

# НЕЙРОИНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

# НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

# NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 004.825:004.932.72'1

## FEATURE VECTOR GENERATION FOR THE FACIAL EXPRESSION RECOGNITION USING NEO-FUZZY SYSTEM

**Bodyanskiy Ye. V.** – Dr. Sc., Professor, Head of Control Systems Research Laboratory, Professor of Artificial Intelligence Department at Kharkiv National University of Radioelectronics, Kharkiv, Ukraine.

**Kulishova N. Ye.** – PhD, Professor of Media Systems and Technologies Department, Kharkiv National University of Radioelectronics, Kharkiv, Ukraine.

**Tkachenko V. Ph.** – PhD, Professor, the Head of Media Systems and Technologies Department, Kharkiv National University of Radioelectronics, Kharkiv, Ukraine.

### ABSTRACT

**Context.** The article is devoted to the problem of a training data set forming for the automatic human emotions recognition system on the basis of a multidimensional extended neo-fuzzy neuron. The aspects of choice the attributes vector's dimension and composition, their influence on the system learning rate are considered. The object of research is the method of multidimensional data clustering. The subject of research is two-dimensional images geometric features systematization.

**Objective.** The main goal of the work is to develop an approach to person's face expression description using geometric features fixed set that can be obtained by video sequence frames processing.

**Method.** To study the facial expressions recognition system it is proposed to form a feature vector consisting of characteristic points coordinates. There were selected points that relate to the location and shape of the eyelids, eyebrows, eye pupils, lips contours, nose wings, nasolabial folds. Such points can be easily found during the automatic image processing using known contour detectors. Also, the possibility of using for the human facial expression description not the coordinates of characteristic points, but the distances between them, was investigated. From these distances a different feature vector was created, the properties of which were compared with the points coordinates vector.

**Results.** The developed recognition system on the basis of a multidimensional extended neo-fuzzy neuron have been implemented in software and investigated for solving the problem of facial expression classification. A comparison between the attribute vectors that are different in composition and dimension is made. The structure for the feature vector, which provides high system learning rate, and does not require the additional structural elements was chosen.

**Conclusions.** The experimental study fully confirms the effectiveness of the developed approach for the human facial expressions recognition using a multidimensional extended neo-fuzzy neuron.

**KEYWORDS:** facial expressions recognition, characteristic features, neo-fuzzy neuron, membership function, fuzzy clustering, computational intelligence.

### ABBREVIATIONS

2D – two-dimensional;  
3D – three-dimensional;  
FACS – facial action coding system;  
AU – action units;  
N – feature points number;  
MENFN – multidimensional extended neo-fuzzy neuron;  
PICS – Psychological Image Collection at Stirling image database;  
CK – Extended Cohn-Kanade image database;

SURF – speeded up robust features detector and descriptor.

### NOMENCLATURE

$x_i$  – input data vector;  
 $w_{li}$  –  $l$ -th synaptic weight;  
 $i$  – number of nonlinear synapse in neo-fuzzy neuron;  
 $l$  – number of membership function in nonlinear synapse;  
 $\mu_{li}(x_i)$  – the membership function on  $i$ -th input;

$\tilde{w}^T$  – synaptic weights vector;  
 $f_i(x_i)$  – output signal of nonlinear neo-fuzzy synapse;  
 $v_j(k)$  – output signal of system intermediate layer;  
 $\tilde{y}_j(k)$  – output signal of system first layer;  
 $m$  – number of emotion to recognize;  
 $\tilde{y}(k)$  – system output signal.

X1, Y1, X2, Y2, ..., X35, Y35 – the coordinates of the proposed 35 feature points;

D1, ..., D24 – the distance indicators of the proposed 35 feature points position;

K1, ..., K14 – the ratio indicators of the proposed 35 feature points position.

## INTRODUCTION

The human face is a communicative system with many inputs and many outputs, which is very flexible. To transfer information in this system, several signal groups are used [1, 2]:

1) static signals are relatively permanent signs, such as bone structure, soft tissues and general proportions. These signals are usually used to identify a person;

2) slow facial signals are changes in the person appearance that occur gradually over time, such as developing permanent wrinkles and changing the texture of the skin. These signals can reduce the distinctness of the facial features and prevent fast facial signals recognition;

3) artificial signals are exogenous facial features, such as eyeglasses and cosmetics. They can hide facial features or, conversely, strengthen;

4) fast facial signals are temporary changes in neuromuscular activity that can lead to visually detectable changes in the face appearance. These person atomic signals are the basis of facial expressions.

Given the significant role of the individual in social life, it is not surprising that the potential benefits of efforts to automate the facial signals analysis, in particular fast facial signals, are diverse and numerous. Especially it concerns human-computer interfaces. Automatic analysis of fast facial signals is used in various vision subsystems, including gaze tracking, lip reading, bimodal speech processing, morphemes visual synthesis, and the formation of commands based on facial expressions. Among the areas where they are applied [3]:

- control the movement of one or more vehicles, aircraft and car control;
- control of high-risk facilities;
- monitoring elderly patients in hospitals;
- video conferences, internet lectures and distance learning.

In numerous facial expressions automatic recognition systems, two approaches to encoding information are used primarily: the detection of facial features changes and the facial muscles action. These approaches are based on the work of psychologists and physiologists. To use them in automatic analysis, it is necessary to find corresponding

changes in the image characteristics (static in the photo or dynamic in the video sequence) for each psycho-physiological feature. Such characteristics include contour, structural and textural image descriptions. For example, these are contour corners, contour extreme points, individual regions color data, the texture properties, etc.

The human-computer interface involves interaction in real time. This means that the number of characteristics that are calculated from the person's face image during the dialogue should not be too large. On the other hand, for a precise identification of a person's state, a set of characteristics cannot be small, since the possible emotional states number is large enough. Therefore, the problem of a data set forming for the automatic person's expression recognition on the image remains unresolved.

**The object of study** is the method of multidimensional data clustering.

**The subject of study** is two-dimensional images geometric features systematization.

**The purpose of the work** is increasing the accuracy of a person's face expression automatic recognition in real-time, when the sizes of the training data are small.

