of the face recognition systems. The first one is to check if a person outstanding in front of a camera is a member of a restricted group of people (*20-500* persons) or not. Usually, such systems are used to access control to buildings, computers, etc. The peculiarities of such systems are real time of reaction and small sensitivity to the checking person position and appearance changing.

Systems of the second type identify a person by photo searching in a large database or confirm its absence. Such a system must work with a database containing *1.000-1.000.000* images. It may work in off-line manner. We try to design a system of the second type.

1.1 Problem Formulation

There is a database of N portrait images and a query image. Find k images most similar to the given face image. The number k may be a constant (for example, *20* for a large database), it may be limited by a similarity threshold, or it may be equal to the number of all pictures of the same person in the database.

1.2 Input Data Limitation

For robust work of the system, images must satisfy the following conditions:

- They are gray-scale or color digital photos.
- The head size in the input image must be bigger than *60x80* pixels; otherwise the fiducial points may be detected with low accuracy.
- Intensity and contrast of the input image allow to detect manually the main anthropometrical points like eye corners, nostrils, lip contour points, etc.
- The head must be rotated not more than at *15-20* degrees (with respect to a frontal face position).

Ideally, the input image is a digitized photo for a document (passport, driving license, etc.).

1.3 Our Approach

There are different approaches to create such systems: eigenface analysis [2], template matching [3], graph matching, fiducial point based approach [4] and others [5]. In most of them faces are considered as flat surfaces and the difference in orientation of the compared faces are ignored. Actually, a face is a 3D convex object with ability to rotation and shape changing. The main difficulty in the face recognition is to find a robust feature set for a unique description of a human face.

In our study we explore geometric features like distances between some facial points. In order to improve the recognition rate we try to find a feature set to be steady to changing of the shooting conditions and the recognizing person face and to minimize it as much as possible. Face similarity is evaluated in several steps, selecting faces "from coarse to fine". First we detect a set of fiducial points in every face, and then approximately k_1 images close to the query face are selected with respect to some geometric features based upon the points, where k1 may be closed to:

$$k_1 = (\ln N)^2 \quad .$$

After that we rotate and scale images, then evaluate correlation between central namely face parts of the k_1 images and the query image. In the end k_2 images from the set of k_1 are selected, where:

$$\ln k_1 \le k_2 \le (\ln k_1)^2 \,.$$

The last set of images has to contain at least one image of the queried person, if any.

2. GEOMETRIC FEATURES

2.1 Fiducial Points

As a basis of our geometric approach we used the set of fiducial points (also they are called as anthropometrical). In experiments we checked *37* such points. Some of them are detected automatically; the others are extracted manually. We are working towards automatic point detection. The used approach to facial keypoints extraction is described in [6]. After the point detection their coordinates may be corrected by operator (manually) to improve their location. Coordinates of the points may be stored with the corresponding image in a database. An example of the explored face keypoints is presented in **Figure 1**, right picture.

2.2 Feature Choosing

Geometric features may be presented by segments, perimeters and areas of some figures formed by the detected points. We studied different subsets of the features looking for the most robust features but due to the great quantity of them we cannot report about strong results of the searching yet. Hence, to present our progress in comparison with known recognition results we tested the feature set described in [7]. It includes *15* segments between the points and the mean values of *15* symmetrical segment pairs. The explored feature set is not the best; the details of its comparison are described in [8].

2.3 The Feature Set Optimization

Once a feature set has been obtained it can be optimized by the presenting technique. How to choose the optimal feature subset? The point was to find the feature space with the maximal distances between the clusters and minimal ones between the patterns of one cluster. In our case all database images of the same person were considered as one cluster. To evaluate the effectiveness of every feature subset the F value was calculated:

$$F = \sqrt{\frac{\sum_{i=1}^{k}\left(M_{D_i} - D_i\right)^2}{\sum_{i=1}^{k}\left(M_{M_i} - M_i\right)^2}} \,,$$

where M_i and D_i are mean and variance of the feature values for k images of the i-th person, $M_{D\,i}$ and $M_{M\,i}$ are mean of D_i and M_i, respectively. The lowest F value corresponds to the better feature set. To validate this technique the distances between the clusters were computed every time of the feature space changing. The experimental results justify our strategy [8].

With the help of this estimation we have selected *28* features from 30 described in [7]. The recognition rate using the optimized feature set was improved.

