NEURAL-LIKE GROWING NETWORKS IN INTELLIGENT SYSTEM OF RECOGNITION OF IMAGES

Vitaliy Yashchenko

Abstract: *The neural-like growing networks used in the intelligent system of recognition of images are under consideration in this paper. All operations made over the image on a pre-design stage and also classification and storage of the information about the images and their further identification are made extremely by mechanisms of neural-like networks without usage of complex algorithms requiring considerable volumes of calculus. At the conforming hardware support the neural network methods allow considerably to increase the effectiveness of the solution of the given class of problems, saving a high accuracy of result and high level of response, both in a mode of training, and in a mode of identification.*

Keywords: *Neural-like networks, images recognition.*

Introduction

The need in intellectual autonomous macro and micro robots, which can replace the man in environment, unacceptable and dangerous to his health and threat of the terrorist acts, steadily grows. The complex control systems are used for the decision of a complex of tasks on management of robots in extreme conditions. To correct orientation in the surrounded environment a robot must "understand" this environment and have its model. One can say that "consciousness" of robot is its ability to reflect a surrounding world in its memory, to model it in the process of its activity. Exactly this characteristic distinguishes intellectual robots from robots of previous generations, whose actions are done "unconsciously", subjected to strict programs.

The solution of such problems by traditional methods of mathematical programming frequently oriented on computer facilities with the consecutive architecture is integrated to temporary costs unacceptable for many applications.

Nowadays, the next rise of intensity of researches in the field of artificial neural networks is observed. The volume of financing of the projects using the technologies of neural networks in Japan, China, USA and Europe is about hundred million dollars.

In the given article we represent the new class of neural networks and its usage in the system of recognition of images.

Neural-like growing networks

The neural-like growing networks are the new class of neural networks, which are designed within the frameworks of bionic approach on the basis of technologies integration of a data processing in growing semantic and neural networks.

Neural-like growing networks (n-GN) were created specially for system engineering of artificial intelligence.

Multidimensional receptor - effector neural-like growing network (ìren-GN) is actually a model of a brain of the man. In them, during functioning (life cycle) of system, the information on the external world accumulates. Simultaneously the own structure of a network is formed at the expense of "birth" (occurrence) of neurons and occurrence and disappearance of connections between them. Actually this structure is a model of the external world in the system. Therefore the class neural-like of growing networks has a wide spectrum of application. We have applied n-GN in the system of recognition of images, about which the speech will go further.

The neural-like growing network is a set of interdependent neural-like elements intended for a reception and transformation of information and in the process of reception of information the network is increasing in size.

The neural-like growing network (n-GN) can be submitted as a non-cyclic digraph with fluidized connections, in which neural-like elements are the tops. The tops which do not have entering arcs, are called the receptors, the rest of them are the neural-like elements.

Depending on the applied area, in which n-GN are used, the receptor can introduce the character, sign, parameter, meaning of physical or economical index, elementary fact from the description of a situation, etc. The neural-like elements correspond to the descriptions of visual images, words, phrases, subjects, objects, processes, plans, situations or phenomena.

The specific peculiarity of n-GN is the capability to adapt for these descriptions, changing the structure accordingly. The adaptation is accompanied by input in the network of new tops and arcs at transition of any group of receptors and neural-like elements in a condition of excitation. The class of neural-like growing networks consists of neural-like growing (receptor) networks themselves, multidimensional neural-like growing (receptor) networks, receptor and effective neural-like growing networks and multidimensional receptor and effective neural-like growing networks.
Formally they are set in the following way.

Neural-like growing (receptor) networks (n-GN) $S = (R, A, D, M, P, N)$, where $R = \{r_i\}$, $i = \overline{1,n}$ - final set of receptors; $A = \{a_i\}$, $i = \overline{1,k}$ - final set of neural-like elements; $D = \{d_i\}$, $i = \overline{1,e}$, - final set of arcs linking receptors with neural-like elements and neural-like elements between themselves; $P=\{P_i\}$, $i = \overline{1,k}$ N = h, where P - threshold of excitation of the top a_i, $P = f(m_i) > P_0$ (P_0 - the minimum permissible threshold of excitation) provided that to the set of arcs D, coming on top a_i, corresponds to the set of weighting coefficients $M = \{m_i\}$, $i = \overline{1,w}$, and m_i can receive both positive and negative meanings.

