Neurocomputing 4 (1992) 93–102
Elsevier

The model of a neural network visual preprocessor

Ilya A. Rybak, Natalia A. Shevtsova and Vladislav M. Sandler
Institute of Neurocybernetics at Rostov State University, 194/1 Stachka Ave., Rostov-on-Don 344104, USSR

Received 17 May 1991
Revised 14 June 1991

Abstract

Rybak, I.A., N.A. Shevtsova and V.M. Sandler, The model of a neural network visual preprocessor, Neurocomputing 4 (1992) 93–102.

The model of a neural network visual preprocessor and a system architecture for visual information processing are proposed. The model of the preprocessor is based on the model of a visual cortex iso-orientation domain which is considered as a neural network with retinotopically organized afferent inputs and anisotropic lateral inhibition formed by feedback connections via inhibitory interneurons. The high-level system uses the preprocessor to process image fragments with different resolutions and to represent the image as a set of contour segments of different sizes.

Keywords. Visual preprocessor; visual information processing; inhibitory interneurons.

1. Introduction

The major consumption of computer time during image analysis by a sequential computer occurs during the primary processing step which is characterized by a large volume of processed information and cumbersome computational procedures. Visual information processing may be accelerated by the implementation of primary processing procedures in a parallel way and by preliminary compression of visual information. However, parallelism in visual information processing does not remove all the problems of artificial vision. According to Minsky [1], for lower levels of information processing, where selection of primary features of images occurs, parallelism is effective. On higher levels of analysis, serious limitations on the usefulness of parallelism appear. It is believed that visual perception is based on two interconnected processes: parallel processing of visual information by automatic mechanisms determined by the neural organization of the retina, LGN, and visual cortex, and sequential processing connected with image recognition mechanisms and controlled by attention processes [2–4]. Detector properties of single neurons and neural groups, such as orientation selectivity of cortical neurons, play an important role in the first process. In the second process, eye movements are of great importance, by means of which the most informative parts of an image are fixated on the fovea which has the highest density of receptors. Thus, the most adequate system for processing and analysis of visual information is one, whose architecture includes a neural network preprocessor simulating parallel processing of information at lower levels of the visual system and a sequential computer controlling the preprocessor for obtaining information required for image recognition.

Prior to the development and design of neural networks preprocessors, it is necessary to investigate the neural organization of lower levels of the visual system. The role of these structures is likely to be a preliminary information processing, filtering and encoding of sensory image features. However, until now the problem of what is encoded by these neural structures, and how, has not been finally resolved. Beginning with the classical work of Hubel and Wiesel [5, 6], the orientation of contrast elements of images has been considered to be one of the main features extracted by neurons of the primary visual cortex. Orientation-selective properties of neurons of the visual cortex allow them not only to encode orientations of image elements, but also to compress initial information [7].

2. The model of the iso-orientation domain of the visual cortex

2.1 *Connectivity in the visual cortex and mechanisms of orientation selectivity*

In accordance with the concept of a columnar organization, neighboring neurons of the visual cortex have a similar orientation tuning, and compose an orientation column or an iso-orientation domain [5, 6]. The visual cortex as a whole is characterized by a retinotopical organization of afferent fibers [8]. As well as a 'global' mapping of the retina to the visual cortex, retinotopically organized projections of the receptive field of hypercolumns on each iso-orientation domain are supposed to exist [9]. It is probable that afferent fibers terminate in the cortex with excitatory synapses only, and inhibition is a consequence of intracortical lateral inhibitory connections due to inhibitory neurons [10–15].

The question of the type of lateral inhibition is discussed next. The inhibition may be afferent (when afferent fibers give excitatory synapses directly onto inhibitory neurons) or recurrent (when inhibitory interneurons are excited by collaterals of axons of cells excited by afferent fibers). The investigations of Watanabe et al. [11], Hayashi [16], and Supin [17] suggest that the lateral inhibition in the visual cortex is of recurrent type.

It is obvious that the mechanism of orientation selectivity of cortical neurons is connected with the spatial anisotropy of a distribution of some fibers. The first explanation for this mechanism by Hubel and Wiesel [5] was based on the concept of an anisotropic distribution of afferent fibers forming the specific receptive fields of the cortical neurons. Another explanation can be based on a spatial anisotropy of the lateral excitation. This idea was realized in the model of Finette et al. [18]. However, the results of experiments of Sillito [19, 20] in which a weakness or disappearance of orientation selectivity by blocking intracortical inhibition was demonstrated, support the concept of Creutzfeldt et al. [14, 21], according to which spatial anisotropy of lateral inhibitory connections plays an essential role in mechanisms of orientation selectivity. Based on the description above, the iso-orientation domain of the visual cortex may be regarded at the first approximation as a screen-type neural structure with a retinotopical organization of afferent inputs and with a spatially anisotropic lateral inhibition between its elements which is realized by feedback connections through inhibitory interneurons.