2.4 Face Recognition Based on The Features

The feature values are stored together with a person identification photographs in a database. When the tested image normalized on the rotation, scale and intensity level the fiducial points are detected and the values of the feature are calculated. All images stored in database are the patterns in the feature space. To find the closest to the tested image we have to evaluate the Euclidean distances from it to all others.

3. FINE RECOGNITION STAGE

In [8] we have introduced a new idea for a gray scale image comparison. Our approach follows the hypothesis according to which if two digital images display the same but slightly changed scene, the images have to be similar locally and have to contain a lot of close pixels with similar gray values.

We consider a gray scale image A as a digital surface, i.e. $A=\{(i,j,a_{ij})\},\ 0 \le i,j \le N_0$.

We proceed from the fact if two pixels belong to different images and display the same element of a real scene, they must be close enough in 3D intensity-spatial space. To evaluate the proximity, we may calculate distance between the pixels by a simple metric like the city-block, the chess-board or any other one. The distance from every pixel of one image to the nearest (in 3D space) pixel of the other image reflects so-called local image dissimilarity. We calculate a set of local dissimilarity values and accumulate them into a global dissimilarity measure, which gives the final similarity evaluation. We calculated global dissimilarity between images A and C by the function

$$D(A,C) = \sqrt{\frac{1}{2ZN}\{\sum_{i,j}^{N}[d(a_{ij},C)+d(c_{ij},A)]^2\}^{1/2}} \,,$$

where

$$d(a_{ij},C) = d(a_{ij},C_w) = \min_{(l,m)\in w}\{d(A_{ij},C_{lm})\}$$

and W means a square observation window of size $(2w+1)\times(2w+1)$ centered at position (i,j) of the image C. In experiments we used w=4 to compare facial templates of size 66×56, like in **Figure 1**. Similarly we calculated $d(c_{ij}, A)$. The basic distance function d may be of any metric. We applied the chess - board metric:

$$d(a_{ij}, c_{lm}) = max\{\ |i-l|,\ |j-m|,\ |Ga_{ij} - Gc_{lm}|\ \},$$

where i,j and l,m are the spatial coordinates of pixels a and c, relatively, and Ga_{ij} and Gc_{lm} are the values of intensity of the pixels.

To reduce the computational cost of the algorithm we used a spiral principle for calculation of $d(a_{ij},C)$. Value $d(a_{ij},c_{ij})$ is determined in the first turn. We suppose that if the images compared are quite similar and the pixel c_{ij} is not equal to the pixel a_{ij}, then a pixel c_{lm} with similar or the same gray level must be close to the pixel c_{ij}. Therefore we check pixels contained in the window W in turn gradually going away from the center of W

located at the position *(i,j)*. The algorithm stops if it discovers a pair of pixels a_{ij} and c_{lm} with the smallest possible distance value between them; otherwise it computes the distance value from the pixel a_{ij} to every pixel c_{lm} in the window *W* and chooses the minimal value.

4. EXPERIMENTS

4.1 The Experimental Database

Our image data set was based on the Olivetti Research Laboratory face database included 400 face images of 40 persons. The portraits were made without sufficient difference in shooting time but with small rotation, orientation and illumination variances. Each photo was an image of 92×112 pixels, quantized to 256 gray levels (**Figure 1**).

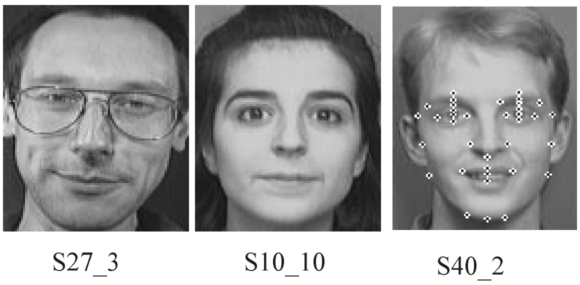

Figure 1. Example of ORL face database and facial keypoints (at the last image) used in the experiments.

The image name-code S*p_t* means the *t*-th image of the *p*-th person. The experimental data set contained 70 images of 12 persons. Two people were presented by 10 photos each, the others – by 5 images. There were 19 images of persons wearing the glasses at the moment of shooting.