Multidimensional neural-like growing (receptor) networks (mn-GN) $S = (R, A, D, P, N)$, where $R \supset R_l, R_r, R_v$, $A \supset A_l, A_r, A_v$, $D \supset D_l, D_r, D_v$, $N = \{n_l, n_r, n_v\}$, here R_l, R_r, R_v - final subset of receptors, A_l, A_r, A_v - final subset of neural-like growing elements, D_l, D_r, D_v - final subset of arcs, P_l, P_r, P_v - final set of thresholds of excitation of neural-like elements, that belong, for example, to linguistic, voice or visual information spaces. N - final set of floating factors of cohesion.

Receptor and effective neural-like growing networks (ren-GN) $S = (R, A_r, D_r, P_r, M_r, N_r, A_e, D_e, P_e, M_e, E, N_e)$, where $R = \{r_i\}$ - final set of receptors, $Ar = \{a_{ri}\}$, - final set of neural-like elements of the receptor zone, $Dr = \{d_{ri}\}$, - final set of arcs of a receptor zone, $E = \{e_i\}$, - final set of effectors, $Ae = \{a_{ei}\}$, - final set of neural-like elements of effecting zone, $D_e = \{d_{ei}\}$, - final set of arcs of effecting zone, N_r, N_e - final set of floating factors of cohesion receptor and effecting zones accordingly, P_r, P_e - the threshold of excitation of tops a_{ri}, a_{ei}, $M = \{m_i\}$ is a set of weighting coefficients, besides m_i can receive both positive and negative meanings.

Multidimensional receptor and effecting neural-like growing networks (mren-GN) $S = (R, A_r, D_r, P_r, M_r, N_r, A_e, D_e, P_e, M_e, E, N_e)$, $R \supset R_v, R_s, R_t$, $A \supset A_v, A_s, A_t$, $D \supset D_v, D_s, D_t$, $P \supset P_v, P_s, P_t$, $M \supset M_v, M_s, M_t$, $N \supset N_v, N_s, N_t$, $E \supset E_r, E_{d1}, E_{dn}$, $A \supset A_r, A_{d1}, A_{dn}$, $D \supset D_r, D_{d1}, D_{dn}$, $P \supset P_r, P_{d1}, P_{dn}$, $M \supset M_r, M_{d1}, M_{dn}$, $N \supset N_r, N_{d1}, N_{dn}$ here R_v, R_s, R_t - final subset of receptors, A_v, A_s, A_t - final subset of neural-like elements of receptor zone, D_v, D_s, D_t - final subset of arcs of receptor zone, P_v, P_s, P_t - final set of thresholds of excitation of neural-like elements of receptor zone, that belongs, for example, to visual, sound, tactile information spaces, N_r- final set of float factors of cohesion of receptor zone, E_r, E_{d1}, E_{dn} - final subset of effectors, A_r, A_{d1}, A_{dn} - final subset of neural-like elements of effecting zone, D_r, D_{d1}, D_{dn} - final subset of arcs of effecting zone, P_r, P_{d1}, P_{dn}- final set of thresholds of excitation of neural-like elements of effecting zone, that belongs, for example, to voice information space and space of operating, N_e - final set of floating factors of cohesion of effecting zone.

The neural-like growing networks are the dynamic structure, which changes depending upon the meaning and time of reception of the information for receptors, and also for the previous condition of the network. In it the information on objects is represented by ensembles of exited tops and connections between them. The storage of the descriptions of objects and situations is accompanied by input in the network of new tops and arcs at transition of any group of receptors and neural-like elements in condition of excitation.
Thus, the combination of stable relationships of depicted object ensuring its integrity and identity to itself, i.e. the saving of the basic characteristics is formalized.

