

UDC 004.9

DOI: 10.15827/2311-6749.20.3.3

## The review of human identification techniques by facial image

A.Yu. Serov <sup>1</sup>, Master of Science, [nippongaijin88@gmail.com](mailto:nippongaijin88@gmail.com)

A.V. Kataev <sup>1</sup>, Ph.D. (Engineering), Associate Professor, [alexander.kataev@gmail.com](mailto:alexander.kataev@gmail.com)

<sup>1</sup> Volgograd State Technical University, Computer-Aided Design (CAD) Department, Volgograd, 400005, Russian Federation

This paper discusses modern methods and technologies for identification techniques by facial image. All known approaches are based on the selection and analysis of face parameters in the image and their further processing. Processing the obtained face parameters is mainly based on neural network and mathematical approaches. The disadvantages of the neural network approach are mathematical problems (retraining, choosing the optimization step, getting into a local optimum).

The problems of the mathematical approach are low speed and low resistance to the image defects, while the neural network method can correct image defects at the stage of image alignment. Also, the mathematical method is very demanding in terms of computing resources.

During the work on the paper, the method of active form models and the neural network method of identifying a person by the image of a person were selected and tested on two data samples. The active shape model method is used to detect facial features and obtain key points on faces. The neural network method using a convolutional neural network retrieves a descriptor that describes a person, which is a vector of 128 features. Further, by determining the distance between the vectors, the most similar vector is located in the database.

During testing, the speed of the method and the accuracy of the work were measured. The test results showed a performance of approximately 2 secs in two samples and an accuracy of 97 %. These studies are related to the development and implementation of a module for streaming identification of people by video stream, where the reaction speed of the method is very important, as well as its accuracy.

**Keywords:** *identification, person identification, person identification by face, face recognition, computer vision.*

The identification and recognizing systems are very widespread, from metro security systems to mobile devices [1]. The initial research in this area began in the 50s of the twentieth century. The development of electronics and computer technology has determined the topicality for the problem of human identification by facial images. Today, some identification systems allow recognizing a person with an accuracy of 97 %, which makes it possible to use these systems in various fields.

At the moment, there are several approaches to solve the identification problem [2]. These solutions are based on various branches of mathematics: neural network models [3], graph theory [4], statistical methods [5]. In this paper, there are four methods for the study: the flexible comparison method on graphs, neural networks, the principal component method, and an active appearance model.

The paper presents the performance test results of the neural network method for identifying a person by face image.

### Human identification techniques and methods by facial image

Figure 1 presents the general structure of the human identification techniques by facial image. The stages of face sensing and equalization are generally the same for widely used methods. The method divergence mainly lies in the stages of feature extraction and comparison, as well as in organizing the database of face features.

Let's make a detailed overview of identification techniques.