To prepare the image database for the fine stage of our recognition procedure we had to make templates of the face central part for every image (see **Figure 4**). In order to normalize the templates for comparison, the images were rotated to made irises been on the horizontal line. The rotation did not have to change the intensity values of the pixels at all. Pursuing this aim the rotation algorithm described in [9] was chosen. It is based on the sequentially shifting of the rows and columns of pixels in the horizontal and vertical directions, respectively. After the rotation, the central parts of the face images were cut from the rotated images. The size of each template was 55*66 pixels. The eyes horizontal positions were fixed relatively to the template borders.

4.2 The Experimental Procedure

There were two aims of the experiments: to test the effectiveness of our estimation of the steady feature set selection and to validate the dissimilarity function *D* applying to face recognition. To find the appropriate quantity of the nearest images k_1 to the query one, Euclidean distances in the feature space from a template image to every image in the database were calculated. In our case (k_1=*6*), the *6* closest images were derived. If there was any photo of the query person, then the result was considered successful. Each image was tested as a query and compared to others. Using the feature set described in [7], we achieved the recognition rate *67/70*. The steps of our study are illustrated by the examples of the nearest images retrieved for *S27_3*. The original image is presented in **Figure 1** on the left. The results of the first experiment are shown in **Figure 2**.

Note we had five pictures of every person in the test data set. So as one of them was the query image then there were maximum *4* images of the same person in the k_1 set. One can see (in **Figure 2**) the 6 closest images do not display the tested person *S27* and the result of the experiment cannot be considered as

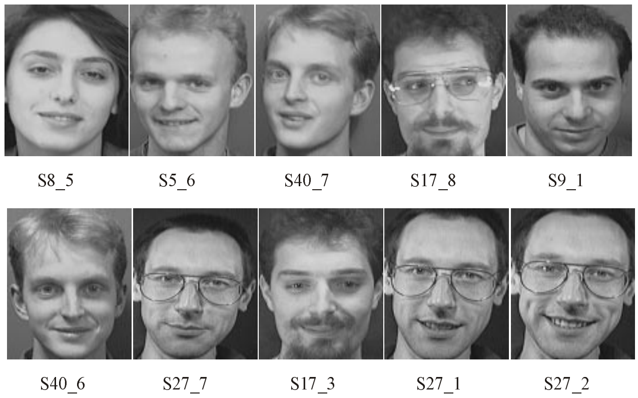

Figure 2. From left top to right bottom: The 10 nearest to S27_3 images using the feature set from [7].

positive. Recognition rate for all attempts was *67/70*, i.e. in *67* cases of *70* possible ones the set of *6* closest images included the photo of wanted person. After the feature set selection and reducing to *28*, the recognition rate improved to *69/70*. It means just in one case of *70* tests there were not any images of the query person through the *6* nearest ones (the worst face was *S10_10*, see **Figure 1**), i.e. the recognition rate was *98.5%.*. Note the best rate in study [7] applied to *100* images from the same database was *89,36 %.* The results of *S27_3* identification using selected features are presented in **Figure 3**. To estimate the effectiveness

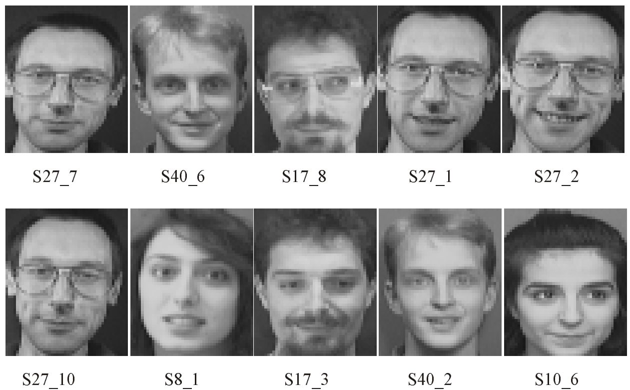

Figure 3. (From left top to right bottom) 10 closest to images S27_3 after the feature set selection.