Information about objects and their classes is represented by ensembles of the associatively interconnected tops distributed on the structure of the net. The input of the new information in the network causes the process of its structure constructing (redistribution of the connections between already existing and again arising tops) with simultaneous excitation of the neural elements. In the result of this process the inclusion of the described object into the class, to which it belongs, is going on, or the new class of the objects is formed. So the classification and choosing the common attributes of the objects is carried out. Algorithm of the network construction establishes automatically the associative connections between descriptions of the objects accordingly their attributes. The description of the object or the class of the objects is located in some part of the network that lets to carry out various operations of associative search effectively. Profitability of the information representation in n-GN is carried out owing to compression of the information on each its level and representation of the identical combinations of attributes of several objects by one common subset of tops of a network. The training of the network is carried out simultaneously with their construction according to rules of construction and functioning of a network.

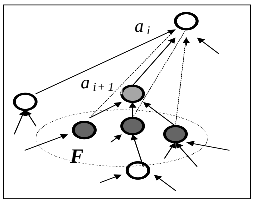

Figure1.Rule 1.

Rule 1. If during the perception of information, a subset of tops *F* from the set of tops, having direct relationship with the top a_i, is excited, and $\overline{\overline{F}} \geq h$, the relationships of a top a_i with tops from the subset *F* are liquidated and a new top a_{i+1} joins the network, whose entries are connected with entries of all tops of the subset *F*, and the exit of a top a_{i+1} is connected with one of the inputs of a top a_i, whereas the input relationships of the top a_{i+1} are assigned weighted factors m_i, corresponding to the weighted factors of liquidated relationships of the top a_i and top a_{i+1} is assigned the threshold of excitation P_i, which equals $f(m_i)$, (function from weighted relationship factors, which fall into the top a_{i+1}). Outcoming relationship of this top is assigned a weighted factor m_i , which equal $f(P_i)$. Relationships, outcoming from receptors, are assigned a weighted factor,and $f(b_i)$, function from the code of sign b_i, correspond to a given receptor (fig.1).

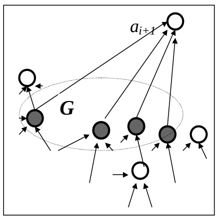

Figure 2. Rule 2.

Rule 2. If during the perception of information, a subset of tops *G* is excited, and $\overline{\overline{G}} \geq h$ a new associative top a_{i+1}, joins the network, which is connected by turning arcs with all tops of the subset *G*. Each of turning arcs is assigned a weighted factor m_i, equal $f(P_i)$ of a corresponding top from the subset *G*, and a new top a_{i+1} is assigned a minimum threshold of excitement P_i, equal to the function of weighted factors m_i of incoming arcs (fig.2).

Example of construction neural-like growing networks

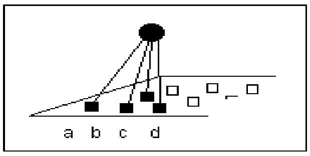

Figure3. Image 1

The principle of constructing n-GN (for simplicity of perception) will be looked at the example of constructing the multiconnection growing network. Formally m-GN is described so: *S=(R, A, D, N).*

Let be learning access, which consists of k-images : *1. a,b,c,d; 2. b,c,d,e,g,h; 3.d,e,f; ... k. d,e,h.*

Let's set up variable coefficient of connectivity $N \geq 3$. In this case when entering the description of the first image (a,b,c,d) on the receptor field, the receptors 1,2,3,4 are changed over to the state of excitation. The vertex *a,b,c,d* is formed and the connections between vertex and excited receptors are set up (fig.3.). The vertex is changed over to the state of excitation. In a definite time the excitation is taken of from receptors and a vertex.

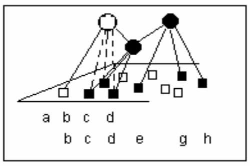

Figure4. Image 2

When entering the description of the second image (*b,c,d,e,g,h*) on the receptor's field *R*, the receptors 2,3,4,5,7,8 are changed over to the state of excitation. The number of signs, coincided with the description of the first image (*b,c,d*)=3, then N=3, the vertex (*b,c,d*) and (*b,c,d,e,g,h*) are formed. The connection of the vertex *a,b,c,d* with receptors 2,3,4 are liquidated. Inputs of the vertex *b,c,d* are connected with receptors 2,3,4 and outputs of this vertex are connected with inputs of the vertex (*a,b,c,d*) and *b,c,d,e,g,h*, and these vertices are changed over to the state excitation (fig.4).In a definite time the excitation is taken off from the vertex (b,c,d) (b,c,d,e,g,h) and receptors.