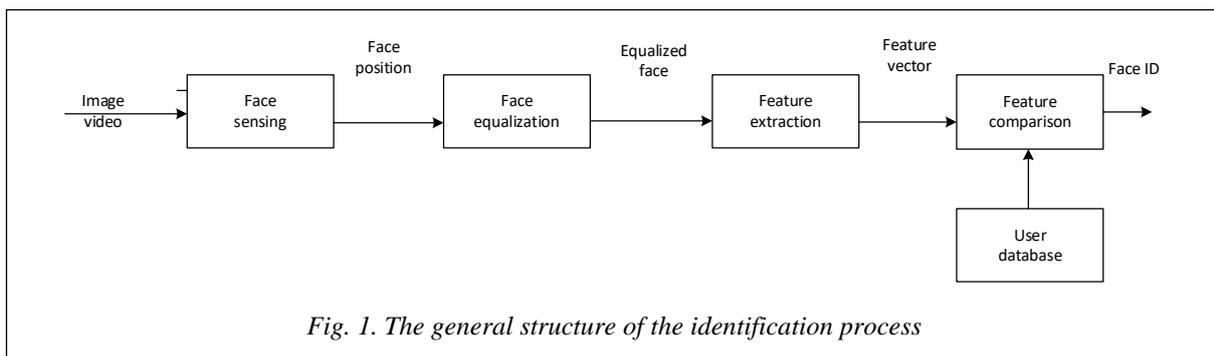
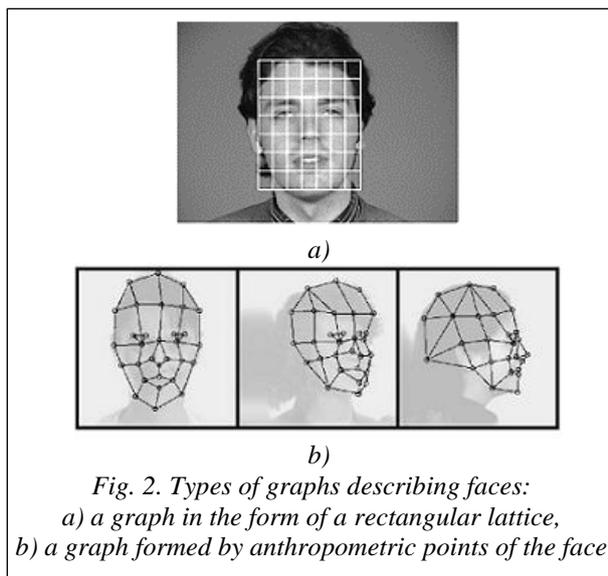


Fig. 1. The general structure of the identification process

### The flexible comparison method on graphs

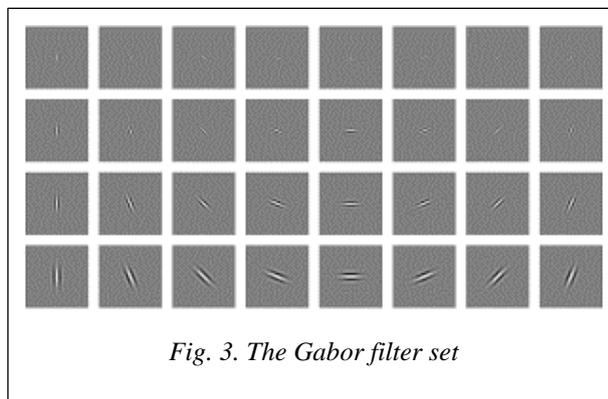


This method uses an algorithm for elastic matching of face graphs. Faces are presented as weighted graphs (Fig. 2). Graphs can have the form of a rectangular lattice or the form of a structure, the vertices of which are anthropometric points of the face. In the course of identification, two graphs are used: the reference graph, which remains unchanged, and the graph, which is adjusted to the reference graph using deformation [6].

The feature values, in this case, are calculated using Gabor filters [7], or Gabor wavelets: sets of Gabor filters that are calculated at the top of the graph by convolution of pixel values with Gabor filters. Figure 3 presents an example of a Gabor filter set.

This method is resistant to the image angle, with high computational power and direct dependence of the operating time from the number of faces in the database, which makes this method unsuitable for systems where a fast response is required.

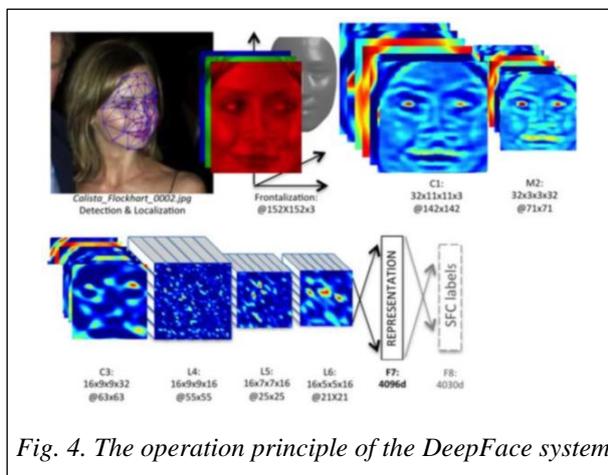
### The identification with neural networks



Currently, there are a large number of neural network varieties. The most common method is a neural network built using a multilayer perceptron. This network allows you to classify the incoming image according to the preliminary network training. The neural network training is based on using a set of training examples. This training is come down to solving an optimization problem using the gradient descent method and adjusting the weights of interneural connections. During the neural network training key features are obtained, their importance is determined and connections are built between them.

