# The Q-Transformation for Recognition of the Facial Images

Y. F. Kutaev[1], V. P. Kozhemyako[2], L. I. Timchenko[2], A. A. Gertsiy[2], S. V. Chepornyuk[2], R. Hamdi[2].

[1]State Scientific Enterprise "Astrophisika", Volokolamskoye shosse 95, Moscow, Russia
[2]Vinnitsa State Technical University, Khmelnitskoye Shosse 95, Vinnitsa, 286021, Ukraine,
E-mail: chsv@cec.vstu.vinnica.ua, timchenko@yahoo.com

## ABSTRACT

Both the methodology and results of experimental research performed for the facial images normalization by means of compact models description that describe image classes are considered in this paper. The problems of preliminary image processing based on the method of generalized Q-transformation are being solved. The image segmentation with the formation of connectivity matrices and the formal description of the resulting components are taken into consideration as well. The calculating algorithms based on the methodology of dichotomous balance of the images being prepared have been used for the classification of human facial images. The high efficiency of recognition caused by the compact description of object images was obtained as a result and in this way it helps to make the processing more noiseproof, to improve invariance to the 2D and 3D- transformations and to reduce the calculation time requirements.

## 1. INTRODUCTION

During the last 15 years the problem of facial recognition has been given more and more attention by researchers. The latest improvements in this field use a great variety of methods. Although each method has its own specific quality, the reliable recognition has not been practically achieved yet, even for the narrow range of conditions. The basic specific problems to be solved are as follows: *3D* rotations, the invariance to facial expression, various sights and the integration of information from different scales.

The main goal of this research is to solve the problem of the integration of the high-level information into the low-level facial presentation [1]. This will result in higher recognition efficiency. The advantage of such recognition is compact face presentation which considers the reduced size of the data sets and faster efficiency.

A human being tends to concentrate upon some high level peculiarities while recognizing unfamiliar faces. On the contrary, most automatic systems process faces as a whole. Kirby and Sirovich [2], Lanitis and others [3] performed the analysis of principal components considering a face as a whole. Blackwell and others [4] noted that the preliminary processing of the whole image cannot solve the problems connected with the integration of large data sets.The other system class processes located descriptors. Lades and others [5] recognized faces without integrating the regularity of a facial class. Penev and Atick [6] performed the analysis of a local peculiarity on the basic of each human's individual view.

## 2. METHODOLOGY

The general methodology of the compact image presentation is given in accordance with the scheme of *coarse - fine* processing. The coarse processing consists of image quantization and is followed by its division into spatially-connected segments [7]. The results of this coarse processing are separate segments structured in some fields according to their connectivity indexes. These image segments are formed from spatially connected pixels of a quantified image and then create connectivity fields. In the case of a human face, coarse processing these fields can reflect its distinctive components (eyes, cheeks, mouth, etc.). Then separate segments undergo the *fine* processing scheme. In this case, the processing of the segments which are less deformed because of different facial expressions becomes possible.

### 2.1 The fine scheme of image processing.

The fine processing scheme is based upon the separate segment analysis and includes image preparing, i.e. transforming the image into the matrix of 3-level binary preparations (positive, negative and zero) and their subsequent transformation by means of forming local and general equalization functions. Further correlation analysis of those functions is provided as well. The preliminary processing includes all subsequent operations. The input *2D* grey-scale image is presented by the matrix of light intensity pixels $A = \left[a_{i,j}\right]$, $i = 1 \div N$, $j = 1 \div M$, where $M \times N$ is the input image dimensions. The general mean of the present matrix A is found as:

$$a = \frac{1}{NM}\sum_{i,j} a_{i,j} \qquad (1)$$

Then sets of differences between pixels and obtained means of the whole image or its segments, are determined:

$$R_{i,j} = a_{i,j} - \overline{a} \qquad (2)$$

In order to obtain the set of preparations this differences are compared with the threshold $\delta$ [8, 9]:

$$q_{i,j} = \begin{cases} 1, \text{ if } R_{i,j} > \delta \\ -1, \text{ if } R_{i,j} < (-\delta) \\ 0, \text{ if } \left| R_{i,j} \right| \le \delta \end{cases} \qquad (3)$$

The experiments have shown that the usage of the function that equally distributes positive, negative and zero preparations is most useful for the facial image preparing. It makes it possible to mark areas with and without brightness differences (according to the threshold) in a proper way. The threshold itself is adapted to the image. It results in segmenting separate fields of an initial image for their further analysis. The threshold $\delta$ is chosen from the following condition [9]:

$$MAX\left(N_t^{(1)}, N_t^{(-1)}, N_t^{(0)}\right) \qquad (4)$$

where $N_t^{(1)}, N_t^{(-1)}, N_t^{(0)}$ is the number of positive, negative and zero preparations within each grey level -*t*.