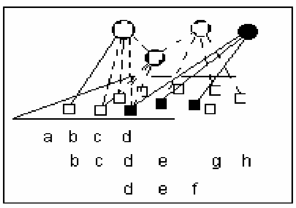

Figure5. Image 3

When entering the description of the third-image on the receptor's field, the new vertex (d,e,f) is formed (fig.5).

When entering the description of k-image on the receptor's field, the new vertex (d,e,k) is formed (fig.6).

In this case the separation of the common signs, described notions are performed.

Thus the description of the notion (the vertex of the network) and signs are stored. Besides, the information, which enter the receptor's fields of the networks, is classified and structured automatically .

When forming new vertex in n-GN, weight coefficients of connection m_i and thresholds of the excitation of the vertex P_i are considered, that is, constructing n-GN is performed analogically with building m-GN, but in accordance with rules, which are described in the materials presented before.

Information in neural-like growing networks is stored as a result of its reflecting in the structure of a network. New information input into the network brings about a process of building its structure.

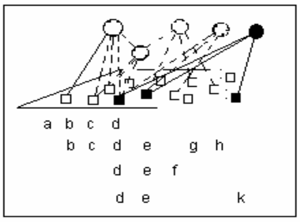

Figure6. Image k

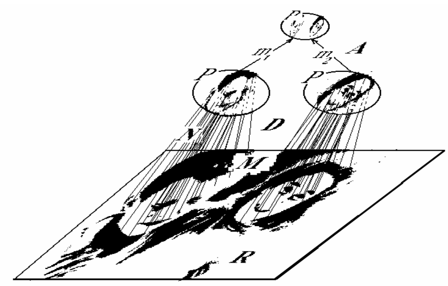

Figure7. Storing the images

Neural-like growing networks are a dynamic structure, which changes depending on values and time of arrivals of image to receptors, as well as former condition of the network. Storing the images descriptions is accompanied by input in to the network of new tops and arcs when turning a group of receptors and neural-like elements became excited (fig.7). The process of excitation spreads on the network, as a wave [1,2].

System of identification of the person

So, first of all there is a problem of detection and allocation of the face on the image.

The program of detection of a face searches for it, sequentially scanning the image in all possible scales, since some minimum size of area (in pixels).

For recognition of faces the different methods are used.

The method of matching with a template is based on matching of the images with the usage of a mask, filter and so on. This method is one of the most simple methods, but it is not steady and responds to the noise.

The method of main components. In the given approach the two-measured image of the face is considered as a vector. If the image width *w* and altitude *h* pixels then, the number of components of this vector will be w*h. In the method the main components the integrated change in a set of faces and this change of several coordinates in new basis is described. Thus the dimension decreases. In the method of the main component statistical image processing will be used [3].

The method of "Distinctive features of the face" (Eigenface Technology). This method uses a set of distinctive features of faces, which represent a combination of light and dark areas. Such a set is formed by means of combination of all images with allocation of individuals having definite similar tags in separate groups. The distinctive features of each face act the role of the pieces of a face. For identification of a face the program compares its distinctive features, which are joint in so-called template of the face with templates of faces from the database, selecting those that coincide most of all [4-6].

Method of the analysis of local elements of the face (Local Feature Analysis). The given method goes from the previous one, but allows to get rid of such problems as sensitivity to deformation of the face, to its relative position, and also to the degree of illuminating intensity. The method of analysis of local elements of the face attracts the attention to concrete details of features of the face, instead of its common view. The method selects the elements, which fully determine each face. These elements play the role of building units, from which all images of the faces can be constructed [5-8].

There are also methods that perform the operations described above. So, the Miros company in its TrueFace ID technology combines the methods *of distinctive features and analysis of local elements of the face* realized in paradigm of artificial neural networks.

The suggested method is based on bionic approach, in which the data processing is realized by means of the neural network.