At the moment the convolutional neural networks have the lead in the person identification problem by using the face image. They are the development of the ideas of such architectures as the cognitron and neo-cognitron. The advantage of a convolutional neural network over a conventional one is the ability to take into account the two-dimensionality of the image. Also, in consequence of the two-dimensional connectivity of neurons, common weights (allows you to detect some features at any place in the image), and spatial sampling, the convolutional neural network is partially resistant to image scaling, displacements, and other contortions.

A convolutional neural network is used in the DeepFace system acquired by the social network Facebook to identify the faces of its users. Figure 4 shows how it works.



The deficiencies of convolutional neural networks are mathematical problems (retraining, choosing an optimization step, getting into a local optimum), as well as difficulties in choosing a network architecture.

### The principal component method

This method is based on Karunen-Loeve's transformation. It was designed from the start in statistics to decrease the feature space without significant information loss. In the identification problem, it is used to represent a face image by a small-sized vector (a principal components' vector).

The method objective is to significantly decrease the feature space in order to achieve a better description of images belonging to multiple faces. This method identifies miscellaneous changes in the training sample of faces and describes them in the basis of orthogonal (proper) vectors. Training proceeds in the following way: a training sample is created  $\Omega = \{I_1, \dots, I_n\}$ , where  $I$  – it is a vector obtained from the image brightness matrix. Next, a matrix is formed from the received data, in which the row corresponds to the training image, and the data is centered and normalized for each column of this matrix. The next step is to calculate the eigenvalues and eigenvectors. Their eigenvectors are chosen  $k$  - vectors corresponding to the first  $k$  largest eigenvalues. The value of  $k$  is selected by the formula

$$\frac{\sum_{i=0}^m k_i = 0\lambda_i}{\sum_{i=0}^m \lambda_i} > 0.95$$

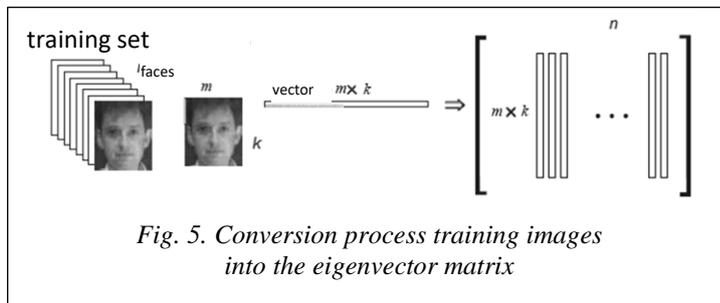


Fig. 5. Conversion process training images into the eigenvector matrix

A matrix is constructed from the obtained eigenvectors. The vectors in it are arranged in decreasing order of eigenvalues. Each image of the input set is projected by multiplying the original vector by the eigenvector matrix. Figure 5 presents the transformation of the training sample to the eigenvector matrix. The resulting vectors are used for identification. The set of eigenvectors obtained once is used to describe all other face images. Using eigen-

vectors, you can derive an archived approximation to the input image, which can be stored as a coefficient vector for subsequent search in the database. During the identification process, a brightness matrix is built for the input image, and then it is converted into a vector. The received vector is multiplied by the matrix of eigenvectors. Next, the Euclidean distance is calculated between the stored vectors and the received ones. The result will be the element to which the Euclidean distance is minimal.

