DOI 10.1007/s10559-023-00537-7
Cybernetics and Systems Analysis, Vol. 59, No. 1, January, 2023

CYBERNETICS

THE ISSUES REGARDING THE ORGANIZATION OF THE HUMAN RETINA: A CYBERNETIC APPROACH

V. P. Boyun

UDC 159.93

Abstract. *The issues in the organization of the human retina are analyzed from the point of view of cybernetics. In particular, hypotheses are formed regarding the conditions of the concentric organization of receptive fields, the extraction of informative features, the possibilities of restructuring the receptive fields and their limitations, the transmission of information through ganglion cells with its binding to the location on the retina, as well as regarding the peculiarities of the visual system in different modes. The conditions for the correct operation of on- and off-centers of receptive fields based on bipolars with a cone as the central zone of excitation and the surrounding rods as an inhibition zone, organized with the help of horizontal cells for the on-center, and vice versa for the off-center, are determined. The conditions for the extraction of informative features based on on- and off-centers in the form of point values of brightness differences between neighboring elements are substantiated. The rigid organization of the receptive fields of ganglion cells and the limited possibilities of its reconstruction are substantiated. A hypothesis on the organization of the transmission of information about brightness differences from bipolar cells through ganglion cells under the control of amacrine cells with reference to the location of the retina is proposed. A hypothesis on the continuous perception of the image by the retina receptors and the discrete perception of information by the brain is formulated. The features of the visual system in the modes of contemplation, searching for an object in an image, tracking it, and detailed examination and recognition are considered.*

Keywords: *peripheral retina, fovea, rods and cones, horizontal, bipolar, amacrine and ganglion cells (neurons), on- and off-centers, receptive fields, concentric organization, neural network, informative features, information transmission, modes of functioning of the visual system.*

INTRODUCTION

The author of this article has been researching computer vision systems for over 20 years, using the basic functions and mechanisms of the human visual analyzer as the prototype for the last 12 years. By taking the research results and their interpretation by neurophysiologists on faith and by using them in the practical development of computer visual systems adjusted for technical implementation possibilities, the author obtained mostly positive results. In particular, here, it is the case of the concentric organization of receptive fields for the isolation of different informative features, the process of growing the rod cell layers on the peripheral retina under the condition of insufficient illumination or of neural layers around the central element under a low contrast, the coarse-to-fine searching for and

V. M. Glushkov Institute of Cybernetics, National Academy of Sciences of Ukraine, Kyiv, Ukraine, *vboyun@gmail.com*. Translated from Kibernetyka ta Systemnyi Analiz, No. 1, January–February, 2023, pp. 3–20. Original article submitted September 29, 2022.

tracking of an object in an image in particular, spatial frequencies, etc. It generated many published articles, issued invention patents, and a number of developed intelligent real-time cameras and video systems. However, when later contrasting the results obtained by many human visual analyzer researchers, the author noticed that some, in his opinion, false result interpretations have migrated from the articles into monographs, especially those related to the dynamic information perception and processing models, the neural plasticity mechanisms, the adaptation and control, etc. Thus, it was decided to analyze the organization of the receptive field, its functioning principles, and the lower-level models of the human visual analyzer based on the experience with constructing computer vision systems from the point of view of cybernetics, which is the field where the author has been working for over 50 years. In the author's opinion, cybernetics as a multidisciplinary science of the general laws of obtaining, storing, and transforming information in complex technical, biological, administrative, and social systems [1, 2] could promote systematic reevaluation of brain processes (particularly those that are related to vision), as well as direct neurophysiologists and other experts towards future research areas.

GENERAL INFORMATION ON THE STRUCTURE AND ORGANIZATION PRINCIPLES OF THE HUMAN RETINA

The retina of the human eye consists of a layer of receptors (rods and cones) and of four layers of neurons, namely, horizontal, bipolar, amacrine, and ganglion neurons. Groups of rods and cones are vertically connected to the bipolar cells, which are themselves connected to the ganglion cells whose axons are part of the optic nerve. There are also two layers of horizontal connections where the first network is created by the horizontal cells between the photoreceptors and the bipolar cells, and the second network of side (lateral) connections is created by the amacrine cells (between the bipolar and ganglion cells). One of the functions of these side connections is to create interactions between the neighboring cells so that bipolar cells can extract informative features from a picture and the ganglion cells can transmit the information, respectively.

The general organization of the retina is presented in Fig. 1. Note that the light enters the retina not from the side of receptors, but passes through the neural network being reflected and refracted multiple times, and is then perceived at an acute angle in this randomized (noisy) form by the side surface of the receptors. This creates a hexagonal diffraction pattern, and it is the best method of adapting to a heterogenous environment. The more random the image, the more information it contains, which ensures a better adaptation to environmental conditions and allows for better survival chances of species in the evolution process.