## 1 PROBLEM STATEMENT

Automatic facial expressions recognition is reduced to solving the data clustering problem. If the data comes from the video, it is obvious that it can contain information about static, slow and exogenous signals that do not change from frame to frame. However, the most important are the fast signals that need to be registered, processed and recognized in each frame. The changing fast signals dynamics also indicate a psycho-emotional state. Therefore, it is necessary to register and process the information about fast signals formation rate in real time. Such rigid limitations to the rate of facial expressions automatic recognition method lead to the necessity of using a compact but informative features vector that will ensure the information preservation and high clustering rate.

The main goal of this work is to choose a way of face features detecting for an online human facial expression recognition using a system based on multidimensional neo-fuzzy neurons.

## 2 REVIEW OF THE LITERATURE

The task of automatic facial expression recognition is complex and multistage. It includes the image pre-processing and face area searching. After the face area is detected, it is possible to recognize the emotion according to the facial characteristic features set. Depending on the chosen approach, the next stage of the problem solving - the calculation of the characteristic vector. If the recognition is performed in real time, then the 3D model use could be inefficient due to the high time costs. For online applications, the 2D image features vector (received as a result of frame-based analysis) will be justified.

The recognition of faces and their expressions uses several descriptor types. The most common are local binary masks (Local Binary Patterns), attributes selection using Gabor Facial Features, active and adaptive models of the shape and appearance (Active Shape Model, Active Appearance Model) [4]. Examples of these descriptors are shown in Fig. 1. In addition, curvilinear contours can be represented in the form of splines, as proposed, for example, in [5]. Local binary masks can be attributed more likely to textural features, since they represent generalized histograms of the face image fixed segments. Such a description will be less informative in emotions recognition.

The Gabor filtering results are definitely interesting for the facial expression recognition. Usually features detected in the image are either geometric shape features of the facial components (eyes, lips, wings of the nose, etc.) or the singular points location (corners of the eyes, mouth, etc.) or appearance elements, representing the skin face texture, including wrinkles, bulges and furrows. Typical examples of geometric characteristics are described in [6], where a 19-point grid is used; in [7], where a model of shape, defined by 58 landmarks for the face is used, and in [8]. Some examples of a hybrid description of the geometric features and appearance model are the methods in [9], where the eye, eyebrows and mouth shapes and wrinkles (crows and nasolabial furrows) were used, and in [10], where 26 facial points around the eyes, eyebrows and mouth were used and the same descriptors as in [9]. Methods based on appearance are the use of Gabor wavelets in [11, 12], and the application of a holistic, monochrome, spatial and temporal patterns of facial coefficients [13]. Adaptive model in this sense is more useful, since it includes not only nodal points, but also contours (eyelids, eyebrows, lips, chin). Such signs, undoubtedly, convey exhaustive information about the emotional state of a person, even if it has a weak degree of expression.

Now, for the labeling of facial actions, the facial action coding system is widely used [1, 2]. FACS links the changes in the face to the muscles actions that produce them. It defines 44 different action units, which are considered to be the minimum visually perceptible face movements. The face expression descriptors are most often used for the six basic emotions (fear, sadness, happiness, anger, disgust, surprise) suggested by Ekman and theorists of discrete emotions. They assume that these emotions are everywhere displayed and recognized from facial expressions. FACS also provides rules for recognizing AU time segments (start, culmination, completion) in the front-end video. However, the main disadvantage of this system for automatic video recognition is the fact that the movements of the facial muscles on 2D frame images can not always be detected by filters, detectors, or other algorithms to numerically evaluate the properties of images. The researchers noted that methods based on geometric features, often exceeded those that are based on the use, for example, of wavelets or Gabor eigenvectors [11].

A separate problem is the definition of dimension of such a model. It is directly related to the clustering method, which is used for recognition. For example, the model in Fig. 1c contains 83 points, taking into account the number of their coordinates, the dimension of the feature vector can be equal to  $N = 166$  (in the 2D case) or  $N = 249$  (in the 3D case). If a neural network or neuro-fuzzy system is selected for clustering, the number of adjusting parameters can be  $2^N$ .

For some types of neural networks, neuro-fuzzy systems, it is possible to reduce the number of configurable parameters to  $N \times M$ , where  $M$  is the dimension of the output vector (it is determined by the number of recognizable emotions and their combinations). Nevertheless, for real-time processing, the represented dimension of the attributes input vector will also be unacceptable.

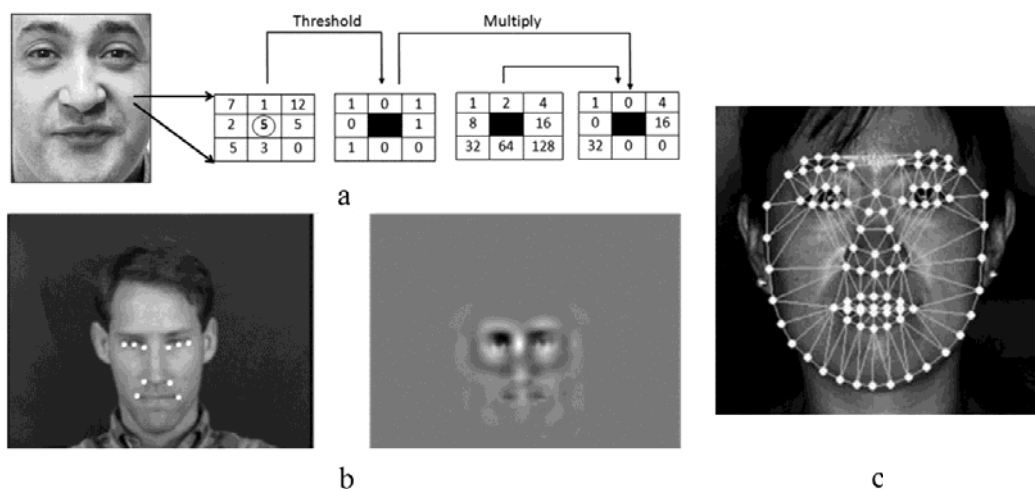


Figure 1 – Examples of approaches to the face characteristic features determining: a – using local binary masks; b – using the Gabor filter; c – using the adaptive model of appearance

### 3 MATERIALS AND METHODS

In this paper, the system of human face expressions automatic recognition in real time, based on the multidimensional extended neo-fuzzy neuron [14], is considered.