The matrix of preparations $Q^x = \left[q_{i,j}^x\right]$ is formed as a result of the above mentioned operations.

## 2.2 The image segmentation.

In order to get certain features [7] from the obtained matrices the connectivity indexes are determined separately for positive, negative and zero preparations. The stages of image segmentation in accordance to the meaning of their connectivity indexes are mentioned below.

The general connectivity index of the whole prepared image is found. The connectivity spectrum of the image with the *M*N*-dimentions is defined as

$$W = \sum_{v=1}^{8} \sum_{i,j}^{N,M} q_{i,j}^{v},$$

where $q_{i,j}^{v}$ is an element of the image with *ij*-coordinates and *v*-connectivity.

## 2.3 The formal description of image parts.

The following decomposition based on dichotomous principle could be used to analyse the obtained connectivity spectrum. Let us divide the whole prepared matrix in a column (row) direction so that two divided parts will have equal connectivity indexes. The operation of column (row) equalization should be introduced for each level of this dichotomous decomposition. The local column equalization function $U^* = \left[u_j^*\right]$ has to meet the following requirement:

$$\sum_{i=1}^{u_j^*} q_{i,j}^{v} = \sum_{i=u_j^*+1}^{N} q_{i,j}^{v}, \quad u_j \in \{1,2,..,N\} \qquad (6)$$

Then, the general column equalization function $U^{**} = \left[u_j^{**}\right]$ is determined as follows:

$$\sum_{j=1}^{u_j^{**}} \sum_{l=1}^{j} q_{i,j}^{v} = \sum_{j=u_j^{**}+1}^{N} \sum_{l=1}^{j} q_{i,j}^{v} \qquad (7)$$

Besides, this equalization process is performed for each row and column of the prepared image. Furthermore, the obtained connectivity indexes are added to each of the divided fields.

The equalization procedure is carried out during each step of this decomposition. The operations being described are performed separately for positive, negative and zero preparations. Then a classifying analysis method has to be used as described in [10] for recognition and identification of human faces with the usage of operations based on the network structure.

# 3. EXPERIMENTAL RESULTS.

Two basic noise distributions most common in the tasks of image processing - Gaussian (normal) distribution and uniform noise distribution have been chosen to analyse a noise invariancy of this method. 7 different gradations with such levels as 0%, 5%, 10%, 15%, 20%, 25%, 30% were organized for the experiment being held. To implement image preparing, the proper threshold value was selected from the range $\delta \in \left\{1,2,..\frac{I_{\max}}{2}\right\}$, $I_{\max} = 255$ gray scale levels.

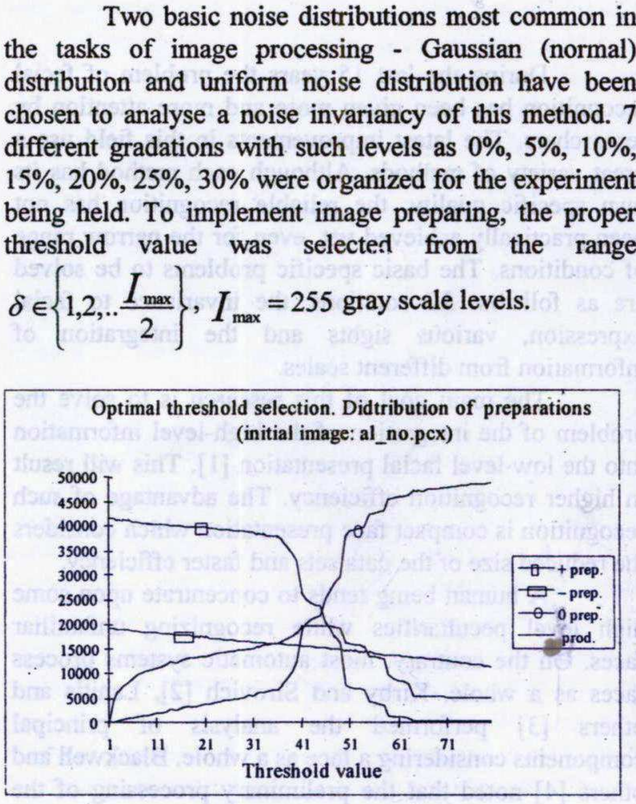


Fig.1 The preparation distributions for the initial image

Those threshold values that exceed this range do not meet the above mentioned requirements because the number of zero preparations prevails. Image preparing was performed as a test for each current value of the threshold with simultaneous calculation of positive, negative and zero preparation. The preparation distributions for the initial image are shown in Fig. 1 and 2.