The advantages of this identification method are its simplicity of implementation, the simplicity of expanding the number of reference faces, and low memory consumption for storing objects. The disadvantage is the high sensitivity to input data. On the change in lighting, head angle, and emotional expression, the algorithm significantly loses its effectiveness. It is related to the fact that the initial problem of the algorithm is to approximate the input data, not to classify.

### The active appearance model

This model is a statistical model that can be adjusted to fit the real image by deformation. Initially, active appearance models were used as a method for assessing the face parameters.

The active appearance model contains the following parameters: shape parameters and appearance parameters. The model is trained on a set of manually labeled images. Labels are numbered. The label defines the key point that the model will look for during adaptation to the image. Figure 6 shows an example of a 68-label markup.

Before starting training the appearance model, the shapes are normalized to compensate for differences in scale, tilt, and displacement. Normalization is performed using orthogonal Procrustean analysis. Figure 7 shows an example of face shape points before and after normalization.

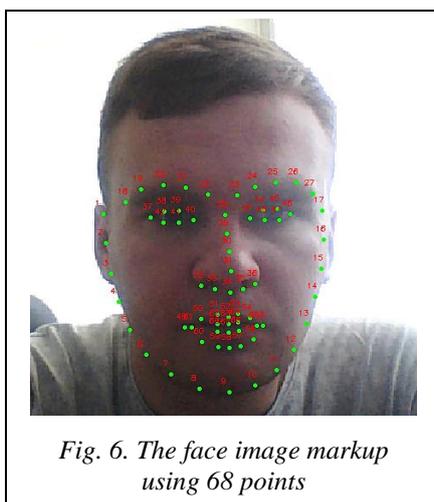


Fig. 6. The face image markup using 68 points

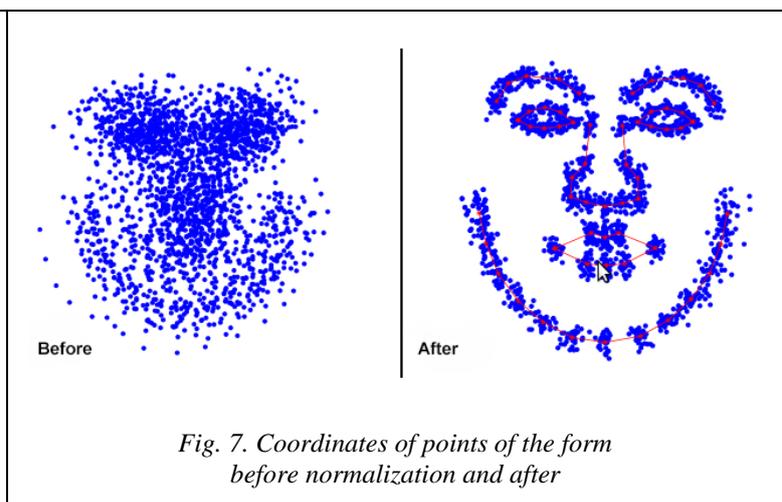


Fig. 7. Coordinates of points of the form before normalization and after

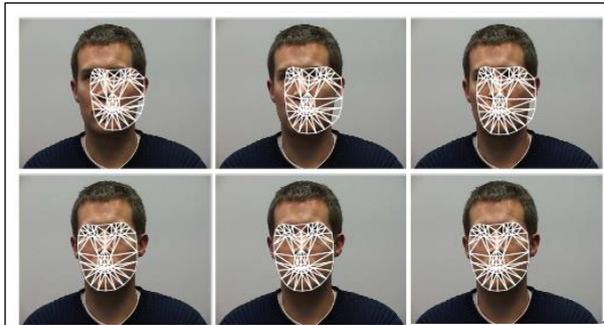


Fig. 8. The model fitting process

Further, principal components are extracted from the normalized coordinates using the principal component method. A matrix of texture pixel values is formed from the pixels within the triangular regions formed by the shape points. After finding the principal components of this matrix, the active appearance model is considered trained.

The model is fitted to the real face image in the process of solving the optimization problem, which minimizes the functionality by the method of gradient descent. Figure 8 represents the model fitting process.