2.2 *The iso-orientation domain model*

The model containing a two-dimensional neural structure and a flat layer of receptors (S-layer) of the same size was considered. For convenience the neural structure was subdivided into two flat layers: the layer of excitatory elements (the E-layer) and the layer of inhibitory elements (the I-layer) (*Fig. 1a, b*). Each element of the E-layer received an excitatory input from the corresponding element of the S-layer (this provided the topical projection S→E) and eight similar excitatory inputs from its neighbors in the E-layer that defined the isotropy of lateral excitation inside the E-layer. It was also inhibited by the corresponding interneuron of the I-layer (*Fig. 1a*). Each element of the I-layer besides the excitatory input from the corresponding element of the E-layer also received a spatially anisotropic excitation from some other elements of the E-layer and provided a spatially anisotropic lateral inhibition (*Fig. 1b*). The anisotropy was provided by a form of the region of spatial summation of excitation from the E-layer for every inhibitory interneuron. Examples of regions for different orientation tuning are shown in *Fig. 2b–e*. *Fig. 2a* schematically shows an inhibitory interneuron with anisotropic dendritic branches which could realize a spatially anisotropic lateral inhibition. The model was described by a system of differential equations which was solved numerically using the Runge–Kutta algorithm.

Bars and edges, oriented along the optimal

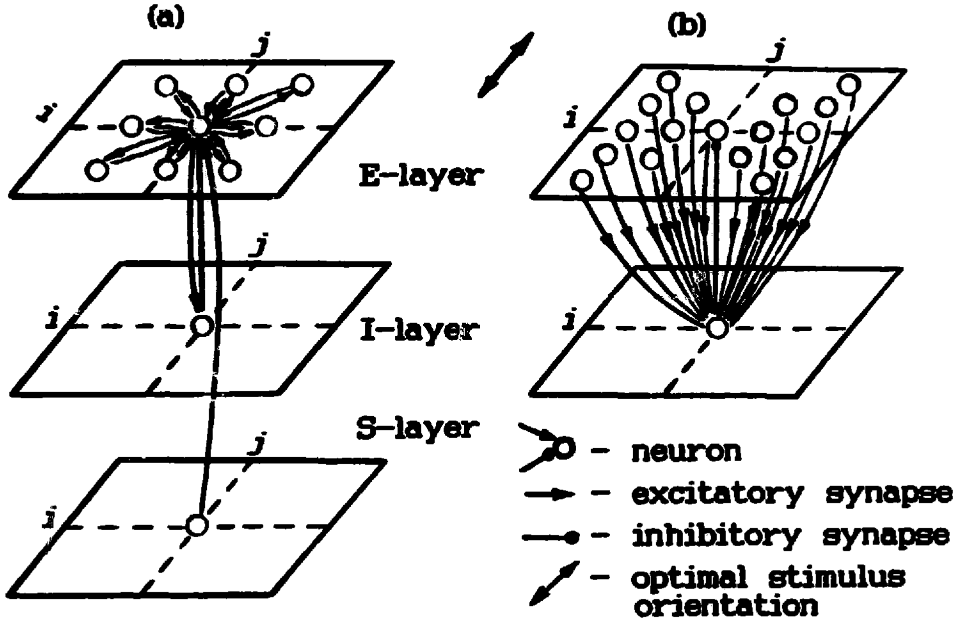


Fig. 1. The model of the iso-orientation domain of the visual cortex.

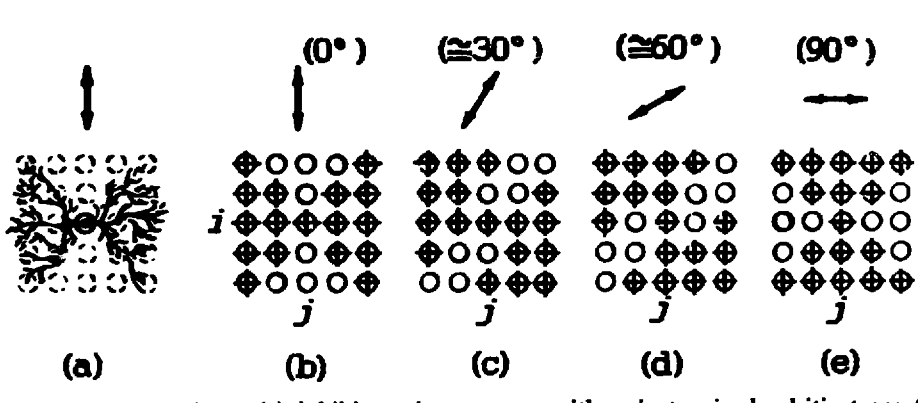


Fig. 2. Anisotropy of lateral connections: (a) inhibitory interneuron with anisotropic dendritic tree; (b–e) regions of spatial summation of excitation from the E-layer by an inhibitory interneuron of the I-layer for different directions of optimal stimulus orientation. The E-layer neurons innervating the interneuron with coordinates (i, j) are marked by the + sign.

direction (orthogonal to the direction of anisotropy of lateral inhibition) and under different angles to the optimal one, were used as stimuli in computer simulations. The purpose of the computer simulations was to investigate the dynamic responses of neurons to presentation of input stimuli.