In bionic approach it is supposed that each conditional point of the image, accepted by man, corresponds to one neuron. Each such a neuron has some dendrites connected with adjacent neurons. The signal from the point of the image, corresponding to the neuron, goes in it through the receptor, strengthening by a positive weighting coefficient of central dendrite, and the signals corresponding to adjacent pixels of the images, which go through receptors, are braked by negative weighting coefficients of lateral dendrites.

Going into neuron the signals multiplied on the corresponding weighting coefficients are summarized and is moving on the output.

On the output of neural network the sequence of numbers takes the place, which according to definite law corresponds to the meanings of codes of colours of input points of the image.

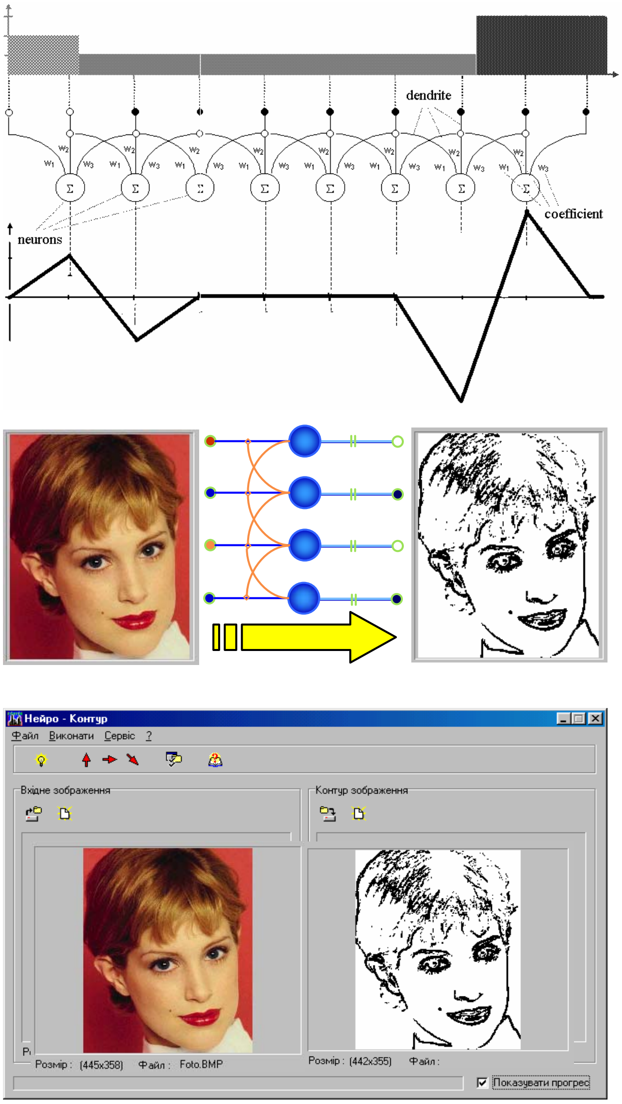

Figure8. Contour of the image

If to figure the data received on an output in the form of schedule, in some places it is possible to note sharp differences of the obtained function, which responds to sharp gangs of colour. Having fixed these differences it is possible to compare their size with the given threshold. The excess of this size of the threshold testifies of availability of a contour of the image in the given point (fig. 8).

Thus, the abstracting from colour of the image, illuminating intensity, and superfluous background "noise" is carried out. It allows to present object for classification more precisely, having deprived the image of elements, which do not bear any information load. And it also allows to reduce a volume of the information about object which is necessary for classification.

The problem of allocation of contours on the image has been solving by classic algorithms for a long time. They are: spatial differentiation, functional approximation, a high frequency filtration.

There is a tendency to esteem the border as the area of a sharp difference of a function of picture level f (x, y) that is general one for all these methods. But they are alike in their mathematical model of the notion edge and search algorithm of edge points [5-8].

There is one main disadvantage of all these algorithms; it a rather low speed of activity, which will not be enough for image processing of the large size at a high frequency of their reception.

In the bionic approach, due to full parallelism of processing of each point of the image, at a hardware representation, the fast enough response time of processing of the input images is reached.

Internal representation of the image

For creation of the internal representation the image, which has passed prior processing, is braked into definite pieces, which are the characteristic peculiarities of the given face. The quantity of such pieces can be changed, selecting optimal.