The purpose of active appearance models is not identification, but a preliminary exact search for anthropometric points of the face for future processing.

### The active shape model

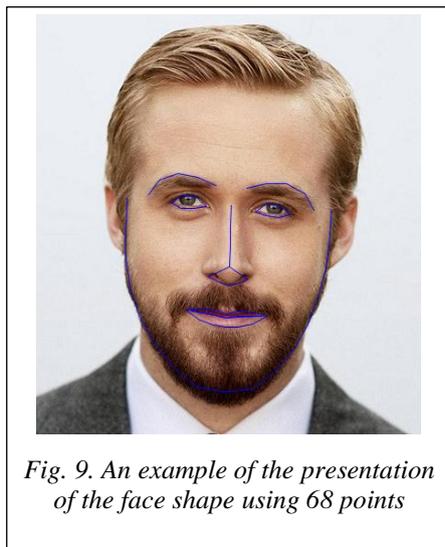


Fig. 9. An example of the presentation of the face shape using 68 points

The problem of the method of active shape models is to take into account the statistical relationships between the anthropometric point spacing on the face. Anthropometric points are manually marked on the training sample of faces shot in full face. In each image, the points are numbered in the same order. Figure 9 depicts an example of a face shape representation using 68 points.

To bring the coordinates on all images to a single system, a generalized Procrustean analysis is carried out, which centralizes all points and brings them to the same scale. Next, the mean shape and the covariance matrix are calculated. Based on the covariance matrix, eigenvectors are calculated and sorted in descending order. The active appearance model is defined by the covariance matrix and the mean shape vector. Localization of the model on a new face image that is not included in the sample occurs in the process of solving the optimization problem. Figure 10 shows an example of the localization process. In most algorithms, a required step before classification is the face image equation to the frontal position relative to the camera or bringing a sample of faces to a single coordinate system. For the phase implementation, it is necessary to localize anthropometric points characteristic of all faces on the image - in most cases, these are the pupil centers or the eye corners.

For the purpose of cost reduction, experts allocate no more than 10 such points. The main goal of an active shape model is not identification, but preparation for it.

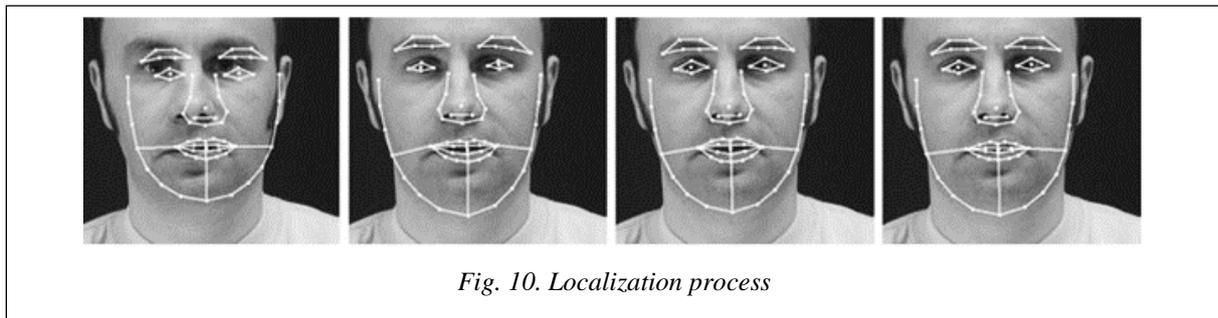


Fig. 10. Localization process

### Testing the performance of the identification algorithm

This testing was carried out on two samples of images: Georgia Tech face database and Faces95 Collection of facial images. The test aim was to obtain the average time required for the process of identifying a person from the face image. The method of active shape models and the ResNet34 neural network trained to extract descriptors participated in the testing. Testing was performed on the following PC configuration:

- CPU: Intel i5-4210U, 2.4 GHz;
- GPU: Nvidia Geforce GT840M;

- RAM: 6 GB;
- HDD: Western Digital WD Blue Mobile 1 TB 5400 rpm.