This system implements a fuzzy Takagi-Sugeno inference of an arbitrary order ( $p-1$ ) in the form:

$$\text{IF } x_i \text{ IS } x_{li} \text{ THEN THE } f(x_i) \text{ IS } w_{li}^0 + w_{li}^1 x_i + \dots + w_{li}^p x_i^p, l = 1, 2, \dots, h. \quad (1)$$

Its architecture is shown in Fig. 2. The input layer of the system consists of extended neo-fuzzy neurons, the intermediate layer consists of elements performing non-linear transform, and the output layer combines the output values into the resulting vector.

In the input layer, the extended neo-fuzzy neurons convert the input signals  $x(k) = (x_1(k), \dots, x_i(k), \dots, x_n(k))^T$  as follows:

$$f_i(x_i) = w_i^T \tilde{\mu}_i(x_i), \quad (2)$$

$$\hat{y} = \sum_{i=1}^n f_i(x_i) = \sum_{i=1}^n w_i^T \tilde{\mu}_i(x_i) = \tilde{w}^T \tilde{\mu}(x), \quad (3)$$

where  $\tilde{w}^T = (w_1^T, \dots, w_i^T, \dots, w_n^T)^T$ ,  $\tilde{\mu}(x) = (\tilde{\mu}_1^T(x_1), \dots, \tilde{\mu}_i^T(x_i), \dots, \tilde{\mu}_n^T(x_n))^T$ ,  $w_{li}$  –  $l$ -th synaptic weight of the  $i$ -

th nonlinear synapse in neo-fuzzy neuron;  $l = 1, 2, \dots, h$ ,  $i = 1, 2, \dots, n$ ;  $\mu_{li}(x_i)$  – the  $l$ -th membership function in the  $i$ -th nonlinear synapse, realizing the fuzzification of a crisp component  $x_i$ .

Elements of the intermediate layer compute the signal:

$$v_j(k) = \psi(\tilde{y}_j(k)) = \tanh(\tilde{y}_j(k)). \quad (4)$$

Output layer detects the maximums in the calculated learning algorithm values  $v_j(k)$ ,  $j = 1, 2, \dots, m$ ;  $m$  – number of emotion to recognize:

$$\tilde{y}(k) = \sup_{j=1}^m \{v_j(k)\}, \quad (5)$$

which is necessary if the learning vector set in the range  $[0, 1]$ .

As a basis for the feature vector, it is proposed to use a set of 35 characteristic points. Their location scheme is shown in Fig. 3.

Taking into account the number of their coordinates  $X1, Y1, X2, Y2, \dots, X35, Y35$ , the dimension of the feature vector can be equal to  $N = 70$  (for the recognition by 2D images). These points can be localized in the face area using contour detectors, for example, using SURF [15] or Shi-Tomasi detectors [16].

The placement of such points can indicate the basic facial actions of the FACS system in the facial dynamics. The action unit codes and the numbers of the corresponding characteristic points are presented in Table 1.

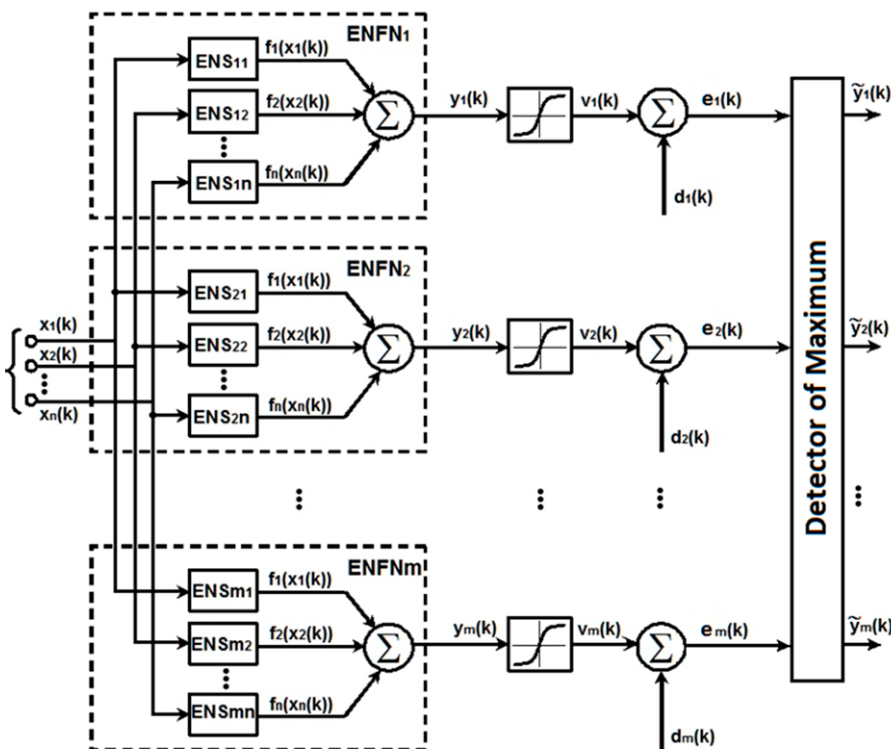


Figure 2 – Multidimensional extended neo-fuzzy neuron

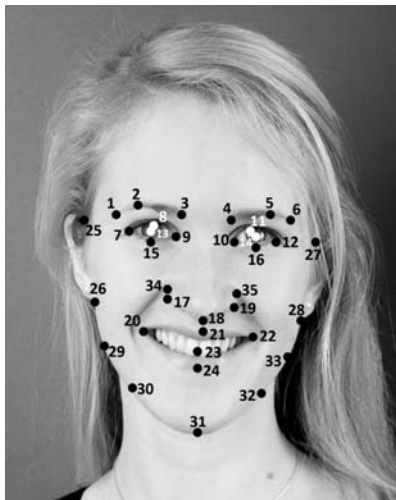


Figure 3 – 35 feature points location

As noted in [1, 2], the basic human emotions (anger, fear, disgust, happiness) correspond to the repetitive sequences of certain AUs:

- Anger A4+A5+A7+A23
- Disgust A9+A15+A16
- Fear A1+A2+A4+A5+A7+A20+ A 26
- Happiness A6+A12
- Sorrow A1+A4+A15
- Surprise A1+A2+A5+A26.