The E-layer neuron responses to optimally and non-optimally oriented bars, edges and a diffuse stimulus are shown in *Fig. 3*. Analysis of the dynamics of the neuron responses has shown:

(i) the responses to the optimally oriented stimulus are of a sustained or biphase character with a well expressed later component;

(ii) later components of neuron responses are much more selective to the stimulus orientation than their initial phases.

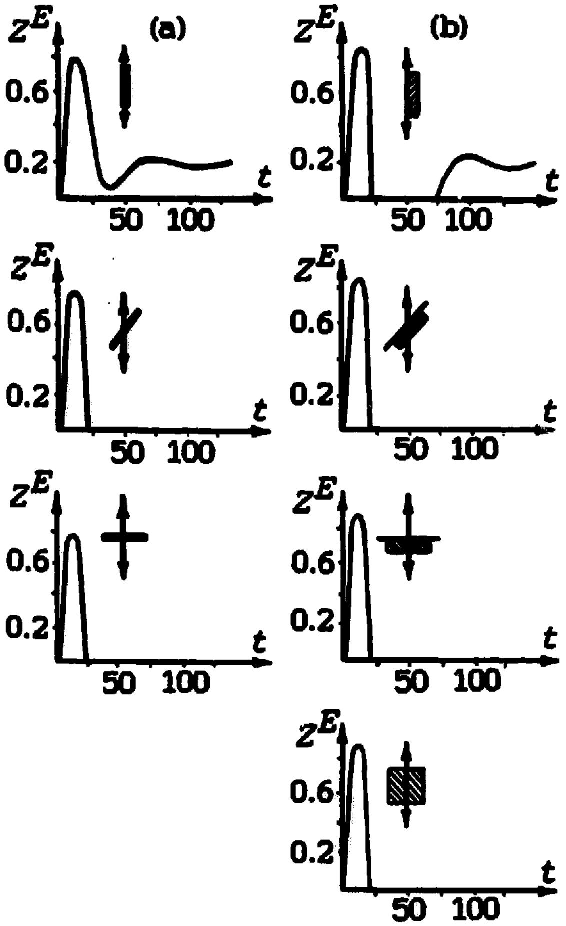


Fig. 3. The responses of the E-layer neurons to optimally and non-optimally oriented stimuli and a diffuse stimulus.

2.3 Experimental results

The experimental neurophysiological studies [22] investigated the dynamics of responses and orientation selectivity of neurons of the visual cortex. The studies were carried out on the guinea-pig visual cortex (the 17 area). The responses of neurons to stationary light bars presented in neuron receptive fields were investigated. Examples of histograms of neuron responses are shown in *Fig. 4*.

The experimental study results confirmed the predictions described at the end of the preceding

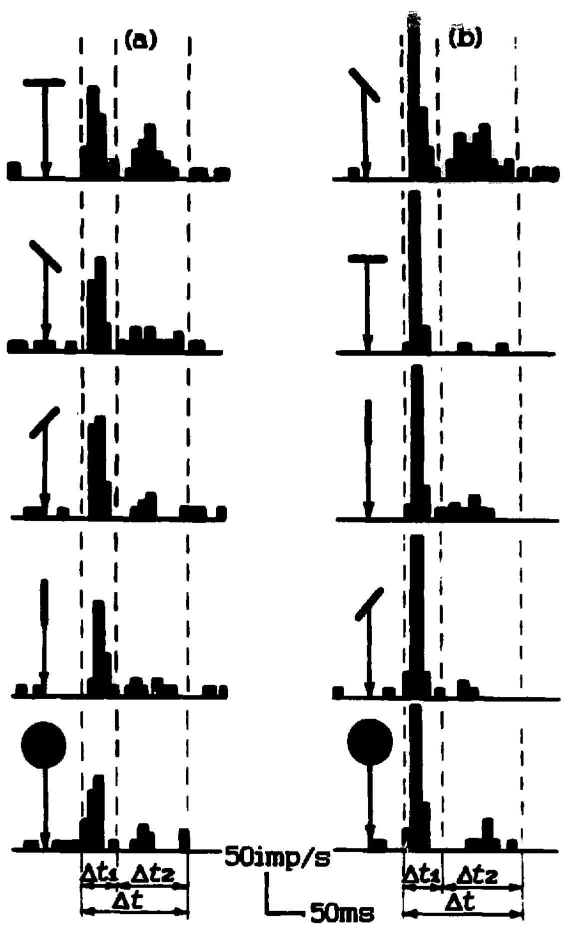


Fig. 4. The histograms of responses of two (a, b) neurons in guinea-pig's visual cortex to optimally (the upper row) and non-optimally oriented light bars and a diffuse stimulus (the bottom row). The time of stimulus presenting is marked by the vertical arrow. Δt, Δt_1, Δt_2 are the time intervals corresponding to the whole response, its initial phase and a later component.

section. Analysis of experimental data showed that 73% of all (or 85% of having a later component of responses) orientation-selective neurons gave the responses that had later components more sharply tuned to the stimulus orientation than their initial phases.