There are two main structural units found in the retina that significantly differ from each other by their organization and their function. These units are the peripheral retina and the fovea centralis. The peripheral retina is mostly made out of rods, which ensure that the light is perceived in the shades of gray, as well as out of a much smaller number of cones, which ensure the perception of light in the defined wave ranges that are interpreted by the brain as three primary colors, namely, red, green, and blue (the RGB color model). The cones are mostly and more densely packed in the fovea. The basic organization principles of the peripheral retina and the fovea is the concentric organization of receptive fields of the bipolar cells. The receptive fields of the bipolar cells have two zones, namely, the central excitation zone and the surrounding it inhibition zone for organizing the so called "on-centers" and the reverse "off-centers," which ensure the extraction of informative features that are bound to the location on the retina from an image. The inhibition zone of bipolar cells is organized by the horizontal cells. The fovea of the retina plays the role of some sort of a sensor that is meant to trace and examine objects more closely. The information of the bipolar cells is transmitted through the ganglion cells into the lateral geniculate nucleus (LGN). The transmission process is controlled by the amacrine cells.

The process of transmitting information from the receptors to the brain is an ascending one. The information from the brain is not directly received by the retina, but it controls the oculomotor movements (descending processes).

The ganglion cells are transmitting information from the peripheral retina with magno (M type) receptive fields and from the fovea with parvo (P type) receptive fields to the LGN, which has larger receptive fields (allowing for the combination and generalization of the information coming from the ganglion cells of the retina) and performs communicative functions by redirecting information to other brain layers.

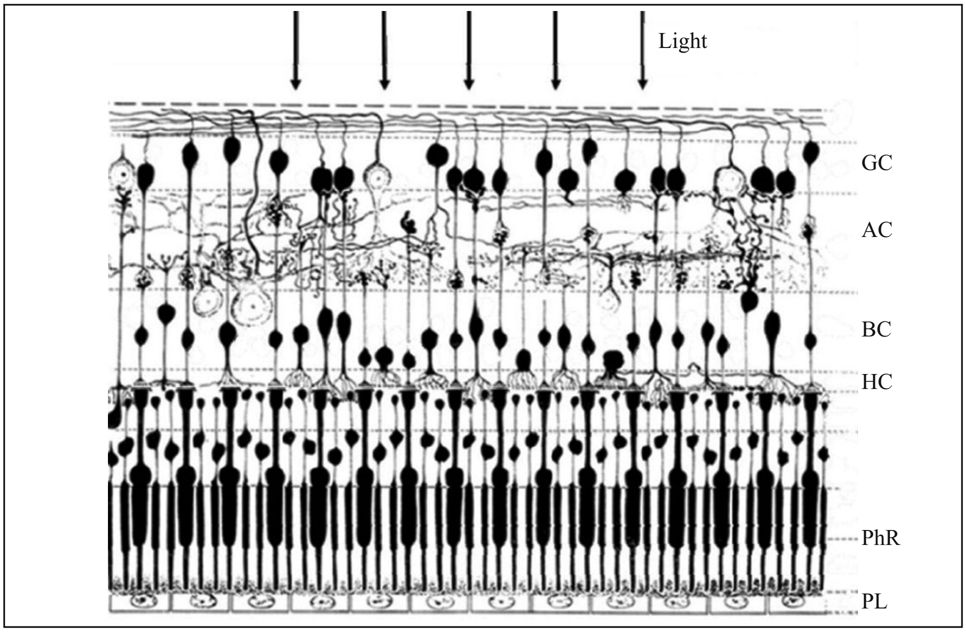

Fig. 1. The inverted retina of the human eye where PL is the pigment layer, PhR is the photoreceptors, HC is the horizontal cells, BC is the bipolar cells, AC is the amacrine cells, and GC is ganglion cells.

It is obvious that the LGN has a mixed structural organization for perceiving and processing information coming from the peripheral retina and the fovea. When it comes to the concentric organization of the LGN neurons, it may be mostly related to the implementation of the trichromatic theory of color by the fovea as the contour information from the peripheral retina is processed by the linear brain structures.

A more detailed information on the structural features of the retina organization, the physical features of light perception, and the mechanisms and processes taking place in the retina are presented in [3–12] in different interpretations.

The aim of the article is the analysis of the challenging questions on the organization of the human eye retina, the formulation of hypotheses on the conditions of the concentric organization of receptive fields, the extraction of informative features, the possibilities of reconstructing the receptive fields and their limitations, the transmission of information through the ganglion cells with its binding to the location on the retina, and the peculiarities of the functioning of the visual system in different modes from the point of view of cybernetics aimed at improving their models and research.

