FACE RECOGNITION USING LOCAL QUANTIZED PATTERNS AND GABOR FILTERS

V. Khryashchev [a], A. Priorov [a], O. Stepanova [a], A. Nikitin [b]

[a] P.G. Demidov Yaroslavl State University, Yaroslavl, Russia – (vhr, andcat, olga1stepanova)@yandex.ru
[b] PicLab LLC, Yaroslavl, Russia – nikitin@piclab.ru

Commission VI, WG VI/4

KEY WORDS: Face Recognition, Eye Center Localization, Local Quantized Patterns, Gabor Filters, Histogram Comparing

ABSTRACT:

The problem of face recognition in a natural or artificial environment has received a great deal of researchers' attention over the last few years. A lot of methods for accurate face recognition have been proposed. Nevertheless, these methods often fail to accurately recognize the person in difficult scenarios, e.g. low resolution, low contrast, pose variations, etc. We therefore propose an approach for accurate and robust face recognition by using local quantized patterns and Gabor filters. The estimation of the eye centers is used as a preprocessing stage. The evaluation of our algorithm on different samples from a standardized FERET database shows that our method is invariant to the general variations of lighting, expression, occlusion and aging. The proposed approach allows about 20% correct recognition accuracy increase compared with the known face recognition algorithms from the OpenCV library. The additional use of Gabor filters can significantly improve the robustness to changes in lighting conditions.

1. INTRODUCTION

The problem of face recognition in a natural or artificial environment has received a great deal of researchers' attention over the last few years. Despite of many efforts devoted to this problem, we have to admit that it is still far from being resolved (Vizilter et. al., 2010).

Being one of biometric-based technologies face recognition are less reliable than other alternative biometric systems of identification. However, face recognition systems are widely used in different applications such as parental control system, interface interaction of humans and robots, sorting of photos and video (Szeliski, 2010).

Different methods are used for person identification using facial image. For example, multilayer neural network shows good results when the number of people is fixed while the task of searching of a particular person in a crowded place requires more complex methods including usage of different feature sets, classifiers, etc.

Despite the variety of different methods of face recognition, a typical algorithm consists of three main components (Vizilter et. al., 2010; Zhao et. al., 2003): (i) projection of face data to a lower dimensional representation; (ii) feature extraction; (iii) classification stage: neural network, metrics, cluster analysis, etc. In addition, a priori information (facial features characteristics) is used for algorithm constructing and experimental information is used for developing and correcting of algorithm.

Different approaches are used for face recognition in computer vision systems: using key facial points and measuring facial characteristics, exploiting geometrical information (Brunelli et. al., 1992), principal component analysis (Turk et. al., 1991), linear discriminant analysis (Belhumeur et. al., 1997), Gabor wavelets based methods (Wiskott et. al., 1997), algorithms using discrete cosine transform (Messer et. al., 2006).

In the recent years a number of approaches based on local binary patterns (LBP) has gained increasing attention from researches in the field of image processing (Ojala et. al., 1994; Ahonen et. al., 2004; Huang et. al., 2011; Petruk et. al., 2011).

The basic idea of the local binary patterns is to avoid image representation as a high-dimensional vector which contains a lot of redundant information, and describe the local structure of an image. Extracted features will have lower dimension.

The original LBP was introduced by T. Ojala in 1994. This theoretically very simple yet efficient multiresolution operator is defined for each pixel by thresholding the 3×3 neighborhood pixel value with the center pixel value. In this way, it can give us a binary sequence defined by local structure of an image.

LBP is computationally efficient because it works with integer arithmetic (it can achieve real-time performance in some tasks), and it is invariant to changes in the brightness of the image caused by shooting in different lighting conditions.