The result of modeling and the experiments allowed to suppose that feature discrimination in the visual cortex takes place not only in space but also in time. The initial phase of the response of the iso-orientation domain perhaps encodes the presence of the stimulus in the receptive field of the hypercolumn and its intensity, and the later temporal component of the response encodes the presence of the edge or contour elements in the same receptive field with the appropriate orientation.

3. Screen-type neuron-like structure as a functional module for a neural network preprocessor

The model of the Screen-type Neuron-like Structure (SNS) developed was based on the model of the iso-orientation domain described above. In principle it should realize in the preprocessor the same function as the iso-orientation domain (orientation column) in the visual cortex. SNS was represented in the model by a flat layer of 25 (5×5) neuron-like elements (*Fig. 5*). Each SNS transformed an 5×5 input signal matrix $\|S_{ij}\|$ into an output matrix $\|Z_{ij}\|$ of the same size. The center 3×3 of this matrix was considered to be functional while the surrounding elements served to eliminate boundary effects. Outputs of nine central elements of SNS entered the output section which encoded features characterizing the processed part of the input raster consisting of nine central elements of the matrix $\|S_{ij}\|$. Each neuron-like element of SNS was an analog of the pair of an E-neuron and an inhibitory I-neuron (*Fig. 6*). An 'afferent' input signal X_{ij} for each (i, j)-element of SNS was formed through the block F^S producing a spatial convolution with the input matrix $\|S_{ij}\|$ according to the type of on-center/off-periphery receptive field. The inertial recurrent inhibition on (i, j)-element of SNS was realized through an inertial block and an F^I-block producing spatial convolution with output matrix $\|Z_{ij}\|$. These convolution coefficients provided the anisotropic character of lateral inhibition.

According to the simulation results described above each (i, j)-element of SNS will exhibit either a monophase response (*Fig. 7a*) or a response with a later component (*Fig. 7b*). The intensities of the initial phase of response and of its later component were estimated at the moments t_0 and t_1 correspondingly. The moment t_0 corresponded to the moments of presenting a stimulus and the moment t_1 was chosen depending on the time constant τ and the maximal possible value of the stimulus intensity. The states of the outputs $\|Z_{ij}\|$ of SNS, which were

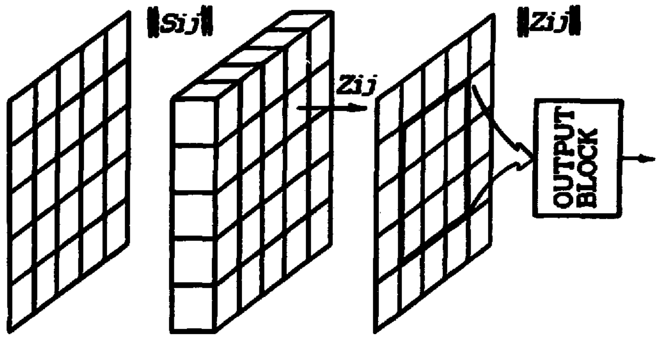


Fig. 5. The Screen-type Neuron-like Structure (SNS).

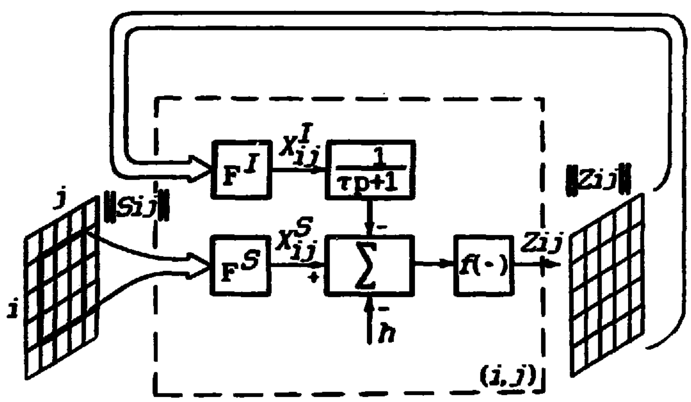


Fig. 6. Scheme of the (i, j)th neuron-like element of SNS.

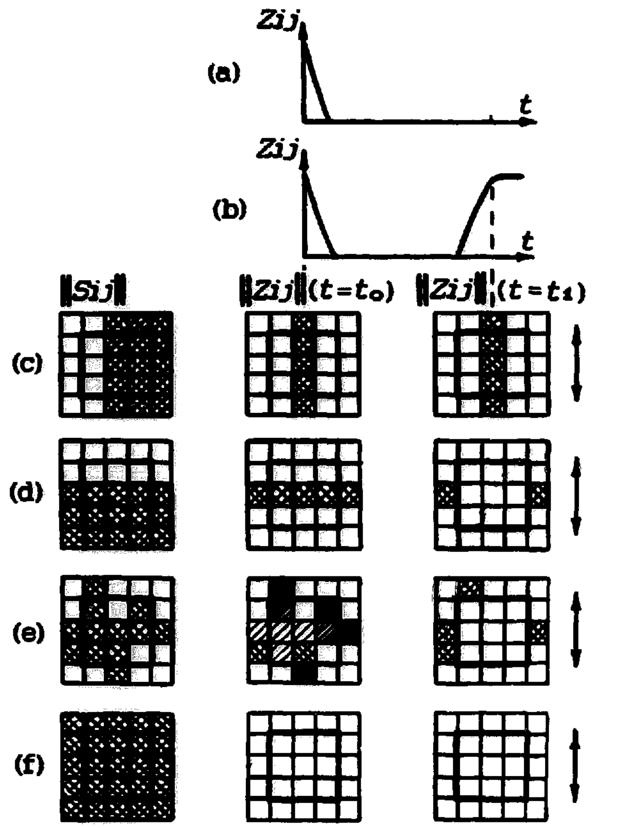


Fig. 7. Dynamics of the SNS responses a, b – two types of response dynamics of the SNS elements; c–f – the output responses of SNS to different input stimuli.

