FOSFI: A System for Face Image Recognition[1, 2]

M. V. Petrushan, A. I. Samarin, and D. G. Shaposhnikov

Kogan Research Institute for Neurocybernetics, Rostov State University,
pr. Stachki 194/1, Rostov-on-Don, 344090 Russia
e-mail: nisms@krinc.ru

Abstract—FOSFI (Foveal System for Face Identification), a hardware/software system for authorized access by face personal identification, is presented based on biologically plausible algorithms of image description and recognition developed earlier. These algorithms perform detection of the most informative image regions, spatially nonuniform representation of visual information, and context encoding of primary features. The testing of the system in actual practice demonstrates that the system does not permit illegal access in 100% of cases, while the probability of false denials of an authorized access is less than 8%.

INTRODUCTION

Presently, in the field of biometric identification of a person, an approach is being intensively developed based on algorithms and systems for face recognition [5]. These systems are oriented toward two main application tasks. The first one is related to the problems of an authorized access when a first-type recognition error (falsely positive identification) should be at least 10^4 times less than a second-type recognition error (falsely negative identification). The second task is related to a search for a certain person among others when the relation of error types should be opposite. As a rule, different algorithms are used for these two tasks because, in the first case, a certain similarity measure is searched for while, in the second case, a measure of dissimilarity should be determined.

The main concern of the present research was the problem of face recognition in application to the task of an authorized access.

The algorithmic basis for analysis and representation of visual features in the present research is the methods and models of active visual perception developed earlier at the Institute for Neurocybernetics at Rostov State University. In these models, a visual foveal sensor with spatially nonuniform resolution is used [1–3]. These algorithms are based on biologically plausible principles aimed at detection of the most informative image regions, spatially nonuniform representation of visual information, and context encoding of primary features [3, 4]. The basic features used in all the designed algorithms are a module and a direction of brightness gradient [1, 3].

[1] The work was supported by the grant of the Institute for Neurocybernetics of Rostov State University "Development of the user software package of the algorithms of face image identification for an authorized access."

[2] The article was translated by the authors.

Received October 25, 2004

GENERAL SCHEME OF THE SYSTEM

To solve a problem of an authorized access to a restricted area, a hardware/software system was developed. The system includes a subsystem of video-input, a subsystem of image processing and formation of template image descriptions, and a subsystem of archive storage and personal data input. The storage of a database of facial images and their descriptions, formation of additional descriptions, and person identification were implemented by means of a high-performance computer (server) in a local network.

The system has three functional modules: Administrator, Security Department, and Identification Sector.

Administrator has complete access to system tuning and image-description forming for identification of objects by manual and/or automatic selection of foveal sensor positioning in a face image. The Administrator module can also form additional descriptions of a given person in the case of falsely negative identification; the feature that marks a new image is then stored.

While adding a new person to the database, the measure of similarity of the person's description with the other subjects in the database is estimated (checking for "twins"). If the measure exceeds an allowable limit (0.75), descriptions of "twin" are added to the database by marking new sensor-positioning points in the "twins'" images until they are definitely distinguished.

Security Department forms personal information in the database and assigns a personal identification code to each subject. A set of template images of this subject in the database corresponds to this code. This user has access to the system protocol, including information about operation of the access system (who came through or tried to and when) and the formation of individual messages to persons passing an access zone.

Identification Sector performs visual control of the system operation in the access zone and allows a person to come into the protected area in the case of a false denial. In this case, current and template images of the person are visualized on the monitor. If an operator rec-

Pattern Recognition and Image Analysis, Vol. 15, No. 2, 2005, pp. 425–427.

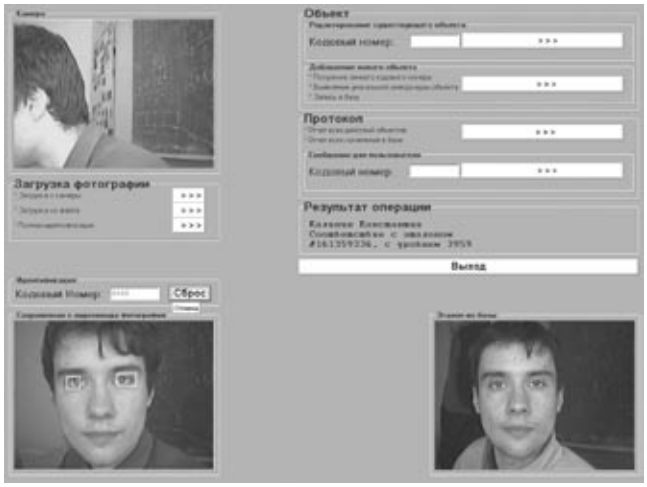

Fig. 1. Operation of the Identification Sector subsystem. The screen is divided into four regions: the top left region is the online video-input window, the top right region is a service region available for the Administrator only, and the bottom left and right regions show the current and template images, respectively.

ognizes this person, he/she sends the current image to the file of possible complementary images.

SUBSYSTEM OF IMAGE PROCESSING AND RECOGNITION

The subsystem of image processing and recognition of the FOSFI includes two basis modules: Face Feature Detection (FFD) and Face Description and Recognition (FDR), implemented as dynamically linked libraries of functions (.dll). Both modules operate with images from 256×256 to 512×512 pixels in size. A facial region is no less than 25% of the image.

In FDD, the cascade method for detecting the most informative regions of a face image (eyes, nose, and mouth) has been realized [2]. The method is based on computationally simple procedures with oriented brightness edges. At a noise level up to 20%, the cascade method provides detection of at least two informative fragments (eye regions) with accuracy sufficient for calculating the rest of the anthropometrical points. The method provides coordinates of the detected fragment centers and their sizes. The time of detection is 100 ms on a Pentium-4 with 256 Mb RAM.