2. FACE RECOGNITION ALGORITHM

In this paper novel face recognition algorithm based on modified LBP is presented. The following steps are offered to apply this algorithm to the problem of face recognition (Fig. 1):

2.1 Normalization of face image using eyes coordinates

The detected face is rotated, scaled and cropped according to the known coordinates of eyes (determined using eye localization algorithm). The chosen resolution of the face image is 90x150 pixels, coordinates of the left eye center are (75, 55), coordinates of the right eye center are (15, 55).
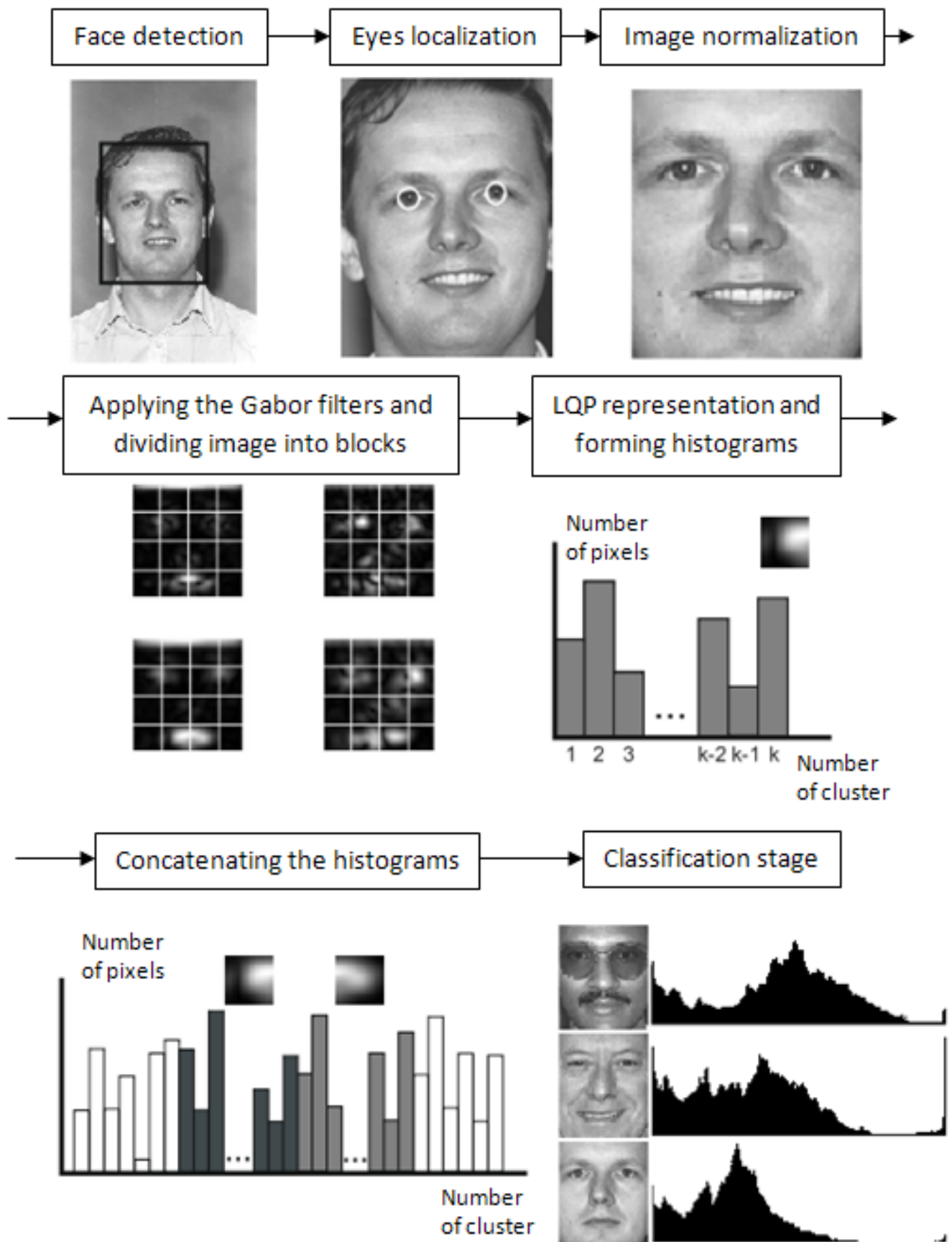
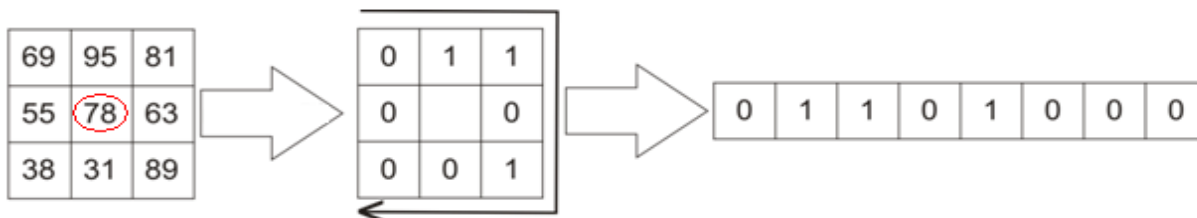Figure 1. Block diagram of face recognition algorithm