Then the emotion of anger, for example, can be detected by changes in the position of proposed points as:

$$A4+A5+A7+A23 = (1, 2, 3, 7, 8) + (7, 8) + (3, 7, 8, 15) + (20, 21, 24, 29, 30, 31).$$

for the right side of the face and

$$A4+A5+A7+A23 = (4, 5, 6, 11, 12) + (11, 12) + (4, 11, 12, 16) + (21, 24, 22, 31, 32, 33).$$

for the left side of the face. Other basic emotions are indicated using the proposed set of points in a similar way.

Adaptive models of shape appearance use not only positions of individual points, but also the location of especial lines. These include nasolabial folds, facial wrinkles and so on. In addition, the relative position of the eyes, corners of the mouth, and the wings of the nose also indicate a change in the state of a person. These indicators can be calculated on the basis of the values of the coordinates of the proposed 35 characteristic points.

Absolute and relative distances from the corners of the lips to the outer corners of the eyes (Fig. 4) are calculated in accordance with the formula (6):

$$D1 = \sqrt{(X17 - X20)^2 + (Y17 - Y20)^2},$$

$$D2 = \sqrt{(X19 - X22)^2 + (Y19 - Y22)^2},$$

$$m = \sqrt{(X18 - X21)^2 + (Y18 - Y21)^2},$$

$$K12 = \frac{D1}{m},$$

$$K13 = \frac{D2}{m}.$$

The distance from lips corners to the outer eyes corners (Fig. 4):

$$D3 = \sqrt{(X7 - X20)^2 + (Y7 - Y20)^2},$$

$$D4 = \sqrt{(X12 - X22)^2 + (Y12 - Y22)^2}.$$

The distance from the lips corners to the eyebrows outer corners (Fig. 4):

$$D5 = \sqrt{(X1 - X20)^2 + (Y1 - Y20)^2},$$

$$D6 = \sqrt{(X6 - X22)^2 + (Y6 - Y22)^2}.$$

Table 1 – The action unit codes and the numbers of the corresponding characteristic points

Facial action units (AU) of the FACS system	Numbers of the corresponding characteristic points of the right side of the face	Numbers of the corresponding characteristic points of the left side of the face
A1	2, 3	4, 5
A2	1,2	5,6
A4	1,2,3, 7,8	4,5,6,11,12
A5	7, 8	11, 12
A6	7,15,34	12, 14, 35
A7	3,7,8,15	4,11,12,16
A9+A10	1,2,3,7,8,9,34,17,20,21	4,5,6,10,11,12,35,19,20,22
A11	17,34	19,35
A12	20,21,24,34,29,30,31	21,24,22,35,31,32,33
A13	20,21,24,34,29,30,31	21,24,22,35,31,32,33
A14	20,21,24,34,29,30,31	21,24,22,35,31,32,33
A15	20,21,24,29,30,31	21,24,22,31,32,33
A16	20,21,24,30,31	21,24,22,31,32
A17	20,21,24,30,31	21,24,22,31,32
A18	20,21,24,34,29,30	21,24,22,35,32,33
A20	20,21,24,29,30,31	21,24,22,31,32,33
A22	20,21,24,34,29,30,31	21,24,22,35,31,32,33
A23	20,21,24,29,30,31	21,24,22,31,32,33
A24	20,21,24	21,24,22
A25, A26,A27, A28	20,21,24,29,30,31	21,24,22,31,32,33
A41,A42,A43	1,2,3,7,8,9	4,5,6,10,11,12
A44,A45,A46	1,2,3,7,8,9,15	4,5,6,10,11,12,16

The distance from the lips corners to the eyebrows inner corners and their ratio (Fig. 4):

$$\begin{aligned} D7 &= \sqrt{(X3 - X20)^2 + (Y3 - Y20)^2}, \\ D8 &= \sqrt{(X4 - X22)^2 + (Y4 - Y22)^2}, \\ K11 &= \frac{D7}{D8}. \end{aligned} \quad (9)$$

The mutual eyebrows and eyes position (Fig. 5):

$$\begin{aligned} D10 &= Y2 - Y8, & D14 &= Y8 - Y15, \\ D11 &= Y5 - Y11, & D15 &= Y11 - Y16, \\ D12 &= Y2 - Y13, & & \\ D13 &= Y5 - Y14, & & \end{aligned} \quad (10)$$

The eyebrows position and curvature (Fig. 5):

$$\begin{aligned} D17 &= Y3 - Y1, \\ D18 &= Y3 - Y2, \\ D19 &= Y4 - Y6, \\ D20 &= Y4 - Y5. \end{aligned} \quad (11)$$

The mutual position of the eyes, wings of the nose and cheeks (Fig. 5):

$$\begin{aligned} D21 &= \sqrt{(X9 - X34)^2 + (Y9 - Y34)^2}, \\ D22 &= \sqrt{(X10 - X35)^2 + (Y10 - Y35)^2}, \\ D23 &= \sqrt{(X17 - X34)^2 + (Y17 - Y34)^2}, \\ D24 &= \sqrt{(X19 - X35)^2 + (Y19 - Y35)^2}. \end{aligned} \quad (12)$$

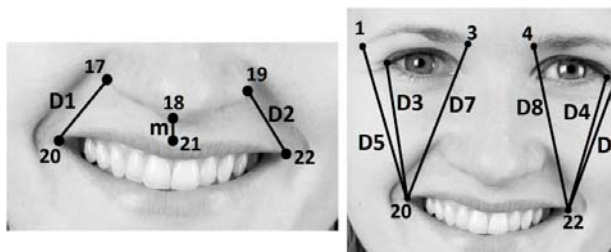


Figure 4 – Scheme for calculating the absolute and relative distances from the lips corners to the outer eyes corners, the distance from the lips corners to the eyebrows outer corners, the distance from the lips corners to the eyebrows inner corners and their ratio