The Georgia Tech face database is a database of human faces containing images of 50 different people, grouped into groups of 15 images with different head tilt, facial expressions, lighting conditions, and scales. On average, a person's face in this sample occupies 8% of the entire image area. Images are taken in the frontal projection of the head. It contains both male and female faces. Figure 11 presents an example of a facial image from this sample.

The Faces95 Collection of facial images is a database of facial images of 72 different people, grouped into groups of 20 images. In a group, images differ from each other in head tilt, lighting, facial expression, and scale. Each image has a resolution of 180x200 pixels. There are both male and female faces. All images are taken in the frontal projection of the head. Figure 12 presents an example of a human face.



Fig. 11. An example of a face image from the Georgia Tech face database



Fig. 12. An example of a face image from a sample of Faces95 Collection of facial images

As follows from the testing on the Georgia Tech face database sample, an average identification time is 1.25 seconds. The percentage of correctly identified persons is 100.

As follows from the testing on a sample of Faces95 Collection of facial images, an average identification time is 1.3 seconds. The percentage of correctly identified faces is 97.8 - it decreased due to too poor illumination in some test images and the low resolution of the sample itself.

### Conclusion

This paper analyzes the main methods for identifying a person by face image used in modern solutions. One identification method was tested using the ResNet34 neural network and the active model method. As the result, this method seems to be possible to use in practice due to its accuracy and operation speed. The study was carried out as part of the development of a module for identifying a re-approached person using a face image, and the selected identification method will be used in its implementation.

*Acknowledgements:* This paper was financially supported by RFBR, project no. 18-07-01308.

### References

1. Yurko I.V., Aldobaeva V.N. Yurko I.V., Aldobaeva V.N. Application areas and operating principles of face recognition and identification systems based on video recording in real time. *International Student Scientific Herald*, 2018, no. 2. Available at: <http://www.eduherald.ru/ru/article/view?id=18416> (accessed August 07, 2020) (in Russ.). DOI: 10.17513/msnv.18416.
2. Sanjeev K., Harpreet K. Face recognition techniques: Classification and comparisons. *IJITKM*, 2012, vol. 5, no. 2, pp. 361–363.
3. Fedyukov M.A. *Algorithms for Constructing a Human Head Model from Images for Virtual Reality Systems*. Moscow, 2016, 22 p. (in Russ.).
4. Kisku D.R., Gupta P., Sing J.K. Applications of graph theory in face biometrics. *CCIS, Proc. BAIP*, 2010, vol. 70, pp. 28–33. DOI: 10.1007/978-3-642-12214-9\_6.
5. Delac K., Grgic M., Grgic S. Statistics in face recognition: Analyzing probability distributions of PCA, ICA and LDA Performance Results. *Proc. ISPA*, 2005. DOI: 10.1109/ISPA.2005.195425.
6. Adeykin S.A. Overview of face recognition methods from 2D images. *Youth Scientific and Technical Bulletin*, 2013, no. 12. Available at: <http://ainsnt.ru/doc/640908.html> (accessed August 15, 2020) (in Russ.).

7. Lavrova E.A. Study of methods for forming facial image features using Gabor filters. *Youth Scientific and Technical Bulletin*, 2014, no. 9. Available at: <http://ainsnt.ru/doc/735601.html> (accessed August 15, 2020) (in Russ.).

8. Arsentev D.A., Biryukova T.S. Method of flexible comparison in graphs as algorithm for pattern recognition. *Vestnl. MGUP*, 2015, no. 6, pp. 74–75 (in Russ.).