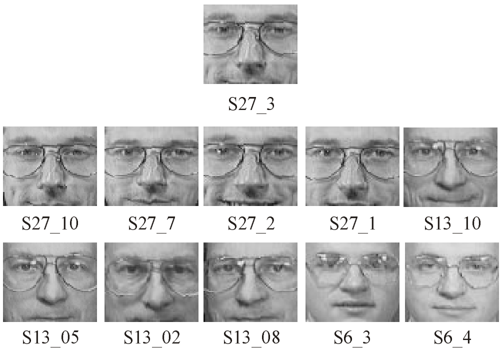

Figure 4. A face template (56×66 pixels) and 10 closest to it rotated subimages acordinally the measure *D*

of dissimilarity measure *D* application to face recognition we compared all *70* faces to each other and the result for *S27_3* is shown in **Figure 4**. All photos of the queried person are on the first positions in the closest image row. The correlation measure is not so exact. The other interesting point is the fact of all pictures in **Figure 4** are the photos of men wearing the glasses.

The other approaches of template matching tested by us (Hausdorff, MSE) did not give such good results. But the time of one image pair comparison when applying the dissimilarity measure D is is about for 500 times longer than the geometric feature based on approach used. That was the reason to combine these two methods for robust and quick recognition system design.

Utilizing the combination of the fast but coarse and the fine but slow approaches we achieved the quiet good and steady result. The example of it is presented in **Figure 5**.

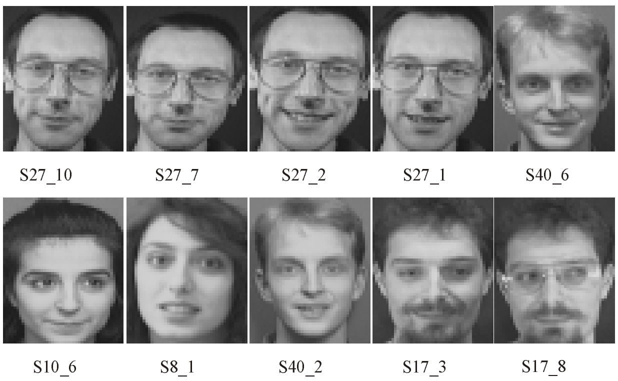

Figure 5. 10 closest images to the query image S27_3 selected by the function D applied to the set of images depicted in **Figure 3**.

5. CONCLUSION

The paper presents the results of the face recognition study based on geometric features and template matching. The features based on the facial key points exhibited good recognition rate. The experimental results demonstrate the geometric features based on approach can be effectively explored for a coarse preliminary face recognition stage. Then for final recognition more precise techniques may be applied as a new low-level approach to grey-scale image similarity evaluation. The experiments with about *100* real images demonstrated good sensitivity of the presented measure *D* to the face photos to the same belong cluster belonging.

6. REFERENCES

[1] "Face Recognition. From Theory to Application", *Proceedings of the NATO Advanced Study Institute,* vol. 163, 637 pages, 1998.

[2] Chellapa P., Wilson C., and Sirohey S., "Human and Machine Recognition of Faces: A Survey," *Proceedings of IEEE,* vol. 83, no. 5, pages 705-740, 1995.

[3] Brunelli R., and Poggio T. "Face Recognition: Features versus Templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pages 1042-1052, 1993.

[4] McKenna S.J., Gong S., *et al*. "Tracking Facial Feature Points with Gabor Wavelets and Shape Models," *Lecture Notes in Computer Science,* Springer Verlag, 1997.

[5] Samal D. and Starovoitov V. "Approaches and methods to face recognition. A survey," *Instute of Engineering Cybernetics, Preprint #8*, Minsk, 54 pages, 1998 (in Russian).

[6] Starovoitov V., Samal D., G. Votsis, and S. Kollias "Face recognition by geometric features", *Proceedings of 5-th Pattern Recognition and Information Analysis Conference*, Minsk, May 1999.

[7] Abay E., Akarun L., and Alpaydyn E., "A Comparative Analysis of Different Feature Sets for Face Recognition," *Proceedings of ISCIS*, Antalya, pages 568-576, 1997.

[8] Samal D. and V. Starovoitov, "Features for recognition choosing based on statistical data," *Digital Processing of Images, Proceedings of IEC*, Minsk, 1999 (in Russian).

[9] Di Gesu V. and Starovoitov V. "Distance-based functions for image comparison," *Pattern Recognition Letters*, Vol. 20, No. 2, pages 207-214, 1999.

[10] Owen C.B. and Makedon F. "Bottleneck-free separable affine image warping," *Proceedings of ICIP*, Vol. 1, pages 683-686, 1997.