Figure 2. Example of counting of LBP

2.2. Gabor filters applying

40 Gabor filters which are determined by (1) are applied to normalized image:

$$\psi_{u,v} = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{\left(-\frac{\|k_{u,v}\|^2\|z\|^2}{2\sigma^2}\right)} \left[e^{i\vec{k}_{u,v}z} - e^{-\frac{\sigma^2}{2}}\right], \quad (1)$$

where
$k_{u,v} = k_v e^{i\phi_u}$, $k_v = \frac{k_{max}}{f^v}$ = the frequency
$\phi_u = \frac{u\pi}{8}$, $\phi_u \in (0,\pi]$ = the orientation
$z = (x,y)$.

It should be noted that variable u in (1) determines the scale of Gabor filters; v controls the Gabor filters orientation. The following parameters are chosen for Gabor filters: $\sigma = 2\pi$, $k_{max} = \frac{\pi}{2}$, $f = \sqrt{2}$, five scales $v \in \{0,1,2,3,4\}$ and eight orientations $v \in \{0,1,2,3,4,5,6,7\}$. These Gabor kernels form a bank of 40 different filters and exhibit desirable characteristics of spatial frequency, special locality, and orientation selectivity (Shan et. al., 2004).

2.3. Dividing image into a number of blocks

Each image is divided into p non-overlapping rectangle blocks each of size $L \times L$ pixels. In this paper $p = 9 \times 15 = 135$.

2.4. Constructing a dictionary for each block

Describing the local structure of an image by LBP operator we avoid image representation as a high-dimensional vector which contains a lot of redundant information. Extracted features have lower dimension.

LBP operator is defined for each pixel by thresholding the $L \times L$ neighborhood pixel value with the center pixel value. An output value of the LBP operator can be obtained as follows (Ojala et. al., 1994):

$$LBP = \sum_{p=0}^{P-1} 2^p s(i_p - i_c), \quad (2)$$

where
i_c = the average intensity of the center pixel

i_p = the intensity of its neighborhood pixel p

s = sigma function:

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}.$$

Basic LBP thresholds eight neighborhood pixel values with the center pixel value (see Fig. 2 for an example). In this way, it can give us a binary sequence defined by local structure of an image (Ahonen et. al., 2004).

Despite their encoding efficiency and simplicity, local binary pattern features remain very local and hence somewhat myopic, where increasing the neighborhood size by including more pixels or increasing the circle diameter increases the histogram size (number of codes) exponentially, e.g. increasing the number of surrounding pixels from 8 to 16 in LBP increases the number of codes from 256 to 65,536. Furthermore, these local patterns use hard-wired codings and fixed layouts, which are most probably not well adapted to the underlying dataset and/or the application domain. These shortcomings limit the encoding capacity of local patterns and prevent them from harnessing all the information available locally.