УДК 004.9

DOI: 10.15827/2311-6749.20.3.3

### Обзор методов идентификации человека по изображению лица

*А.Ю. Серов*<sup>1</sup>, магистр, [nippongaijin88@gmail.com](mailto:nippongaijin88@gmail.com)

*А.В. Катаев*<sup>1</sup>, к.т.н., доцент, [alexander.kataev@gmail.com](mailto:alexander.kataev@gmail.com)

<sup>1</sup> *Волгоградский государственный технический университет, кафедра «Системы автоматизированного проектирования и поискового конструирования», г. Волгоград, 400005, Россия*

В данной статье рассмотрены современные методы и технологии идентификации человека по изображению лица. Все известные подходы основаны на выделении и анализе параметров лица на изображении и на дальнейшей их обработке. Обработка полученных параметров лица в основном базируется на нейросетевом и математическом подходах. Недостатками нейросетевого подхода являются математические проблемы (переобучение, выбор шага оптимизации, попадание в локальный оптимум).

Проблемы математического подхода заключаются в низком быстродействии и слабой устойчивости к дефектам изображения, в то время как нейросетевой метод способен исправить дефекты изображения на этапе выравнивания изображения. Также математический метод очень требователен к вычислительным ресурсам.

В ходе исследования был выбран и протестирован на двух выборках данных метод активных моделей формы и нейросетевой метод идентификации человека по изображению лица. Метод активных моделей формы используется для детектирования черт лица и получения ключевых точек на лица. Нейросетевой метод с использованием сверточной нейронной сети извлекает дескриптор, описывающий лицо, представляющий собой вектор из 128 признаков. Далее путем определения расстояния между векторами в базе находится самый похожий вектор.

В процессе тестирования измерялось быстродействие метода и точность работы. Результаты тестирования показали быстродействие примерно 2 секунды на двух выборках и точность в пределах 97 %. Данные исследования связаны с разработкой и реализацией модуля для потоковой идентификации людей по видеопотоку, где очень важна скорость реакции метода, а также его точность.

**Ключевые слова:** *идентификация, идентификация человека, идентификация человека по лицу, распознавание лиц, компьютерное зрение.*

**Благодарности.** Работа выполнена при финансовой поддержке РФФИ, проект № 18-07-01308.

### Литература

1. Юрко И.В., Алдобаева В.Н. Области применения и принципы работы систем распознавания и идентификации лиц по видеофиксации в реальном времени // *Междунар. студ. науч. вестн.* 2018. № 2. URL: <http://www.eduherald.ru/ru/article/view?id=18416> (дата обращения: 07.09.2020). DOI: 10.17513/msnv.18416.

2. Sanjeev K., Harpreet K. Face recognition techniques: Classification and comparisons. *IJTKM*, 2012, vol. 5, no. 2, pp. 361–363.

3. Федюков М.А. Алгоритмы построения модели головы человека по изображениям для систем виртуальной реальности: автореф. дис... канд. наук. М.: ИПМ им. М.В. Келдыша РАН., 2016. 22 с.

4. Kisku D.R., Gupta P., Sing J.K. Applications of graph theory in face biometrics. *CCIS, Proc. BAIP*, 2010, vol. 70, pp. 28–33. DOI: 10.1007/978-3-642-12214-9\_6.

5. Delac K., Grgic M., Grgic S. Statistics in face recognition: Analyzing probability distributions of PCA, ICA and LDA Performance Results. Proc. ISPA, 2005. DOI: 10.1109/ISPA.2005.195425.
6. Адейкин С.А. Обзор методов распознавания лиц по 2D-изображениям // Молодежный научно-технический вестник. 2013. № 12. URL: <http://ainsnt.ru/doc/640908.html> (дата обращения: 15.08.2020).
7. Лаврова Е.А. Исследование методов формирования признаков изображения лица с помощью фильтров Габора // Молодеж. науч.-технич. вестн. 2014. № 9. URL: <http://ainsnt.ru/doc/735601.html> (дата обращения: 15.08.20).
8. Арсентьев Д.А., Бирюкова Т.С. Метод гибкого сравнения на графах как алгоритм распознавания образов // Вестн. МГУП им. Ивана Федорова. 2015. № 6. С. 74–75.