ANALYZING THE RETINA ORGANIZATION PRINCIPLES

A lot of research on the morphology of retina cells and their physiological properties, and on the organizational and functional principles of the receptive fields of the human visual analyzer as a multidisciplinary object is already done by neurobiologists, psychologists, cognitive specialists, and researchers from many different fields (such as optics, biophysics, electrochemistry, mechanics, neurophysiology, communication, informatics, system analysis, etc.) [3–12]. However, the extremely high complexity of the analyzer organization and functioning does not allow for the processes of information processing and transmission, of adaptation to different conditions, and of control of the commutation and of the transmission of the information at all the levels to be studied in real time using the methods of the above sciences as these processes are simultaneously taking place in hundreds of billions of retina neurons with thousands of connections between them. That is why the general picture of the functioning of the human vision analyzer, especially in real time, stays practically undetermined.

It is worth remembering the stages and approaches to the examination of the processes of perception and sensation of the visual information by the human visual analyzer. These are empiricism, structuralism, Gestalt psychology, the constructivist, ecological, informational, and neurophysiological approaches, cognitive neurology, functional magnetic resonance imaging, etc. [3–15].

Thus, it is possible to understand these processes only by familiarizing ourselves in detail with the modern results obtained by the researchers of many fields, their comprehension, and the development of hypotheses for modeling and research.

The main challenging questions of the organization of the human eye retina are as follows:

— the concentric organization of the receptive fields of the bipolar retina cells;

— the extraction of informative features from the receptive fields of the eye retina (spots or information for contours);

— the organization of the ganglion cells based on the bipolar cells using the amacrine cells;

— the binding of the information to the location of the retina;

— the plasticity of the neural network and the possibility of its reconstruction;

— the transmission of the information through the ganglion cells to the LGN;

— the operational mode of the human visual system.

These questions are deeply interconnected and very important for the correct understanding of the processes of functioning of the visual system.

Despite the complex organization of the retina, neuroanatomists identified the main structural features of its construction, while neurophysiologists determined its general function.

Apart from the layer of receptors (mostly rods), the peripheral retina has four neural layers with different functions and different branching degrees of the dendrites and axons. Moreover, the functions of many of them remain unclear. The retina performs the function of perceiving the video information through the rods and cones, extracting informative features from the image, and transmitting it to the LGN and other brain layers. The amount of knowledge available today is insufficient to describe the process of information perception in detail. Moreover, taking into account the organizational complexity of the retina and of the brain, this is hardly possible in the foreseeable feature. Therefore, let us set aside the details and try to at least generally describe the components of the process of information perception by the retina.

The concentric organization of the receptive fields discovered by Stephen Kuffler is formed on the bipolar cells of the retina using horizontal cells and is based on the use of two zones, namely, the central excitation zone and the inhibition zone in the form of a circle around it. This is confirmed by all the researchers and is without a doubt. The on- and off-centers are formed on this basis, and they ensure that the brightness differences between the lighter zones of the retina on the dark background (an on-center) and vice versa (an off-center) are emphasized. The centers are an exceptionally powerful and useful tool for extracting informative features from an image, but the use of the concentric organization is not always correct.

Researchers indicate, on the one hand, that the inhibition zone should be much larger than the excitation zone, and that the simultaneous illumination of both of the zones by the diffused light should not evoke a reaction from the on- and off-centers on the other. However, the total signal of the larger inhibition zone is going to almost always suppress the total signal from the smaller excitation zone, evoking a significant reaction of the on- and off-centers in the case of diffused illumination of the center and its surrounds. It may not have been noticed in the case of conducting research aimed at determining receptive fields by illuminating the excitation zone with a bright light, while a large area of the inhibition zone remained dark. Thus, when summing up signals from the excitation zone, its brightness sum was exceeding that of the inhibition zone of the receptive field and evoking the correct reaction of the on- and off-centers. It is known that the receptor signal strength is directly proportional to the illumination brightness and the receptor area; thus, on- and off-centers on uniform receptors (rods or cones) are not going to function without coordination and without taking into account the ratio of excitation and inhibition zones, i.e., they are not going to emphasize brightness differences of neighboring elements of the receptor image.

In order for the excitation/inhibition centers to function correctly, the ratio of the excitation and inhibition areas should be accounted for, i.e., the influence of the excitation center on the area ratio coefficient should be increased. Moreover, in the case of the diffused illumination of both of the centers, zero value at the bipolar cell output is ensured.