In this paper perspective approach described in (Hussain, Triggs, 2012; Hussain, Napoléon, Jurie, 2012) is offered to solve many of the above mentioned problems of local
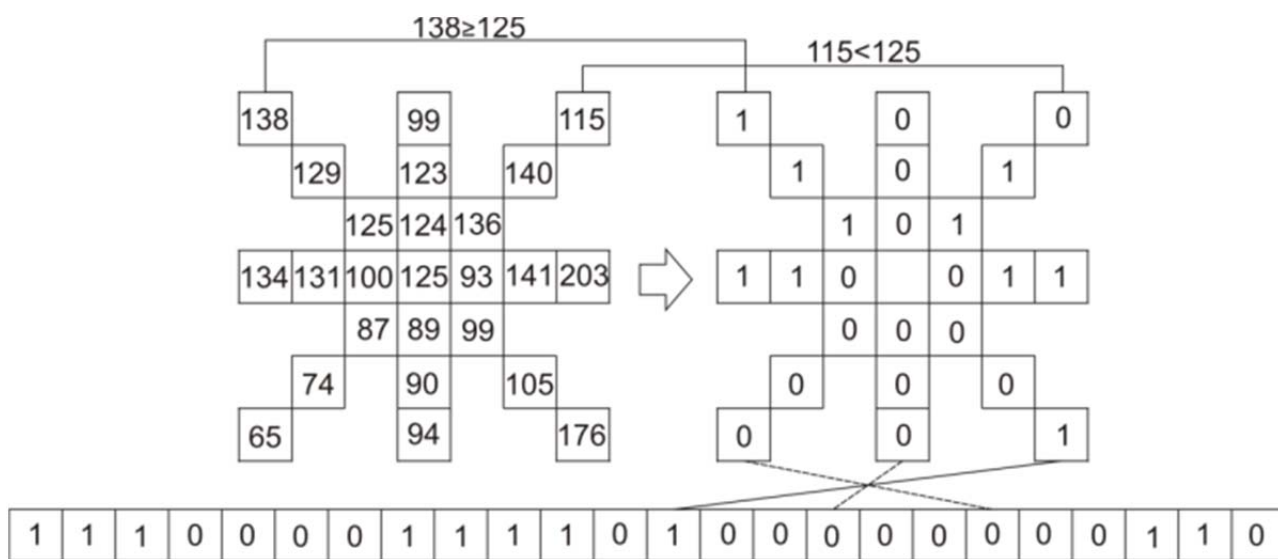

Figure 3. Processing of extended 24-bit local quantized pattern

patterns. To maintain the speed and simplicity of local pattern features and to make the process of vector quantization fast, local quantized patterns (LQP) are used. The basic idea of LQP is to extend processing region and to divide distinct codes into groups by clustering (Mandel, 1988). Groups of codes form new feature set of lower dimension.

Many different neighborhood geometries are possible for LQP and one of its main advantages is its flexibility in this respect. In this paper 24 pixel local pattern neighborhood is used. After standard thresholding this neighborhood with the center pixel value we get 24-bit binary code for each pixel of image (see Fig. 3).

These 24-bit codes form huge array of binary sequences. For example, 655,360 codes are formed from training database that consists of 10 images with resolution 256x256 pixels. To divide these codes into groups we should count the number of occurrences of each code over suitable image regions and take into account only codes that occur more than N times. The resulting codes can be easily stored in a table. LQP builds these tables by mapping all the codes to the nearest cluster centers using K-Means, and thus guarantees fast vector quantization with no extra overhead at runtime. As a result codes that are divided into k clusters form a dictionary. Example of dictionary for $k = 100$ is presented in table 1.