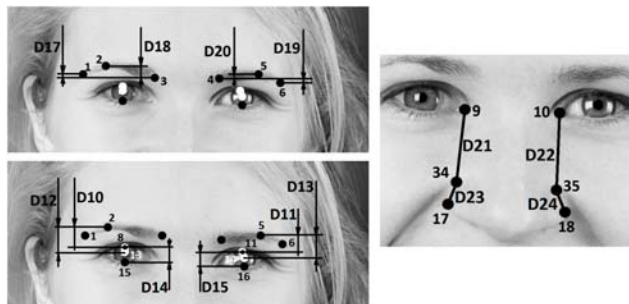


Figure 5 – Scheme for calculating the mutual eyebrows and eyes position, the eyebrows position and curvature, the mutual position of the eyes, wings of the nose and cheeks

The relative mouth opening and width (Fig. 6):

$$\begin{aligned} a &= Y24 - Y21, \\ b &= Y31 - Y18, \\ c &= X19 - X17, \\ d &= X22 - X20, \\ K1 &= \frac{a}{b}; K2 = \frac{d}{c}; K14 = \frac{d}{a}. \end{aligned} \quad (13)$$

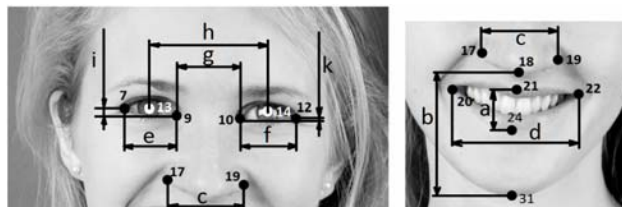


Figure 6 – Scheme for calculating the relative mouth opening and width, the relative eye size and inclination

The relative eye size (Fig. 6):

$$\begin{aligned} e &= X9 - X7, \\ f &= X12 - X10, \\ K3 &= \frac{e}{c}; K4 = \frac{f}{c}. \end{aligned} \quad (14)$$

The eye inclination (Fig.6):

$$\begin{aligned} g &= X10 - X9, \\ h &= X14 - X13, \\ i &= Y9 - Y7, \\ k &= Y12 - Y10, \\ K5 &= \frac{e}{h}; K6 = \frac{f}{h}; K7 = \frac{i}{g}; K8 = \frac{k}{g}. \end{aligned} \quad (15)$$

#### 4 EXPERIMENTS

For a comparative evaluation of the different in composition and dimension characteristic points vectors, photographs from two open image databases that were designed to solve the problems of facial expressions recognition were used: Psychological Image Collection at Stirling [17], partly from the Extended Cohn-Kanade database [18]. Another important question was the verification of the recognition system on the basis of the neo-fuzzy neuron, if the training sample is small (up to 1000 samples). Photographs were selected from the image databases, where the person emotional state expression degree is different – from barely visible, to strongly noticeable.

The total number of photos in the training sample was 344, their distribution by basic emotions is shown in Table 2.

Table 2 – Dimensions of photos training sets for single emotions

Emotion	Anger	Disgust	Fear	Happiness	Sorrow	Surprise	Neutral
Data set size	49	66	35	45	19	50	80

For these photographs, the coordinates of the proposed 35 points were found:  $X_1, Y_1, X_2, Y_2, \dots, X_{35}, Y_{35}$ , and then calculated the indicators  $D_1, \dots, D_{24}$  and  $K_1, \dots, K_{14}$ . To test the possibility of the feature vector dimension reducing, two sets of data were additionally formed. The first set includes the coordinates of four points from the original vector. These are points 13, 14, 18, 23 (Fig. 7a). In the second set – the coordinates of 22 points (Fig. 7b).

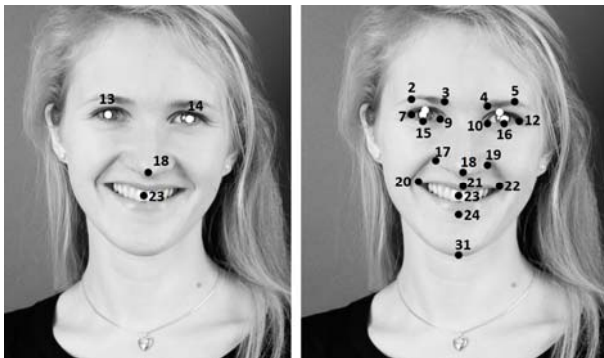


Figure 7 – Disposition of points in additional data sets

The resulting feature vectors were used to learn a facial expression automatic recognizing system based on a multidimensional extended neo-fuzzy neuron (Fig. 2). In the power polynomial of this neuron nonlinear synapse, the terms number equal to three is assumed. Thus, this system implements the Takagi-Sugeno second order fuzzy inference. The number of membership functions and learning epochs was variable. The recognition system learning results are given in Tables 3–6.

## 5 RESULTS

As a result of the experiment, it was found that the reduction of the feature vector dimension can be compensated by increasing the membership functions number in the multidimensional extended neo-fuzzy neuron structure. Characteristics in the form of distances between feature points reduce the system learning rate. This problem can also be solved by increasing the membership functions number.

Table 3 – Percentage of unrecognized emotions, when the feature vector includes the coordinates of 4 points

Number of learning epochs	Basic emotions							The average percentage of unrecognized emotions
	Anger	Disgust	Fear	Happiness	Sorrow	Surprise	Neutral	
Number of membership functions = 7								
10	55	80	100	86.7	100	92	53.8	93
30	67.4	72.7	100	82.2	89.5	82	50	77,3
50	63.3	65.2	100	77.8	89.5	76	50	75,3
70	65.3	63.6	100	77.8	89.5	72	48.8	86
150	61.2	59.1	100	77.8	84.2	68	47.5	84
2000	57.1	42.4	92.9	75.6	57.9	46	51.3	46
Number of membership functions = 9								
2000	59.2	37.9	80	73.3	36.8	36	43.8	58
Number of membership functions = 11								
2000	57.1	50	68.6	66.7	36.8	40	47.5	45

Table 4 – Percentage of unrecognized emotions, when the feature vector includes the coordinates of 22 points

Number of learning epochs	Basic emotions							The average percentage of unrecognized emotions
	Anger	Disgust	Fear	Happiness	Sorrow	Surprise	Neutral	
Number of membership functions = 7								
10	63.3	63.6	94.3	77.8	89.5	62	66.3	73.8
30	53	48.5	82.9	68.9	84.2	40	58.8	62.3
50	51	40.1	77.1	66.7	84.2	30	53.8	57.5
70	48.9	43.9	65.7	48.9	73.7	28	42.5	50.2
150	36.7	33.3	65.7	44.4	36.8	20	36.3	39.0
2000	18.4	4.6	25.7	15.6	15.8	6	17.5	14.8
Number of membership functions = 9								
2000	12.2	10.6	17.1	6.7	10.5	4	18.8	11.4
Number of membership functions = 11								
2000	6.1	4.6	14.3	2.2	5.3	2	11.3	6.5