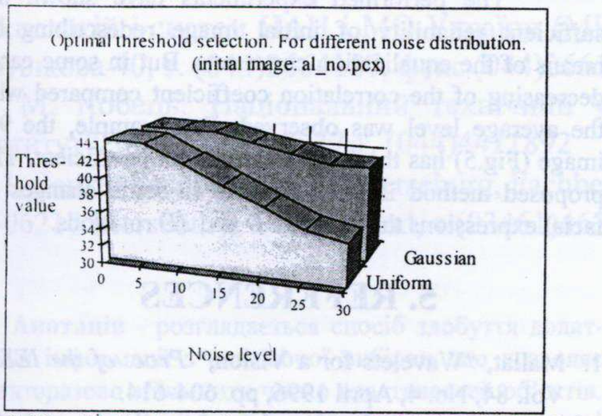


Fig.2. The distributions of optimum threshold for noise-added images

Fig. 1 shows that the optimum value for the initial image is 44. The analogous operations for all noise levels were performed during the experiments. The higher noise level causes decreasing the threshold value, especially in case of the uniform noise distribution. The vertical and horizontal equalization curves were obtained for the initial and noise-added images during the experiment.

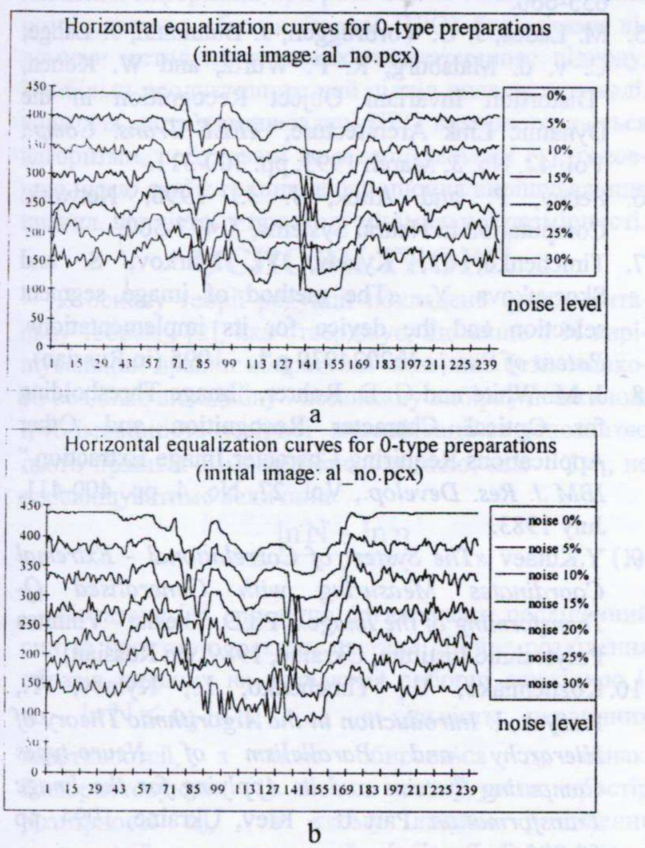


Fig.3 Horizontal equalization curves. a) Gaussian noise distribution; b) Uniform noise distribution

The resulting curves with offset are shown in Fig.3, where the highest curve corresponds with the initial image without noise. The results of the correlational comparison of the initial and noise-added image curves can be found in Fig. 4.

The correlative comparison of equalization curve sets (Fig.5) obtained in the row direction was carried out for different facial expressions. These images are shown in Fig.6. The image with normal expression (al_no1) was used as a reference for further comparison with other images which contain 9 other facial expressions and *3D* rotations of the same face (al_mim2..al_mim10). The optimal thresholds were obtained for these images. The results of correlational comparison are shown in Fig. 7.