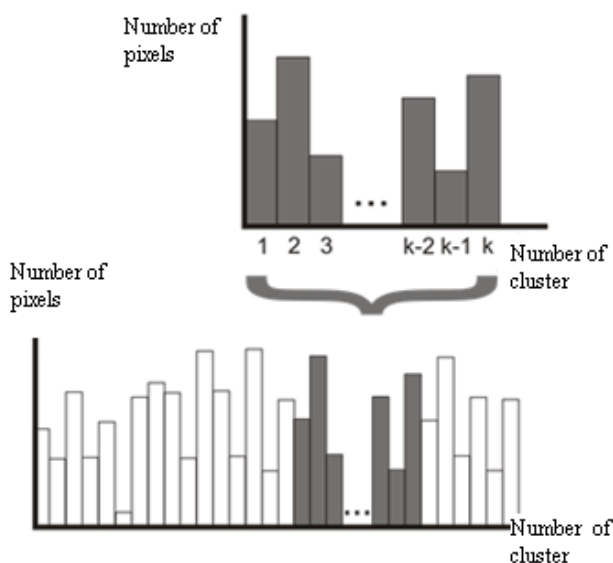| 24-bit binary code	Number of cluster
0000 0000 0000 0000 0000 0000	53
0000 0000 0000 0000 0000 0001	53
.... …. …. …. …. ….	…
0110 0011 0100 1000 0110 0001	1
0110 0011 0100 1000 0110 0010	1
0110 0011 0100 1000 0110 0011	27
.... …. …. …. …. ….	…
1010 0001 0110 1011 0111 0011	78
1010 0001 0110 1011 0111 0100	100
1010 0001 0110 1011 0111 0101	100
.... …. …. …. …. ….	…
1111 1111 1111 1111 1111 1111	91

Table 1. Example of dictionary for $k = 100$

Figure.4. Concatenating the histograms

After the first Gabor filter is applied the histogram with k bins and with height proportional to the number of pixels that fall into each cluster is formed for each block of every image from training dataset. The same procedure is repeated after applying the rest Gabor filters.

2.5. Concatenating the histograms

Histograms for all blocks of every Gabor image are concatenated (see Fig. 4). The resulting histogram may have quite big amount of bins. For example, histogram has 135x200x40 = 1 080 000 bins for $p = 9x15 = 135$ blocks and $k = 200$ clusters.

2.6. Classification stage

For recognition, histogram comparing is used to measure the similarity of different facial images and exploited for final classification. From the statistical hypothesis point of view, it cannot be said that any of the metrics would be the best one with a high (>0.95) probability.

However, histogram intersection and χ^2 measures are clearly better than log-likelihood when the average number of labels per histogram bin is low but log-likelihood performs better when this number increases. The log-likelihood measure has been preferred for texture images (Wiskott et. al., 1997) but because of its poor performance on small windows in our experiments it is not appealing for face recognition. The χ^2 measure performs slightly better than histogram intersection so we chose to use it despite the simplicity of the histogram intersection.

3. EVALUATION

To illustrate the performance of the proposed LQP-based algorithm, we conduct experiments on FERET database. It consists of 14,000 images with resolution 384x256 pixels. The FERET images are all taken indoors, with good resolution, image quality, and limited variations in lighting, size and position of faces in images. Pose of the faces in these images is typically very close to frontal.

The FERET database is subdivided into categories: The gallery images are 1196 Fa images captured during different photo sessions, with one image per person. The Fb probes consist of 1195 images with alternative facial expression; The Fc probes contain 194 images taken with a different camera and lighting; The duplicate I probes involve 722 images with corresponding gallery image taken from on a different day and in a different place, and thus they might contain "unite variation" due to illumination, facial expression, accessories, and aging, etc. The duplicate II probes, a subset of duplicate I, is considered the most challenging probe, since they contain 234 images from subjects whose gallery match was taken more than 18 months beforehand.

The experimental results for algorithm based on LQP and LQP-G (which uses Gabor filters) comparing with other popular algorithms from Open Source Computer Vision Library are shown in Table 2.

Algorithm	Testing set				
	fb	fc	dup-I	dup-II	Average
LQP (without using Gabor filters)	98,6%	45,4%	61,6%	50,4%	64,0%
LQP-G (using Gabor filters)	98,6%	96,4%	73,4%	71,4%	85,0%
LDA (OpenCV)	76,1%	9,8%	36,3%	19,2%	35,4%
PCA (OpenCV)	75,8%	6,7%	34,1%	15,4%	33,0%
LBP (OpenCV)	76,7%	21,1%	38,1%	20,9%	39,2%

Table 2. The experimental results on FERET database.