tuned to the vertical orientation of the stimulus at the moments t_0 and t_1 with different input stimuli are shown schematically in *Fig. 7c–f.* If there was an optimally oriented edge in $\|S_{ij}\|$ at t_0 (*Fig. 7c*) the functional area contained three excited elements, otherwise the number of excited elements was less than three. At t_0, the initial phase of the response was characterized by the presence of excited elements in all cases except the case of uniform intensity in $\|S_{ij}\|$, i.e. when all non-uniformities were lower than the threshold (*Fig. 7f*). Thus, during information processing, SNS at the moment t_0 provides an output signal indicating the presence of suprathreshold contrasts in this area and then, at the moment t_1, produces a signal indicating the presence of visual image elements (bars or edges) whose orientation corresponds to the tuning of SNS.

4. The neural network preprocessor in the system for processing and analysis of visual information

4.1 *Multiresolution encoding of contour features and parallel-sequential processing of visual image*

In drawing a representational picture from nature, an artist is trying to achieve similarity, using an algorithm of drawing 'from-coarse-to-fine'. In this process he draws a set of auxiliary lines that do not correspond to exact real contours of the subject, but are observable by consideration of the subject with a less visual resolution. These generalized contour elements and the spatial relations between them are of great importance for the exact reproduction of subjects. Evaluation of the spatial relations (distances and orientations) between smaller elements of the contours (found by finer definition) and these generalized auxiliary lines are not less important. Also, the artist does not complete the process of detailing but elaborates in detail only the most important or significant parts of the picture.

It is attractive to assume that the artist intuitively uses an algorithm of successive analysis of images in the visual system connected with the sequential fixations of the moving eyes. In the process of image recognition the human eyes are permanently moving and fixate on the most informative fragments of the picture (*Fig. 8*). When the gaze is fixated, the fragment is projected onto the fovea, which has the largest representation in the visual cortex. It may be noted that usually one does not attach great importance to the fact that during the process of successive fixations one and the same fragment of an image is repeatedly projected onto the retina with different distances from its center and processed in the visual cortex with a different degree of resolution. In this case, the system of iso-orientation domains of the appropriate parts of the visual cortex can extract and encode the

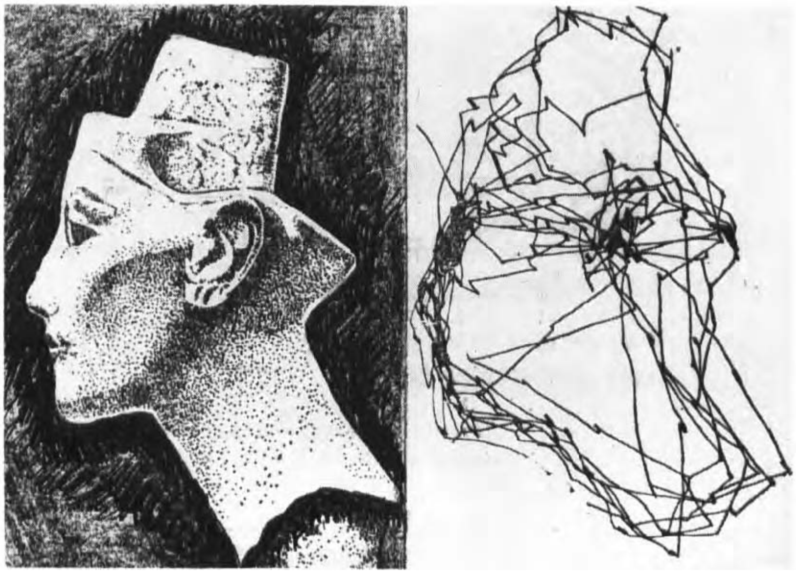


Fig. 8. Trajectories of eye movements and gaze fixations in image recognition by a human observer [23].

edge and contour elements of different extent of generality. Encoding some generalized contour elements and the spatial relations between them and contour elements extracted in more detailed subsequent analysis (when the parts of the image are projected on the center of the retina) can provide the set of primary features for the invariant representation of the shape of analyzed images. As described above, the two-phase character of the response of the iso-orientation domains of the visual cortex allows to presume that the intensive first phase (that carries information about the presence of a pronounced relief of brightness in the processed fragment of an image) may be a signal for the system controlling an automatical transfer of the gaze. In the secondary phase of the responses of the system of iso-orientation domains the orientation filtration and the discrimination of contour elements and edges take place. It should be noted that reafference from the oculomotor system to the visual cortex is of great importance in the encoding of spatial relations.