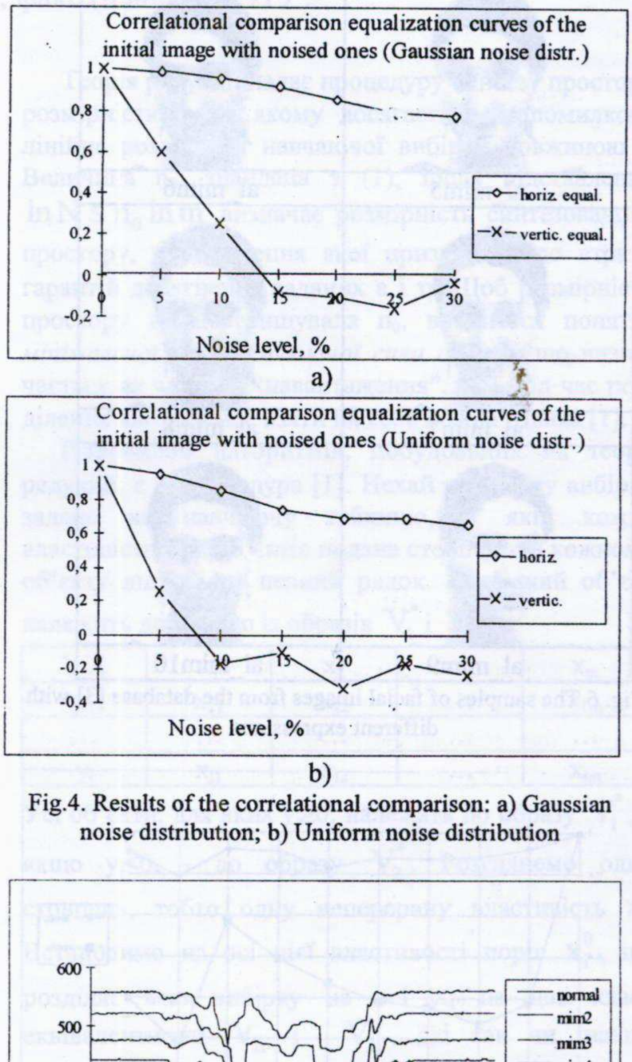


Fig.4. Results of the correlational comparison: a) Gaussian noise distribution; b) Uniform noise distribution

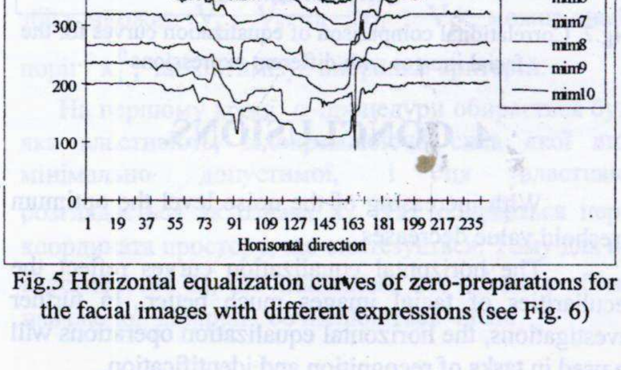


Fig.5 Horizontal equalization curves of zero-preparations for the facial images with different expressions (see Fig. 6)