The results lead to the following conclusions:
- The proposed algorithm LQP-G allows about 20% correct recognition accuracy increase compared with the known face recognition algorithms from the OpenCV library. Medium rate of recognition accuracy is 85% (98,6% on FB dataset) when number of classes in training is 1196.
- The additional use of Gabor filters can significantly improve the robustness to changes in lighting.
- Using Gabor filters allows processing the most challenging dataset dup-II. Standard algorithms show poor recognition accuracy on this dataset.

4. CONCLUSION

This paper proposes novel LQP-based algorithm. The effectiveness of the proposed method comes from several aspects including the multi-resolution and multi-orientation Gabor decomposition, the local quantized patterns, and the local spatial histogram modeling. Experimental evaluations of the proposed approach on the FERET face database have evidently illustrated the effectiveness and robustness of algorithm to the general variations of lighting, expression, occlusion and aging.

ACKNOWLEDGEMENTS

REFERENCES

Ahonen T., Hadid A., Pietikäinen M., 2004. Face recognition with local binary patterns. In: *European Conference on Computer Vision*, pp. 469-481.

Belhumeur, P.N., Hespanha, J., Kriegman, D., 1997. Eigenfaces vs. fisherfaces: recognition using class specific linear projection In: *IEEE Transactions on Pattern Analysis and Machine Intelligence 19*. pp. 711–720. (ECCV).

Brunelli, R., Poggio, T., 1992. Face recognition through geometrical features. In: *European Conference on Computer Vision (ECCV)*. pp. 792–800.

Huang, D., Shan, C., Ardabilian, M., Wang, Y., Chen, L., 2011. Local binary patterns and its application to facial image analysis: a survey. In: *IEEE transactions on systems, man, and cybernetics* – part C: applications and reviews. V. 41, no. 6. pp. 765–781.

Hussain, S., Triggs, B., 2012. Visual recognition using local quantized patterns. In: *European Conference on Computer Vision (ECCV)*. pp. 716–729.

Hussain, S., Napoléon, T., Jurie, F., 2012. Face recognition using local quantized patterns. In: *British Machine Vision Conference*, pp. 1–11.

MacQueen, J.B., 1967. Some methods for classification and analysis of multivariate observations. In: *5th Berkeley Symposium on mathematical statistics and probability*. University of California Press, 1967. pp. 281–297.

Mandel, I.D., 1988. *Cluster analysis*. M.: Finance and Statistics,. p. 176.

Messer, K., Kittler, J., Short, J., 2006. Performance characterization of face recognition algorithms and their sensitivity to severe illumination changes. In: *International Conference on Biometrics (ICB)*. pp. 1–11.

Nikitin, A.E., Khryashchev, V.V., Priorov, A.L., Matveev, D.V., 2014. Development and analysis of face recognition algorithm based on local quantized patterns. In: *Nonlinear world*. № 8. pp. 35-42.

Petruk, V.I., Samorodov, A.V., Spiridonov, I.N., 2011. Application of Local binary patterns to solving the problem of face recognition. In: *Bulletin of the Moscow Bauman State Technical University*. Instrument. 2011. Special. vol. Biometric technology. pp. 58-63.

Shan, S.G, Gao, W., Chang, Y.Z., Cao, B., Yang, P., 2004. Review the strength of Gabor features for face recognition from the angle of its robustness to mis-alignment. In: *International Conference on Pattern Recognition (ICPR)*. pp. 338–341.

Szeliski, R., 2010. *Computer Vision: Algorithms and Applications*. Springer.

Turk, M., Pentland, A., 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, no. 3. pp. 71–86.

Vizilter, Y.V., Zheltov, S.Y., Bondarenko, A.V., Ososkov, M.V., Morzhin, A.V., 2010. *Image processing and analysis tasks in machine vision: Lectures and practical classes*. M .: Fizmatkniga.

Wiskott, L., Fellous, J., Krüger, N., Malsburg, C., 1997. Face recognition by elastic bunch graph matching. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence 19*. pp. 775–779.

Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A., 2003. Face recognition: A literature survey. In: *ACM Computing Surveys (CSUR)*, V. 35. № 4. pp. 399-458.