4.2 An example of processing a test image

The structural scheme of the system for processing and analysis of visual information is presented in *Fig. 9.* An initial image is given in the raster of 243×243 pixels. The programmed switch module which connects the preprocessor with one or another part of the image plays the role of the oculomotor system. Five levels of resolution were examined, on each of which the raster contained $3^n \times 3^n$ pixels (where n is the number of the resolution level, $n = 1, \ldots, 5$). The stimulus intensity at every point of the lower resolution raster was determined by averaging the intensity along nine (3×3) higher resolution raster points.

During each time step of sequential processing a part of the image of size 5×5 elements was

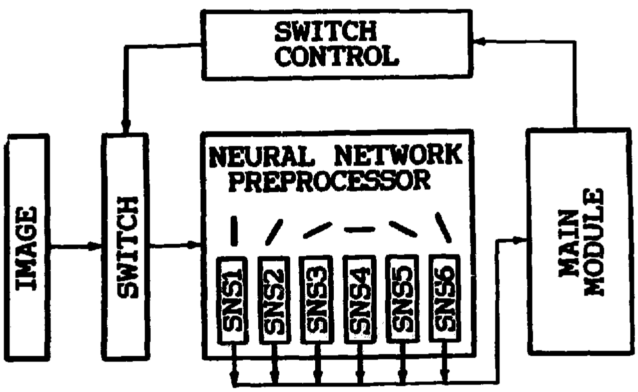


Fig. 9. Scheme of the system for processing and analysis of visual information with the neural network preprocessor.

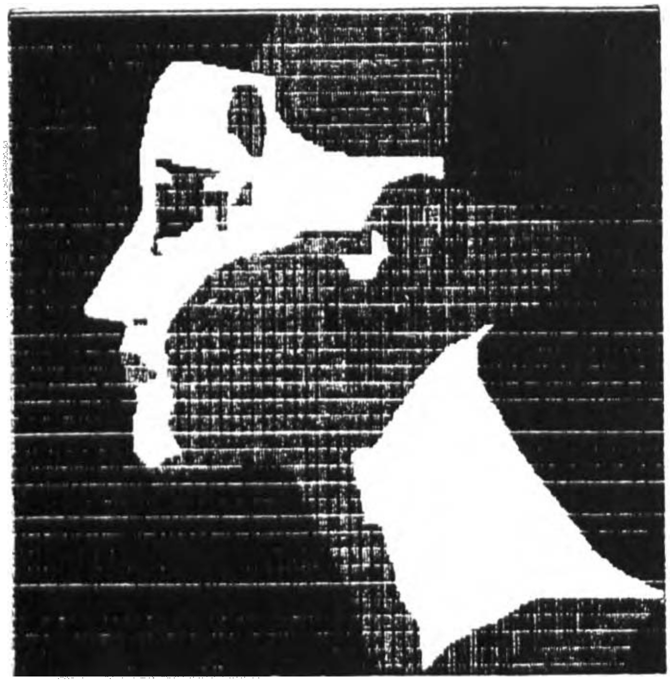
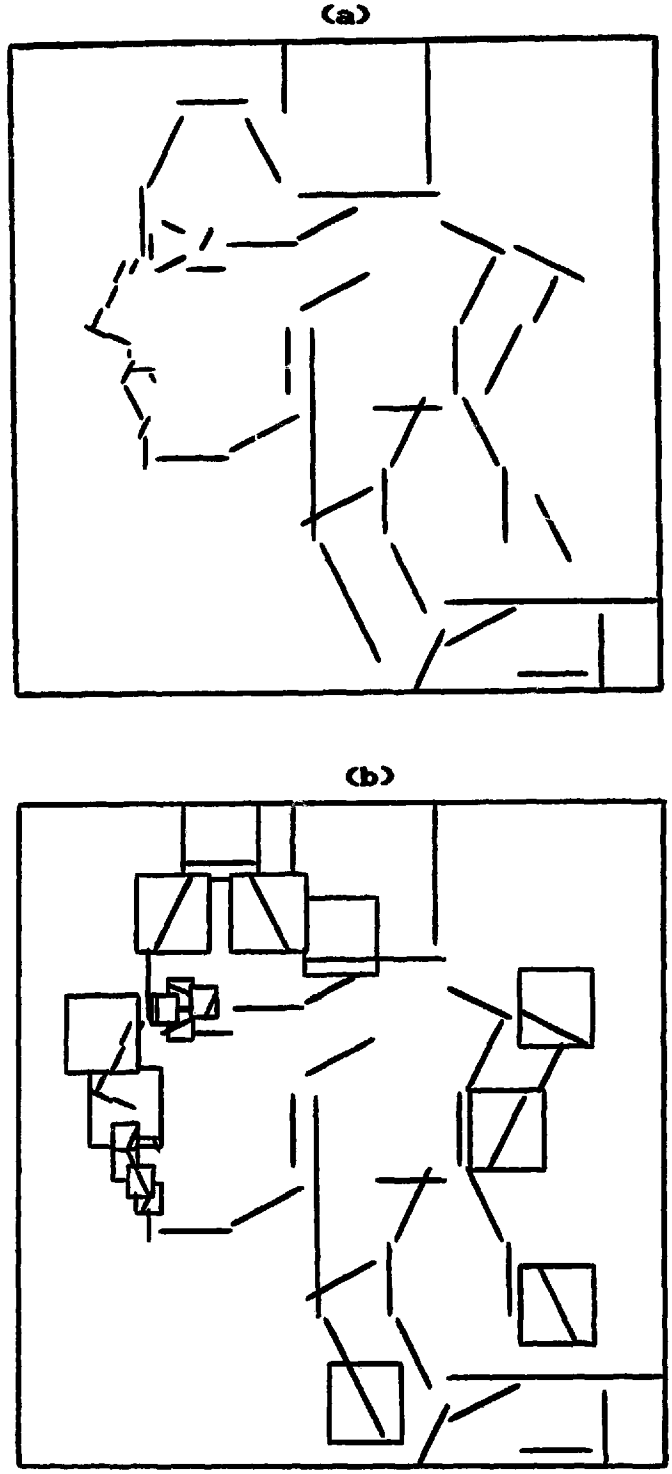


