Face recognition methods

A I Dzhangarov¹, M A Suleymanova² and A L Zolkin³

¹ Department of Programming and Information and Communication, Chechen State University, Sheripova Street 32, Groznyy 364093, Russia
² Department of Business Informatics, Chechen State University, Sheripova Street 32, Groznyy 364093, Russia
³ Department of Water Transport, Volga State University of Water Transport (Samara branch), Molodogvardeiskaya Street, 62-64, Samara 443099, Russia

Abstract. Nowadays, a science as cybernetics is gaining popularity. World society is on the verge of digitalization a technological and social sectors of human life. One of the most popular areas of cybernetics is computer vision. Systems of video assistants that recognize images and objects are widely distributed, and social test and control systems are introduced in test mode. In this regard, the relevance of research in this area is extremely high. This paper provides an analysis of the available methods for recognizing faces in images and discusses the software implementation of one of them.

1. Introduction
Face recognition is a component of a higher level task - identification of objects, their details and signs. A similar methodology has already been widely adopted in various structural accounting systems for legal entities, video recording systems and access control, in institutions such as airports and metro. Also in intelligent human-computer interfaces; in photographic equipment for automatic focusing on a person’s face, as well as for image stabilization of faces in order to facilitate recognition of emotions; to expand the stereo vision zone when creating 3D display systems.

There are four methods of face detection:

- empirical method;
- scanning unique characteristics;
- recognition using patterns defined by the developer;
- method which uses an external signs of learning system.

2. Empirical method
This approach involves the use of an algorithm that implements several rules, which, in turn, must be answered by the analyzed fragment of the image. These rules are a formalization of empirical knowledge about how a person’s face should look in the images and what the user should be guided in when deciding whether it is a human face or not.

The basic rules are as follows:

- the brightness of the face (of the central part) should be uniform;
• the face should have symmetrically placed eyes, nose and mouth, the brightness of which should differ significantly from other parts of the face;
• the brightness between the central and upper parts of the fragment should seriously differ [1].

![Figure 1. Brightness zones.](image)

Also, this approach involves a significant reduction in the input image, to smooth out interference and optimize computational operations. As a result, in such an image it is much easier to recognize the zone of uniform distribution of brightness and then check the presence of zones that differ significantly in brightness (Figure 1). Observance of all the described rules and properties makes it possible with different probability to judge whether the object in question is a human face.

However, it is worth noting that this method has lost its widespread use with the advent of more powerful computers. Despite the fact that the empirical approach gave impulse to the development of new and much more advanced recognition algorithms, it did not produce the most accurate results and was extremely sensitive to tilting and turning the head, not to mention complex scenes (the presence of a large number of foreign objects in the background).

3. Scanning unique characteristics

In fact, this method was designed to solve the problems that arose in an empirical study. It took a different approach to the problem - there was no attempt to explicitly formalize all the processes that logically arise in the human mind. On the contrary, an attempt was made to establish patterns and properties of fragments of faces implicitly, to find unique characteristics of the face that would ignoring the angle of inclination of head and its position.

The main algorithm can be divided into the following stages:

• identification on the image fragment of obvious features of the face;
• detection of the border of the face, its shape, brightness, texture and color;
• the union of all detected invariant features and their verification.

The approach of recognition faces in complex scenes involves deep search for the correct geometric locations of facial features. For this, a Gaussian derivative filter which contain a lot of scales and orientations is used. After that, a search is made for the correspondence of the revealed features and their relative position by random sorting [2].

![Figure 2. Geometric locations of facial features.](image)

The essence of the feature grouping approach is to use the second derivative of a Gaussian filter to search for areas of interest in the image fragments (Figure 2). The edges are then grouped around each such area using a threshold filter. Also the following estimate is used to help Bayesian network to produce a selection of facial features.
Similar approach have become widespread and used, as they allow one to detect faces in various positions. However, it also has a number of significant drawbacks. With a slight clutter of the face with various objects, the appearance of noise or light, the percentage of reliable recognition drops significantly. Also plays a significant role the presence of a complex background image.