Table 5 – Percentage of unrecognized emotions, when the feature vector includes the coordinates of 35 points

Number of learning epochs	Basic emotions							The average percentage of unrecognized emotions
	Anger	Disgust	Fear	Happiness	Sorrow	Surprise	Neutral	
Number of membership functions = 7								
10	63.3	75.5	65.2	91.4	84.4	94.7	64	77.5
30	53	65.3	39.4	85.7	62.2	78.9	36	68.8
50	51	53.1	30.3	62.9	51.1	63.2	30	62.5
70	48.9	44.9	27.3	65.7	48.9	52.6	26	51.3
150	36.7	30.6	18.2	51.4	33.3	42.1	24	40
2000	18.4	12.2	6.1	20	6.7	15.8	4	11.3
Number of membership functions = 9								
2000	2	4.6	8.6	4.4	5.3	0	6.3	4.4
Number of membership functions = 11								
2000	2	0	5.7	4.4	0	0	0	1.7

Table 6 – Percentage of unrecognized emotions, when the feature vector consists of the parameters D1, ..., D24 and K1, ..., K14

Number of learning epochs	Basic emotions							The average percentage of unrecognized emotions
	Anger	Disgust	Fear	Happiness	Sorrow	Surprise	Neutral	
Number of membership functions = 11								
2000	26.5	6.1	28.6	22.2	15.8	4	15	16.9
4000	24.5	3	25.7	24.4	15.8	4	11.3	15.5
8000	22.5	3	25.7	17.8	15.8	2	12.5	14.2
Number of membership functions = 13								
8000	18.4	6.1	17.1	11.1	15.8	2	7.5	11.1
Number of membership functions = 17								
8000	12.2	4.6	5.7	4.4	5.3	4	5	5.9
Number of membership functions = 21								
8000	4.1	1.5	5.7	0	0	0	1.3	1.8

## 6 DISCUSSION

The experiment shows that for the system of emotions recognition based on a multidimensional expanded neo-fuzzy neuron, there is not enough a feature vector that includes only the coordinates of four characteristic points (position of the eyes, nose and mouth). The recognition accuracy cannot be increased by increasing the training duration or the nonlinear synapses number. Obviously, it would be easier to find four characteristic features in the video sequence using automatic detectors, but for the clustering system will not be enough of these data.

On the other hand, the use of the distances between facial features to describe the facial expression dynamics leads to the need to increase the duration of the system training. Since these distances are parameters that are to be calculated from the coordinates (that is, in fact, they are indirect data), their application for automatic analysis of facial expressions is less preferable.

As a result of the experiment, a high efficiency of vector consisting of 35 characteristic points coordinates was established for automatic recognition of facial expressions using a system based on a multidimensional expanded neo-fuzzy neuron. High clustering accuracy is achieved already for 2000 learning epochs with 11 nonlinear synapses in the neuron structure. The reduction of the feature vector dimension (up to 22 points) leads to an increase of training duration and can be compensated by increasing the synapses number to 21. Let's note here, that the number of training samples does not exceed 1000. This is quite important for many practical problems, when

training data of tens and hundreds of thousands of samples are inaccessible.

## CONCLUSIONS

The problem of basic human emotions automatic recognition in real time is considered. For this task, a multidimensional extended neo-fuzzy neuron is used.

**The scientific novelty** of the work lies in the fact that the feature vector dimensionality is determined for the proposed computational architecture. It can include both the coordinates of individual characteristic points, and the distances between them. These data vector correlate well with the well-known facial action coding system. The system learning rate can be increased by increasing the number of neo-fuzzy neuron membership functions. Detection of characteristic points in the automatic recognition system can be realized in real time using standard detectors. Despite the small volumes of the learning sample set, the system provides high recognition accuracy. This factor is especially important in such practical applications, where it is not always possible to obtain thousands and tens of thousands of training samples.

**The practical significance** of obtained results is that the size of the feature vector has been found that provides high accuracy of automatic recognition of basic emotions from 2D images. The experimental results allow using this vector to study the dynamics of facial expression by video.



**Prospects for further research** are to study the recognition of subtle facial expressions (fatigue, pain, uncertainty) using a system based on a multidimensional neo-fuzzy neuron.