Fig. 10. Test image.



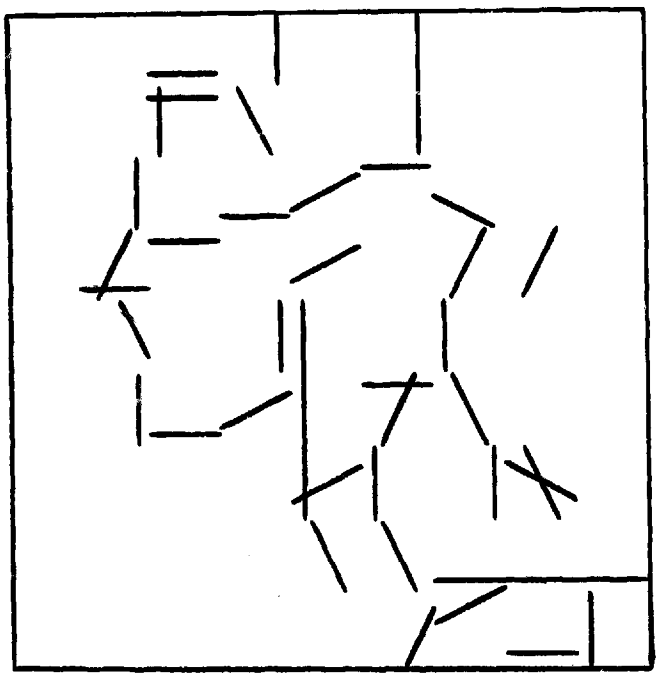


Fig. 11. Result of processing of the test image at the third level of resolution.

Fig. 12. Result of initial image processing of the typical fragments with enough degree of detail. Processing was carried out by scanning the raster at the third level of resolution with the additional processing of some fragments at the third and fourth levels. In (b), additionally processed fragments are framed.

presented to the input of the neural network preprocessor consisting of six SNS whose selective orientation tuning was varied successively in steps of 30°. The set of SNS initially gave information on the presence of intensity non-uniformity and then discriminated the orientation of the edges or contour elements belonging to this part of the image. The preprocessor output (from the set of SNS) should be taken to the main module which should store this information and, using data about the location and the level of resolution, calculate invariant relations between contour elements detected in different steps of analysis. The main module should also compare the detected features (invariant spatial relations) with the template ones and form a strategy for the next steps of image processing. This strategy should be realized by the module of switch control.

The test image (Nefertiti head) is given in *Fig. 10*. *Figure 11* represents the result of test image processing on the third level of resolution. In *Fig. 12*, the result of processing on the third level with detailing of some parts on the fourth level of resolution is shown. The degree of detail appears to be high enough to recognize the object.

References

[1] M. Minsky, A framework for representing knowledge, in: P.H. Witston, ed., *The Psychology of Computer Vision* (McGraw-Hill, New York, 1975) 211–277.

[2] V. Neisser, *Cognitive Psychology* (Appleton, New York, 1967).

[3] B. Julesz, Experiments in the visual perception of texture, *Scientific American*, 232 (1975) 34–43.

[4] R.M. Shiffrin and W. Schneider, Controlled and automatic human information processing. 2. Perceptual learning, automatic attending, and a general theory, *Psychol Rev*. 84 (1977) 127–190.

[5] D.H. Hubel and T.N. Wiesel, Receptive fields, binocular integration and functional architecture in the cat's visual cortex, *J. Physiol*. 160 (1962) 106–154.

[6] D.H. Hubel and T.N. Wiesel, Sequence, regularity and geometry of orientation columns in the monkey striate cortex, *J. Comparative Neurol*. 158 (1974) 267–293.

[7] M. Kunt, A. Ikonomopoulos and M. Kocher, Second-generation image coding techniques, in: *Proc. IEEE* 73 (1985) 549–574.