The basis of the considered approaches - empiricism, is both their strengths and weaknesses. Recognition difficulty the dependence of the image fragment: on the shooting and lighting make it possible to attribute face identification in the image to high-complexity tasks without hesitation. And since it is not easy to effectively transfer informal human experience and knowledge into a set of formal rules, serious difficulties arise in these methods[3].

4. Recognition using templates defined by the developer
The use of templates establishes a standard understanding of the image of the face, by describing its individual areas and their probable relative position. Face identification using a template consists in sequentially checking all areas of the image for compliance with a given template. Uniqueness of this approach can be considered the presence patterns at the implementation stages (non-deformable and deformable). These patterns are pre-programmed and untrained - a comparative correlation is used to find the face in the image. This approach of identification faces using three-dimensional forms involves the use of a template in the form of pairs of brightness ratios in two areas. To detect the location of the face, it is necessary to analyze the whole image with a given template. Moreover, it is necessary to do this with a different scale (Figure 3) [4].

![Figure 3. Template example.](image)

The control points distribution models are statistical and reviewed objects whose shape can change. In addition, this method is the able to distinguish the form of variable objects within the training set with a small number of form parameters (Figure 4). This compact and precise parameterization can be used to develop effective classification systems.

![Figure 4. Face shape fixation.](image)

The main advantages of this type of recognition include simplicity of implementation and good results in images with a not very complex background. Moreover, the main drawback is the need to configure the template close to the face. The great complexity of calculating templates for various angles and face rotations calls into question the appropriateness of their use [5].

5. Face identification approach by external signs
When using such approaches, the image is assigned in a certain way a calculated feature vector, which is used to classify images into two classes - face / not face. Very often, approaches based on the use of a mathematical model of images of a human face are used in research of this type, consists in a complete search of all rectangular fragments of the image and checking each of the fragments for the presence of
a face. Since the search scheme has such obvious disadvantages as redundancy and high computational complexity, approaches are used to reduce the number of fragments under consideration. The main features:

- **scholasticism**: the image is scanned by rectangular areas and represented by feature vectors;
- **block structure**: the image is divided into intersecting or disjoint sections (Figure 5) of various scales and an assessment is made using algorithms for estimating the weights of vectors [6].

![Figure 5. Splitting face areas.](image)

For training algorithms, a library of manually prepared special images (which have fragments of faces and another things), any other images, is required. Therefore, the most important part of this approach – certain classifiers to search for signs in the image is also extremely urgent.

The main methodology for introducing classifiers is often the so-called neural network. Today, the creation of a neural network is the easiest way to solve the problems of facial recognition in images. This network is a developed mathematical model that is a system of connected and interacting components called neurons. Neural networks need to be trained. Technically, training consists in finding the coefficients of connections between neurons.

There is a method of support vectors, which is used to reduce the dimension of the feature space, without leading to a significant loss of information content of the training set of objects. Application of the principal component method to a set of linear space vectors allows one to go over to such a basis of space that the main dispersion of the set will be directed along the first few axes of the basis, called the principal axes. The subspace spanned by the main axes thus obtained is optimal among all spaces in the sense that it best describes the training set. This set is formed by capturing the key points of multiple images. The more images the network receives at the input, and the greater the variability of capture and recognition of key features, the higher the likelihood of face identification [7].

### 6. Software implementation

As part of the research, it was decided to use one of the most accurate tools in face recognition with the FaceNet system, which is implemented in the Python programming language. FaceNet is a neural network based on the so-called triple loss function. This function minimizes the difference between the input image and images that contain similar features. And vice versa - maximizes the difference if there are few common signs.

For testing was chosen the Python programming language, which is the most commonly used when it comes to computer vision and setting up self-learning networks. The fact is that convenient language tools allowed creating extremely useful and convenient libraries for realizing recognition tasks. The tests used TensorFflow libraries (used for deep machine learning) and Keras [8]. At the first stage, the FaceNet network was assembled and for this the following libraries and tools were additionally imported:

- glob;
- numpy;
- cv2;
- tensorflow;
- keras;
- fr_utils.
The initialization of the network began with determining input image sizes:

```python
FIDmodel = faceIdentification(in_size=(3, 150, 150))
```

here, 

\[(3, 150, 150)\]

indicates that the image is transmitted through three RGB channels and has a size of 150x150 pixels.