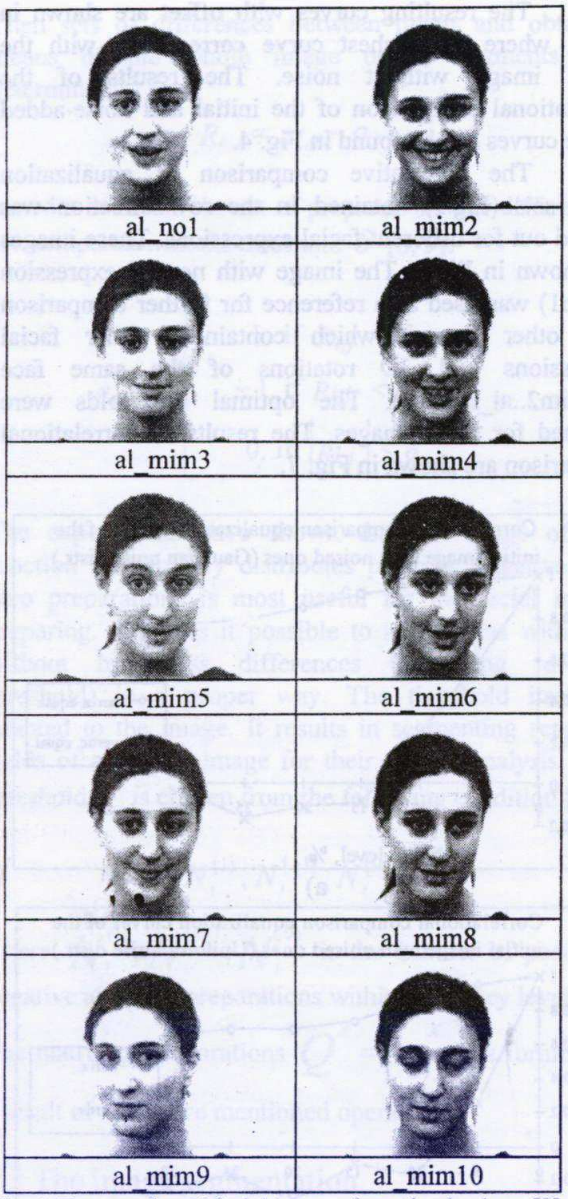


Fig. 6 The samples of facial images from the database [3] with different expressions

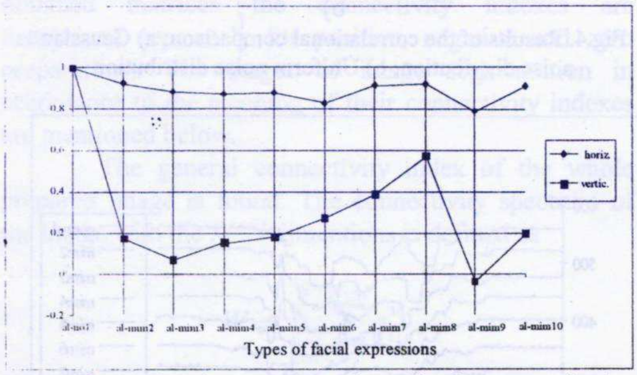


Fig.7. Correlational comparison of equalization curves for the facial images with different expressions

# 4. CONCLUSIONS.

With increasing of the noise level the optimum threshold value decreases.

The horizontal equalization curves reflect the peculiarities of facial images much better. In further investigations, the horizontal equalization operations will be used in tasks of recognition and identification.

In the case of the noise level having been increased to 30 %, the correlation coefficient remains rather high - 0.8 (Gaussian noise) and 0.65 (uniform noise); that is caused by the usage of operations of addition which minimize the noise influence.

Due to the transformation of the initial *2D* image into a *1D* equalization curve set the further correlational calculations becomes faster.

The performed experiments have shown the sufficient reliability of initial image redescribing by means of the equalization curve sets. But in some cases decreasing of the correlation coefficient compared with the average level was observed. For example, the 9th image (Fig.5) has the lowest correlation coefficient. The proposed method is more reliable to some changes of facial expressions than to the *2D* and *3D* rotations.

# 5. REFERENCES

1. Mallat, "Wavelets for a Vision," *Proc. of the IEEE*, Vol. 84, No. 4, April 1996, pp. 604-614.
2. M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-12,1990.
3. A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic Face Identification System Using Flexible Appearance Models," *Image and Vision Computing*, Vol. 13, No. 5, June 1995, pp. 393-401.
4. Blackwell, K. T., Vogl, T.P., Hyman, S.D., Barbour, G.S., and Alkon, D.L. 1992. Pattern Recognition,25, 655-666.
5. M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. v. d. Malsburg, R. P. Wurtz, and W. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Trans. Comp.*, Vol. 42, No. 3, March 1993, pp. 300-311.
6. Penev, P. and Atick, J. J., 1996, Network: Computation in Neural Systems, 7, 477-500.
7. Timchenko, L., Kytaev, Y., Markov, S. and Skoryukova, Y., «The method of image segment selection and the device for its implementation». *Patent of Russia* №2024939 p.1. - 1995.(in Russian).
8. J. M. White and G. D. Rohrer, "Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction," *IBM J. Res. Develop.*, Vol. 27, No. 4, pp. 400-411, July 1983.
9. Y.Kutaev *«The System of Correlational - Extremal Coordinates Measuring with Generalised Q-Transforming of the Image»*. Ph.D. Thesis: - Vinnitsa Polytechnic Institute, Ukraine, 1989, (in Russian)
10. Kozhemiako, V., Timchenko, L., Kytaev, Y., Ivasyuk, I. *Introduction in the Algorithmic Theory of Hierarchy and Parallelism of Neuro-types Computing Systems and its Applying for the Image Transformation.* Part II. - Kiev, Ukraine, 1994, pp 65-205 (in Russian)