#### REFERENCES

1. Ekman P., Friesen W. V. Facial Action Coding System. Palo Alto, USA, Consulting Psychologist Press, 1978.
2. Ekman P., Friesen W. V., Hager J. C. Facial Action Coding System, A Human Face. Salt Lake City, USA, 2002.
3. Pantic M., Bartlett M. S. eds: K. Delac and M. Grgic Machine Analysis of Facial Expressions, Face Recognition, InTech Education and Publishing, 2007, Available from: [http://www.intechopen.com/books/face\\_recognition/machine\\_analysis\\_of\\_facial\\_expressions](http://www.intechopen.com/books/face_recognition/machine_analysis_of_facial_expressions).
4. eds: S. Z. Li, A. K. Jain. Second edition Handbook of Face Recognition. London, Springer-Verlag, 2011, 695 p.
5. Kryvonos Ju. G., Krak Ju. V., Jefimov G. M. ta in. Modeljuvannja ta analiz mimichnyh projaviv emocij, Dopovidi NAN Ukraïny, 2008, No. 12, pp. 51–55.
6. Gokturk S. B., Bouquet J. Y., Tomasi C., Girod B. Model-based face tracking for view independent facial expression recognition, *Proc. IEEE Int'l Conf. Face and Gesture Recognition*, 2002, pp. 272–278.
7. Chang Y., Hu C., Feris R., Turk M. Manifold based analysis of facial expression, *J. Image & Vision Computing*, 2006, Vol. 24, No. 6, pp. 605–614.
8. Pantic M., Rothkrantz L.J.M. Facial action recognition for facial expression analysis from static face images, *IEEE Trans. on Systems, Man and Cybernetics*, 2004, Part B, Vol. 34, No. 3, pp. 1449–1461.
9. Tian Y. L., Kanade T., Cohn J. F.; eds Li S. Z., Jain A. K. Facial Expression Analysis. Handbook of Face Recognition. New York, Springer, 2005, pp. 247–276.
10. Zhang Y., Ji Q. Active and dynamic information fusion for facial expression understanding from image sequence, *IEEE Trans. Pattern Analysis & Machine Intelligence*, 2005, Vol. 27, No. 5, pp. 699–714.
11. Bartlett M. S., Littlewort G., Frank M. G., Lainscsek C., Fasel I., Movellan J. Fully automatic facial action recognition in spontaneous behavior, *Proc. IEEE Conf. Automatic Face & Gesture Recognition*, 2006, pp. 223–230.
12. Guo G., Dyer C. R. Learning from Examples in the Small Sample Case – Face Expression Recognition, *IEEE Trans. Systems, Man, and Cybernetics*, 2005, Part B, Vol. 35, No. 3, pp. 477–488.
13. Anderson K., McOwan P. W. A Real-Time Automated System for Recognition of Human Facial Expressions, *IEEE Trans. Systems, Man, and Cybernetics*, 2006, Part B, Vol. 36, No. 1, pp. 96–105.
14. Hu Z., Bodyanskiy Ye. V., Kulishova N. Ye., Tyshchenko O. K. A Multidimensional Extended Neo-Fuzzy Neuron for Facial Expression Recognition, *International Journal of Intelligent Systems and Applications (IJISA)*, 2017, Vol. 9, No. 9, pp. 29–36.
15. Bay H., Ess A., Tuytelaars T., Van L. Gool SURF: Speeded up robust features, *Computer Vision and Image Understanding*, 2008, Vol. 110, pp. 346–359.
16. Shi J., Tomasi C. Good Features to Track, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 593–600.
17. [http://pics.psych.stir.ac.uk/2D\\_face\\_sets.htm](http://pics.psych.stir.ac.uk/2D_face_sets.htm)
18. Lucey P., Cohn J. F., Kanade T., Saragih J., Ambadar Z., Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression, *Proceedings of IEEE workshop on CVPR for Human Communicative Behavior Analysis*, San Francisco, USA, 2010.

Received 08.05.2018.

Accepted 27.05.2018.

УДК 004.825:004.932.72'1

#### ФОРМУВАННЯ ВЕКТОРУ ХАРАКТЕРНИХ ОЗНАК ДЛЯ РОЗПІЗНАВАННЯ ВИРАЗІВ ОБЛИЧЧЯ ЗА ДОПОМОГОЮ НЕО-ФАЗЗИ СИСТЕМИ

**Бодяньський Є. В.** – д-р техн. наук, професор, керівник Проблемної науково-дослідної лабораторії автоматизованих систем управління, професор кафедри штучного інтелекту Харківського національного університету радіоелектроніки, Харків, Україна.

**Кулішова Н. Є.** – канд. техн. наук, доцент, професор кафедри медіасистем та технологій Харківського національного університету радіоелектроніки, Харків, Україна.

**Ткаченко В. П.** – канд. техн. наук, професор, завідувач кафедри медіасистем та технологій Харківського національного університету радіоелектроніки, Харків, Україна.

#### АНОТАЦІЯ

**Актуальність.** Стаття присвячена вивченню проблеми формування набору навчальних даних для системи автоматичного розпізнавання емоцій людини на основі багатовимірного розширеного нео-фаззи нейрона. Розглядаються аспекти вибору розмірності і складу вектору ознак, їх впливу на швидкість навчання системи.

Об'єктом дослідження є метод кластеризації багатовимірних даних. Предмет дослідження – систематизація геометричних ознак двовимірних зображень.

Основна мета роботи – розробка підходу до опису виразу обличчя людини за допомогою фіксованого набору геометричних ознак, які можуть бути отримані в ході обробки кадрів відеоряду.

**Метод.** Для навчання системи розпізнавання виразів обличчя пропонується утворити вектор ознак, що складається з координат характерних точок. Вибрані такі точки, які пов'язані з розташуванням і формою зіниць очей, контурами губ, повік, брів, крил носа, носогубних складок. Подібні точки досить легко можна виділяти в автоматичному режимі обробки зображень за допомогою відомих контурних детекторів. Також розглянуто можливість використання для опису вираження обличчя не координат характерних точок, а відстаней між ними. З цих відстаней створено інший вектор ознак, властивості якого було порівняно з властивостями вектору з координат точок.

**Результати.** Розроблена система розпізнавання на базі багатовимірного розширеного нео-фаззи нейрона була реалізована у програмному забезпеченні та досліджена для вирішення проблеми класифікації виразів обличчя. Зроблено порівняння векторів атрибутів, що відрізняються за складом та розмірами. Обрано таку структуру вектору ознак, що забезпечує високу швидкість навчання системи, та не вимагає введення додаткових структурних елементів.

**Висновки.** Експериментальне дослідження повністю підтверджує ефективність розробленого підходу для розпізнавання виразів обличчя людини з використанням багатовимірного нео-фаззи нейрона.

**КЛЮЧОВІ СЛОВА:** розпізнавання виразів обличчя, характерні ознаки, нео-фаззи нейрон, функція приналежності, нечітка кластеризація, обчислювальний інтелект.

УДК 004.825:004.932.72'1

## ФОРМИРОВАНИЕ ВЕКТОРА ХАРАКТЕРНЫХ ПРИЗНАКОВ ДЛЯ РАСПОЗНАВАНИЯ ВЫРАЖЕНИЙ ЛИЦА С ПОМОЩЬЮ НЕО-ФАЗЗИ СИСТЕМЫ

**Бодянский Е. В.** – д-р техн. наук, профессор, руководитель Проблемной научно-исследовательской лаборатории автоматизированных систем управления, профессор кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники, Харьков, Украина.

**Кулишова Н. Е.** – канд. техн. наук, доцент, профессор кафедры медиасистем и технологий Харьковского национального университета радиоэлектроники, Харьков, Украина.

**Ткаченко В. Ф.** – канд. техн. наук, профессор, заведующий кафедрой медиасистем и технологий Харьковского национального университета радиоэлектроники, Харьков, Украина.