Then the abstracting from the sizes of the entering image and seal of the obtained data of each piece is made, that allows to get rid of the unnecessary or unessential information. Then the obtained packed data are integrated in a so-called internal form of the image of the face.

The size of one such representation depends on quantity of fragments. The conducted researches have shown that internally the representation with optimal for further recognition by quantity of units takes no more than 25-100 bytes. So, for 100 000 images it is enough less than 2,4 - 100 Mbytes of a disk space.

Training and identification of the images in neural-like growing network

While training, internally the representation of the image containing the special features of the face is moved on receptors of neural-like growing network. Then, the excitation is transmitted on neural-like network elements, which correspond to the most similar faces. Further on, according to the rules of formation of neural-like growing network, the exited elements are integrated in one that corresponds to this representation.

Thus, there a case of learning, classification and accumulation of the obtained information process is observed. And the structure of neural-like growing network is built in such a way that the accumulation of any duality of the existing information is eliminated. At the stage of identification, on the input of neural-like growing network the internal representation of the identified image is moving. As a result, definite neural-like elements corresponding to the most eligible images that are accumulated in the network are excited.

According to the degree of excitation of elements the factor of similarity is determined, which in case of excess of the predetermined threshold value indicates a positive result of identification (fig.9).

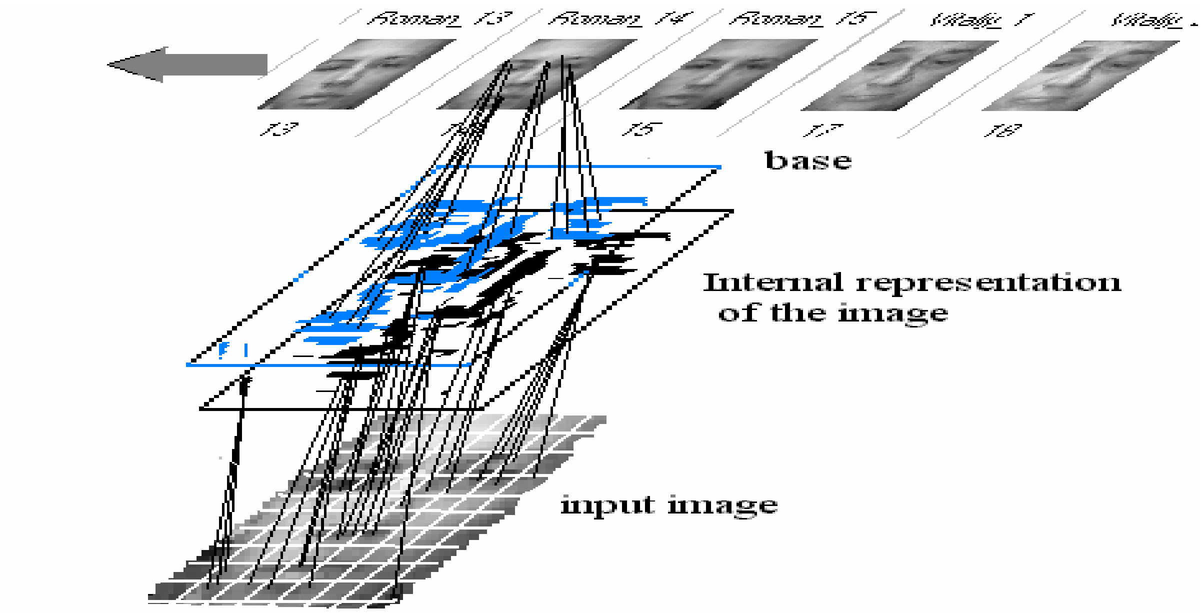

Figure9. Factor of similarity.

At hardware representation of neural-like growing network, where the information went on each receptor is processed in parallel way, the classification descends immediately.

Program implementation

The program system (fig.10) is designed in Borland C ++ Builder 4.0 under platforms Windows 95/98/NT/2000.

System decides a problem of identification of the person according to the image of his face, having the database of the stored images. The images of faces on a program input are moved by means of the digital video camera, TV and also from the file of the beforehand recorded image or from a clipboard of Windows system.