FDD also provides the marking of basic anthropometrical points used for criminalistic facial identification [6].

FDR is used to specifically describe a 2D image by a set of 49-dimensional vectors [1, 3] while a foveal sensor is positioned in the anthropometrical points. This module provides image encoding invariant to scaling and projective transformations and to the changing of a viewing point (within the limits of $\pm 25°$) by forming complementary image descriptions. The processing time of forming the image feature description is 0.25 s. A specific set of 49-dimensional template vectors is formed in the database as a result of operation of the module in the mode of template image description. In the mode of recognition, the module forms a presented image feature description and calculates a coefficient of similarity for the current description and the templates from the database.

The algorithms for detecting oriented edges, used both for detection of informative regions and for specific description of an image, make it possible to parallelize computations. Moreover, since the FOSFI system is aimed at estimating the authenticity of a given person, the processing time depends only on the number of template descriptions of this person in the database.

TESTING OF THE SYSTEM

The testing of the hardware/software system, including the FOSFI, were carried out at the Institute for Neurocybernetics at Rostov State University. For one month, when coming to the institute, employees entered an individual identification code received from the experimental database. As soon as the first digit of the code was entered, the person's face image was taken by a video camera and sent to the server. The server provided the image processing and comparison of its feature vectors with the template vectors from the database corresponding to the individual code. If the coefficient of similarity exceeded an experimentally determined value of 0.75 (the correct recognition), the server sent a signal about recognizing the entering person.

Figure 1 shows an example of operation of the subsystem Identification Sector in the case of correct recognition. It can be seen that the entered image differs from the template one in size. As the testing showed, the system provided a stable recognition if the size of a facial image varied in a range from 0.8 to 1.3 with respect to the size of a template image from the database.

If the similarity coefficient is less than 0.75 (the system did not recognize the current image), the template image corresponding to the personal identification code from the database is visualized along with the current image and a message about the event is stored. A user of the subsystem Identification Sector made a decision, by visual comparison of the two images, about whether the person should be allowed to enter the institute or not. At the end of a working day, the Administrator module analyzed all the events of non-recognition and, if necessary, added a complementary description of a nonrecognized person into the database.

Inserting a new image into the database was carried out as follows. The Administrator chose images to be included in the database from the nonrecognized images and manually marked up two basic points (centers of eye pupils), the remaining anthropometrical points were marked automatically. If their location did

Fig. 2. Interface of the Administrator subsystem in the mode of marking the basic anthropometrical points (black circles).

not correspond to the real anthropometrical points, it was manually corrected. As testing revealed, the automatic algorithm of marking did not require corrections in 60% of cases.

Figure 2 shows an example of the Administrator subsystem functioning in the mode of marking new images. The interface of this subsystem allows a user to observe both the image as a whole with all the anthropometrical points and the context region of a point that should be marked. Then, the foveal sensor is positioned in each of these points and feature vectors are formed and stored in the database.

One month of system testing yielded the following results. For a database of template descriptions of 50 persons and for the 150 persons tested, including those from the database, the system provided 100% identification of "outsiders" and gave a false alarm while admitting a person with an authorized access in 8% of cases.

It was also determined that the minimally sufficient number of templates in the database was from 2 to 5 descriptions per person. The average time of image processing and recognition, taking into account information exchange with the server, did not exceed 1 s, which corresponded to the time of walking from the video camera to the door.

CONCLUSIONS

The testing demonstrates that the developed system of personal identification by facial image analysis provides a high degree of protection from illegal access. The system can plausibly be adapted to variations of the individual behavior of a person in the access zone and accumulates minimally sufficient descriptions in 15–20 entries. In comparison with the known systems (see, for example, www.neurotechnologija.com) the developed system has a series of advantages. In particular, it does not require an exact positioning of a person's face during video-input for reliable recognition.

The algorithms and methods that underlie the recognition system demonstrate an improvement in recognition quality in comparison with the ORL image database (www.orl.co.uk/facedatabase.html) that was used in the course of development of the original algorithms. This is due to the increase in image size and improvement of image resolution.

In the future, we plan to expand the FOSFI possibilities to solving problems related not only to determination of a person's authenticity in the access zone but also to recognition of a person in the case of their free behavior.

REFERENCES

1. Y. K. Gavrilei, A. I. Samarin, and M. A. Shevchenko, "Active Image Analysis in Systems with Foveal Perception," Neurocomputers: Design and Application **7–8**, 34–46 (2002).
2. A. Golovan', N. Shevtsova, L. Podladchikova, S. Markin, and D. Shaposhnikov, "Detection of Face Informative Regions by Using Local Features," Neurocomputers: Design and Application **1**, 50–57 (2001).
3. D. Shaposhnikov, A. Golovan', L. Podladchikova, N. Shevtsova, X. Gao, V. Gusakova, and Yu. Gizatdinova, "Application of the Behavioral Model of Vision for Invariant Recognition of Facial and Traffic-Sign Images," Neurocomputers: Design and Application **7–8**, 21–33 (2002).
4. E. L. Schwartz, D. N. Greve, and G. Bonmassar, "Space-Variant Active Vision: Definition, Overview, and Examples," Neural Networks **8** (7/8), 1297–1308 (1995).
5. H. Wechsler, *Face Recognition: From Theory to Applications* (Springer, 2002).
6. A. M. Zimin and L. Z. Kirsanova, *Criminalistic Photography Facial Examination (Manual)* (VNKC Press, Moscow, 1991) [in Russian].