Next, was determined the triple loss function:

```python
def triplel(q_true, q_same, q_other, size = 0.3):  
    point, true, false = q_pr[0], q_pr[1], q_pr[2]  
    min_distance = calc_sum(tensf.square(tensf.subtract(point, true)), axis=x=1)  
    max_distance = calc_sum(tensf.square(tensf.subtract(point, false)), axis=y=1)  
    triple = tensf.add(tensf.subtract(min_distance, max_distance), size)  
    loss = tensf.calc_sum(tensf.maximum(triple, 0.0))  
    return loss
```

**Figure 6.** Triple loss function.

The definition of a function is described in the TensorFlow framework. After determining the loss function, it is possible to compile the entire system.

After the initialization of the system, the final stage takes place – preparing the database for training the network. As already mentioned, FaceNet minimizes the difference between similar images and increases it between different ones. In the above implementation, this information is used to identify the presence of a human face in the image. Images from the database are processed directly in the faces method:

```python
def recognition(img, base, fragment):  
    encoding = img_to_encoding(img, fragment)  
    pos_distance = 90  
    id = None  
    for (face, t_encode) in base.items():  
        distance = np.linalg.norm(t_encode - encoding)  
    if pos_distance < neg_distance:  
        pos_distance = distance  
        id = face  
    if neg_distance > 0.63:  
        return None  
    else:  
        return id
```

**Figure 7.** Face recognition function.

Then, after the function processes the images using FaceNet and returns the encoded image, it becomes possible to compare it with the images of the faces of people who are in the database and make an assumption about the likely identity of the person:

```python
if pos_distance < neg_distance:  
    pos_distance = distance  
    id = face  
if neg_distance > 0.63:  
    return None  
else:  
    return id
```

**Figure 8.** Face identification.
About three hundred images (150 faces and 150 “not faces”) were entered into the created database for tests. After training the network, the number 0.63 was obtained, which means the percentage of confidence in the identification of a person. This number may vary depending on the data provided and the settings used [9].

7. Conclusion
As can be seen from the review of existing methods, the direction of computer vision, and specifically the tasks of identifying objects, is constantly evolving and improving. More precise and more productive methods, libraries, frameworks are appearing. The process of global digitalization, which today affects more and more different states, once again emphasizes the importance and relevance of research in this area. Directions related primarily to the analysis of big data. Based on the current trend, we can easily assume that the tasks of building neural networks will soon take one of the leading roles in many areas of our lives. That is why it is so important to study the available tools and approaches to their implementation [10].

References
[1] Viola P and Jones M J 2004 Robust real-time face detection International Journal of Computer Vision 57(2)
[2] Lanitis A, Taylor C J, Ahmed T and Cootes T F 1995 Wolfson Image Anal Classifying variable objects using a flexible shape model Image Processing and its Applications
[3] Annalyn Ng and Kenneth Soo 2017 Numsense! Data Science for the Layman: No Math Added Kindle
[4] Cathy O’Neil and Rachel Schutt 2013 Doing Data Science
[5] Sameer Dhanrajani 2017 AI and Analytics: Accelerating Business Decisions Kindle
[6] Simon S Haykin 2010 Neural Networks and Learning Machines (PHIL: Third Edition edition)
[7] Dzhangarov A I and Akhmetova H A 2010 Inženernyj vestnik Dona 1 Retrieved from http://www.ivdon.ru/ru/magazine/archive/n4y2019/5877
[8] Müller A C and Guido S 2018 Introduction to Machine Learning with Python: A Guide for Data Scientists Kindle
[9] Jake VanderPlas 2016 Python Data Science Handbook Jupyter
[10] Christopher M 2006 Bishop Pattern Recognition and Machine Learning Information Science and Statistics
[11] Varaev Z M and Magomedov I A 2019 Prediction of diabetes using the artificial neural network approach The development of legal science and the problem of bridging the gaps in law pp 1-3
[12] Mentsiev A U 2019 J. Phys.: Conf. Ser. 1399 044042