On the stage of prior processing, as it was described above, the problem of allocation of contours of the image is decided, using the bionic approach. The quality of further recognition fully depends on quality of the image preparation on this stage. So a flexible system of parameters selecting of allocation of contours is realized there.

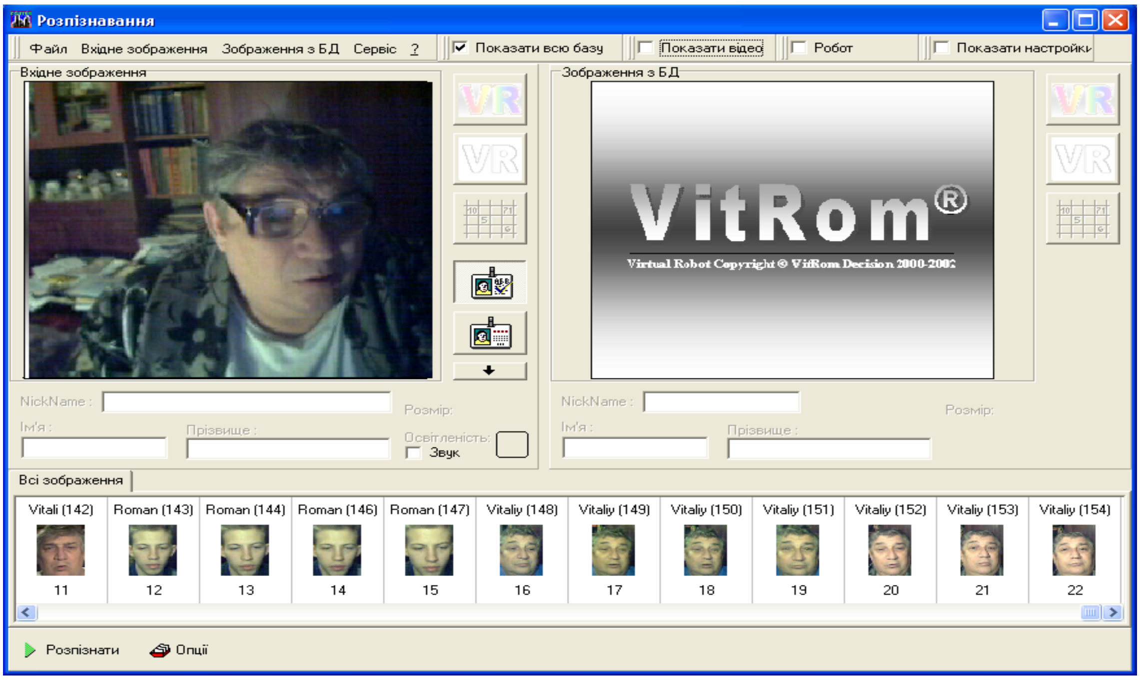

Figure 10. The program system

Conclusions

One of the most complex problems, that faces specialists of the artificial intelligence, is a problem of perception of information and organization of intellectual system or robot behavior.

Considered here types of neural-like growing networks allow to the system, containing n-GN structure, in the process of its activity to form in it a model of external environment and work out adequate actions.

The neural-like growing networks give an opportunity to form the notions as the objects and relations among them when constructing the network itself. For this each sense (notion) gets a separate component of the network as an vertex connected with the other vertexes. In general, this fully corresponds to the structure reflect able in the brain, where each explicit notion has represented as a definite structure and has its denoting symbol. Thus, the neural-like growing network is represented as a convenient apparatus for simulating the mechanisms of purposeful thinking as performing the specific psychophysiological functions.

In comparison with known intellectual systems and robots, behaving according to the preprogrammed functions, the intellectual systems based on the new technology, provide possibility to generate own functions of behavior through the analysis of external information.

The usage of the bionic approach on the stage of prior processing of the image and the technology of data processing in neural-like growing networks allows to reduce considerably the volumes of computing operations. All operations made above the image on the stage of prior processing and also the classification and information storage about the images and their further identification are made only by the mechanisms of neural-like networks without the usage of complex algorithms, which demand a large volume of calculations. By the corresponding hardware support the neural network methods allow to increase considerably an effectiveness of the solution of the given class of problems, saving a high accuracy of result and high level of response, both in the mode of training, and in the mode of identification. The described method can be applied in security systems and different technical systems for recognition of current situations with the purpose of decision making and fulfillment of actions based on the processing of the visual information.