[8] S.A. Talbot and W.H. Marshall, Physiological studies on neural mechanisms of visual localization and discrimination, *Amer. J. Physiol*. 24 (1941) 1255–1263.

[9] F.L. Schwartz, A quantitative model of the functional architecture of human striate cortex with application to visual illusion and cortical texture analysis, *Biol. Cybernet*. 37 (1980) 63–76.

[10] C. Stefanis and H. Jasper, Recurrent collateral inhibition in pyramidal tract neurons, *J. Neurophysiol*. 27 (1964) 855–877.

[11] S. Watanabe, M. Konishi and O.D. Creutzfeldt, Postsynaptic potential in the cat's visual cortex following electrical stimulation of afferent volleys, *Experimental Brain Res*. 1 (1966) 272–283.

[12] M. Ito, Neuronal linkage in the cat visual cortex, *J. Physiol. Soc. Japan* 32 (1970) 550–551.

[13] L.G. Garey and T.P.S. Powell, An experimental study of the termination of the lateral geniculo-cortical pathway in the cat and monkey, in: *Proc. Royal Soc. London*, Ser. B, 119 (1971) 41–63.

[14] L.A. Benevento, O.D. Creutzfeldt and U. Kuhnt, Significance of intracortical inhibition in the visual cortex, *Nature, New Biol*. 238 (1972) 124–126.

[15] D. Ferster and S. Lindstrom, An intracellular analysis of geniculo-cortical connectivity in area 17 of the cat, *J. Physiol*. 342 (1983) 181–215.

[16] Y. Hayashi, Recurrent collateral inhibition of visual cortical cells projecting to superior colliculus in cats, *Vision Res*. 9 (1969) 1367–1380.

[17] A.J. Supin, Feedback excitation and inhibition in the visual cortex, *Neurophisiologia* 2 (1970) 418–422 (in Russian).

[18] S. Finette, E. Harth and T.J. Csermely, Anisotropic connectivity and cooperative phenomena as a basis for orientation selectivity in the visual cortex, *Biol. Cybernet*. 3 (1978) 231–34.

[19] A.M. Sillito, The contribution of inhibitory mechanisms to the receptive field properties of neurons in striate cortex of the cat, *J. Physiol*. 250 (1975) 305–329.

[20] A.M. Sillito, Functional consideration of the operation of GABergic inhibitory processes in the visual cortex, in: *Cerebral Cortex*, vol. 2 (Plenum Press, New York/London, 1984) 91–117.

[21] O. Creutzfeldt, U. Kuhnt and L.A. Benevento, An intracellular analysis of visual cortical neurons to moving stimuli: Responses in a cooperative neuronal network, *Experimental Brain Res*. 21 (1974) 251–274.

[22] I.A. Rybak, Study of response dynamics and orientation selectivity of visual cortex neurons by methods of

mathematical modelling and neurophysiological experiment, Doctor Thesis, USSR, Rostov-on-Don, 1988.
[23] A.L. Yarbus, *The Role of Eye Movements in Vision Process* (USSR, Moscow, Nauka, 1965).

Ilya Rybak founded and is the Head of the Laboratory of the Neural Network Modeling in Vision at the Institute of Neurocybernetics at Rostov State University, Rostov-on-Don.

He received the M.D. in Electrical Engineering in 1975 at Odessa Polytechnic Institute. He has worked in the Institute of Neurocybernetics since 1977. His investigations included study and modeling of a single nervous cell, modeling of neural networks with supervisor learning, study and modeling of visual cortex, computer vision based on neural network approach. Dr. Rybak received the Ph. D. degree in biophysics at Leningrad State University in 1988. His doctor thesis was devoted to the study of response dynamics and orientation selectivity of the visual cortex neurons by mathematical modeling methods and neurophysiological experiments.

Dr. Rybak has been a chief of several research projects devoted to modeling of the visual cortex and to applications of artificial neural networks to computer vision. His research interests include vision neurophysiology, neural network models, and computational vision. He has published more than 50 papers.

Natalia Shevtsova is a senior researcher at the Laboratory of the Neural Network Modeling in Vision at the Institute of Neurocybernetics at Rostov State University, Rostov-on-Don.

She received the M.D. in Mathematics in 1973 at Rostov State University. She has worked at the Institute of Neurocybernetics since 1974. Her investigations included modeling of neural networks, study and modeling of the visual cortex, computer vision based on a neural network approach.

Dr. Shevtsova has been a leader of several research projects devoted to modeling of the visual cortex and to applications of artificial neural networks to computer vision. Her research interests include vision neurophysiology, neural network models, and computational vision. She has published more than 30 papers.

Vladislav Sandler received the M.D. in Biology in 1990. Being a student, he took part in research carried out at the Laboratory of Neural Network Modeling in Vision at the Institute of Neurocybernetics at Rostov State University, Rostov-on-Don. His investigations included study and modeling of the visual cortex, computer vision based on a neural network approach, vision neurophysiology, neural network models, and computational vision. He has published 12 papers. Presently, he is a post-graduate student of the Tel Aviv University, Israel.