### АННОТАЦИЯ

**Актуальность.** Статья посвящена изучению проблемы формирования набора обучающих данных для системы автоматического распознавания эмоций человека на основе многомерного расширенного нео-фаззи нейрона. Рассматриваются аспекты выбора размерности и состава вектора признаков, их влияния на скорость обучения системы.

Объектом исследования является метод кластеризации многомерных данных. Предмет исследования – систематизация геометрических признаков двумерных изображений.

Основная цель работы – разработка подхода к описанию выражения лица человека с помощью фиксированного набора геометрических признаков, которые могут быть получены в ходе обработки кадров видеоряда.

**Метод.** Для обучения системы распознавания выражений лица предлагается создать вектор признаков, состоящий из координат характерных точек. Выбраны такие точки, которые связаны с расположением и формой зрачков глаз, контур губ, век, бровей, крыльев носа, носогубных складок. Подобные точки достаточно легко можно выделять в автоматическом режиме обработки изображений с помощью известных контурных детекторов. Также рассмотрена возможность использования для описания выражения лица не координат характерных точек, а расстояний между ними. Из этих расстояний создан другой вектор признаков, свойства которого были сравнены со свойствами вектора из координат точек.

**Результаты.** Разработанная система распознавания на основе многомерного расширенного нео-фаззи нейрона реализована в программном обеспечении и исследована для решения проблемы классификации выражений лица. Выполнено сравнение между векторами атрибутов, которые отличаются по составу и размерности. Выбрана такая структура вектора признаков, которая обеспечивает высокую скорость обучения системы, и не требует введения дополнительных структурных элементов.

**Выводы.** Экспериментальное исследование полностью подтверждает эффективность разработанного подхода для распознавания выражений лица человека с использованием многомерного нео-фаззи нейрона.

**КЛЮЧЕВЫЕ СЛОВА:** распознавание выражений лица, характерные признаки, нео-фаззи нейрон, функция принадлежности, нечеткая кластеризация, вычислительный интеллект.

### ЛІТЕРАТУРА / LITERATURA

- Ekman P. Facial Action Coding System / P. Ekman, W. V. Friesen. – Palo Alto, USA : Consulting Psychologist Press, 1978.
- Ekman P. Facial Action Coding System, A Human Face / P. Ekman, W. V. Friesen, J. C. Hager. – Salt Lake City, USA, 2002.
- Pantic M., Bartlett M. S. Machine Analysis of Facial Expressions, Face Recognition/ eds: K. Delac and M. Grgic. – InTech Education and Publishing, 2007. – Available from: [http://www.intechopen.com/books/face\\_recognition/machine\\_analysis\\_of\\_facial\\_expressions](http://www.intechopen.com/books/face_recognition/machine_analysis_of_facial_expressions).
- Handbook of Face Recognition / eds: S.Z. Li, A. K. Jain. Second edition. – London: Springer-Verlag, 2011. – 695 p.
- Кривонос Ю. Г. Моделювання та аналіз мимічних проявів емоцій / Ю. Г. Кривонос, Ю. В. Крак, Г. М. Єфімов та ін. // Доповіді НАН України. – 2008. – № 12. – С. 51–55.
- Gokturk S. B. Model-based face tracking for view independent facial expression recognition / S.B. Gokturk, J. Y. Bouguet, C. Tomasi, B. Girod // Proc. IEEE Int'l Conf. Face and Gesture Recognition. – 2002. – P. 272–278.
- Chang, Y. Manifold based analysis of facial expression / Y. Chang, C. Hu, R. Feris, M. Turk // J. Image & Vision Computing. – 2006. – Vol. 24, No. 6. – P. 605–614.
- Pantic M. Facial action recognition for facial expression analysis from static face images / M. Pantic, L. J. M. Rothkrantz // IEEE Trans. on Systems, Man and Cybernetics. – 2004. – Part B, Vol. 34, No. 3. – P. 1449–1461.
- Facial Expression Analysis. Handbook of Face Recognition / [Y. L. Tian, T. Kanade, J. F. Cohn]; eds Li S. Z., Jain A. K. – New York : Springer, 2005. – P. 247–276.
- Zhang Y. Active and dynamic information fusion for facial expression understanding from image sequence/ Y. Zhang, Q. Ji // IEEE Trans. Pattern Analysis & Machine Intelligence. – 2005. – Vol. 27, No. 5. – P. 699–714.
- Bartlett M. S. Fully automatic facial action recognition in spontaneous behavior / M. S. Bartlett, G. Littlewort, M. G. Frank, C. Lainscsek, I. Fasel, J. Movellan // Proc. IEEE Conf. Automatic Face & Gesture Recognition. – 2006. – P. 223–230.
- Guo G. Learning from Examples in the Small Sample Case – Face Expression Recognition / G. Guo, C.R. Dyer // IEEE Trans. Systems, Man, and Cybernetics. – 2005. – Part B, Vol. 35, No. 3. – P. 477–488.
- Anderson K. A Real-Time Automated System for Recognition of Human Facial Expressions / K. Anderson, P.W. McOwan // IEEE Trans. Systems, Man, and Cybernetics. – 2006. – Part B, Vol. 36, No. 1. – P. 96–105.
- Hu Z. A Multidimensional Extended Neo-Fuzzy Neuron for Facial Expression Recognition / Z. Hu, Ye. V. Bodyanskiy, N. Ye. Kulishova, O. K. Tyshchenko // International Journal of Intelligent Systems and Applications (IJISA). – 2017. – Vol. 9, No. 9. – P. 29–36.
- Bay H. SURF: Speeded up robust features / H. Bay, A. Ess, T. Tuytelaars, L. Van Gool // Computer Vision and Image Understanding. – 2008. – Vol. 110. – P. 346–359.
- Shi J. Good Features to Track / J. Shi, C. Tomasi // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 1994. – P. 593–600.
- [http://pics.psych.stir.ac.uk/2D\\_face\\_sets.htm](http://pics.psych.stir.ac.uk/2D_face_sets.htm)
- Lucey P. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression / P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews // Proceedings of IEEE workshop on CVPR for Human Communicative Behavior Analysis, San Francisco, USA. – 2010.