The technologies of recognition of the man are already submitted in the software market.

The company Miros has presented TrueFace ID system of recognition of the person of the man [3]. It is a new product Miros, constructed on the basis of technology TrueFace. The system TrueFace ID is constructed on the basis of the appropriate software, which on the reclame of the developer, allows to identify the person of the man in a database from thousand images in some seconds. TrueFace ID makes comparison of the images of the persons received directly with the help of a videocamera or written down on a videofilm, with stored(kept) in a database, and precisely identifies the man.

The corporation Visionics offers a means of recognition of under the name FaceIt PC [4], which is intended for amplification (strengthening) protection of independent PCs equipped OC Windows 95. The company DCS (Dialog Communication Systems) Inc. represents technology BioID, on the basis of which the identification of the person is made on three biometric parameters: the person, vote, movement of lips [5]. The testing of our system has shown results not making a concession of systems TrueFace ID and BioID. However, taking into account an orientation of new technology – n-GN on mass parallelism, the hardware realization of system will give significant advantages above systems TrueFace ID and BioID

References:

[1] V.A. Yashchenko Neural-like growing networks - new class of the neural networks // Proceedings of the International Conference on Neural Networks and Brain Proceeding, pages 455 -458, Beijing, China, Oct. 27-30' 98.

[2] V.A. Yashchenko Receptor-effector neural-like growing network - an efficient tool for building intelligence systems // Proceedings of the second internatiional conference on information fusion, July 6-8, 1999, Sunnyvale Hilton Inn, Sunnyvale, California, USA, Vol.11, pp. 1113-1118

[3] Technology TmeFace ID of the company Miros. - http: // www.miros.com, 2000

[4] Technology Facelt of the company VISIONICS. - http: // www.faceit.com, 2000

[5] Technology BioID of the company DCS Inc.- http: // www.bioid.com, 2000

[6] Method main компонент.-http:// ww.mechmath. psu.ru, 2000

[7] P. Duda, item Харт Pattern recognition and analysis сцен.- Пер. With engl., under ed. V.L. Stefanuk m. The world, 1976E. Hall Computer Image Processing and Recognition - N.Y.: Academic Press, 1979

[8] K.S. Fu, J. Mu Pattern Recognition, 1981,v.l3.№l

Author information

Vitaliy Yashchenko - Institute of Mathematics of Machine and Systems; Kiev, Ukraine;
e-mail: mis@immsp.kiev.ua

SEGMENTATION OF A SPEECH SIGNAL WITH APPLICATION OF FAST WAVELET TRANSFORMATION

T. Yermolenko

Abstract: the article describes the method of preliminary segmentation of a speech signal with wavelet transformation use, consisting of two stages. At the first stage there is an allocation of sibilants and pauses, at the second – the further segmentation of the rest signal parts.

Key words: wavelet transformation, wavelet coefficients, approximation, segmentation.

Introduction

As known, the speech signal will consist of quasi-stationary parts corresponding to voice and sibilant phonemes, alternated by parts with rather fast changes of signal spectral characteristics (interphoneme transitions, explosive and occlusive phonemes, interword transitions speech - pause). It is possible to say, that the speech signal is characterized by nonlinear fluctuations of various scales. Therefore multireslution analysis and wavelet – transformation is considered to be rather effective for the analysis of a speech signal.

Segmentation of a speech signal (SS) means allocation of signal parts corresponding to separate structural units of SS. Considering phonemes as such units the task of segmentation is reduced to detection of interphoneme transitions. Within the framework of traditional approaches the decision of this task is rather problematic. However WT allows to solve this problem at least for the phonemes corresponding to rather extensive quasi-stationary SS parts. The matter is that on interphoneme transitions the signal undergoes significant changes at once on many research scales, and, accordingly, is characterized by increase of wavelet coefficients for many levels of decomposition while on stationary parts of phonemes wavelet coefficients appear grouped near to the certain scales. Thus, search of interphoneme borders can be reduced to search of moments of wavelet coefficients increase for a significant amount of levels of resolution.