However, in the case of even the smallest brightness difference relative to the inhibition zone appearing in the excitation center, the brightness difference value increases by the area ratio coefficient. This includes both the increase and the decrease in the brightness difference, i.e., both the on- and off-centers are activated, and they greatly increase the difference scale. In other words, high sensitivity to the contrast highlighting in the pinpoint excitation zone is ensured.

Note that the concentric organization principle is widely used in computer vision systems to extract a high variety of informative features. In particular, many different masks are developed to determine object boundaries, contours, angles, and orientation, filtering from different kinds of interferences, etc [16–27].

To ensure the optimal object recognition, cortical neurons of the brain have to obtain informative features from the retina invariantly, i.e., irrespective of the differences in image features that are quite important but not crucial for the spatial analysis, such as the light background level, the contrast of the object with the background, the time duration of its functioning, etc. The organization of receptive fields according to the "center-surrounds" principle promotes the reduced influence of lighting differences when it is concerning both the center and its surrounds. This is why they are more effective when determining the pinpoint information on the boundaries of differently illuminated objects in an image.

Some scientific articles [4, 8–11] consider the use of the large excitation centers on the peripheral retina as it supposedly corresponds to the organizational and functional principles of the retina and allows for the extraction of light or dark spots where, despite them being different from the background, objects can still be found. However, this is hardly correct as it necessitates a considerable reconstruction of the receptive fields, which is a very demanding mechanism for a vision system, and it is not suitable for extracting spots with different brightness, especially elongated ones. In these cases, spots are extracted imprecisely, they necessitate the transmission of huge additional information volumes (such as addresses, dimensions, and junctions not only of the bipolar cells of the ganglion cells, but of the ganglion cells themselves if a spot exceeds the boundaries of the receptive field of one ganglion cell, etc.) through the ganglion cells to the LGN, and do not correspond to the retina organization principles. Moreover, important details can be missed in the case of large spots [5, 6, 19, 20, 27, 28].

In the case of the computer vision systems with a homogenous receptive field, scientists were exploring the possibility of extracting spots of different dimensions that differ from the background using different masks, such as, for example, the Laplace mask of the size 3×3. The mask elements were the corresponding image fragments of the size 2×2, $3\times 3, \dots$, i.e., corresponding to the dimensions of the spot being extracted. Moreover, the coefficient of the increasing influence of the excitation center on the area ratio of excitation and inhibition zones remained constant and was equal to the Laplace mask, namely, 8. Here, the following conclusions [20, 21, 27, 28] can be drawn:

(i) the spot being extracted is represented by the mean brightness or color value in the excitation zone, i.e., it does not always ensure a precise extraction of the object location, especially when it comes to elongated objects. In the case of larger excitation zones, the object location centers are determined imprecisely, which the author finds unacceptable in a vision system. The use of elongated masks in order to search for an object necessitates a large number of manipulations with the mask size and orientation, which is not conducted in vision systems;

(ii) a powerful apparatus of the concentric organization of the excitation/inhibition centers over the minimal image elements ensures the possibility of a precise emphasizing of brightness differences in an image, i.e., extracting the information on object boundaries as these differences include around 80% of the information necessary for object recognition. Moreover, this approach is implemented at the first information processing stage when organizing the excitation/inhibition centers at separate peripheral retina points, which decreases the amount of information necessary for consequent processing by one or two orders.

Since the retina cannot determine the area ratio of the inhibition and excitation zones, as well as perform the multiplication of the total signal value in the excitation zone by this coefficient, the centers located on homogenous cells, for example, rods, cannot function. However, nature found another way to solve this problem that was not noticed for a long time.

It is stated in many scientific articles by the oculomotor system researchers that there are mostly rods and a small percentage of cones located on the peripheral retina, and that the manner in which they should be considered is unclear, which led to the decision of not considering them at all. The solution to the situation with the rods on the peripheral retina can be promoted by examining a photograph of a receptive field section (Fig. 2) from [3] where little cones can be seen surrounded by little rods.

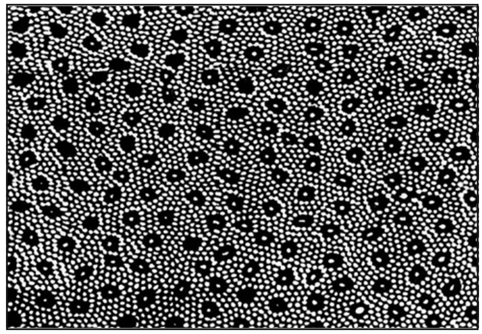

Fig. 2. A cross-section of the peripheral retina of a human eye.

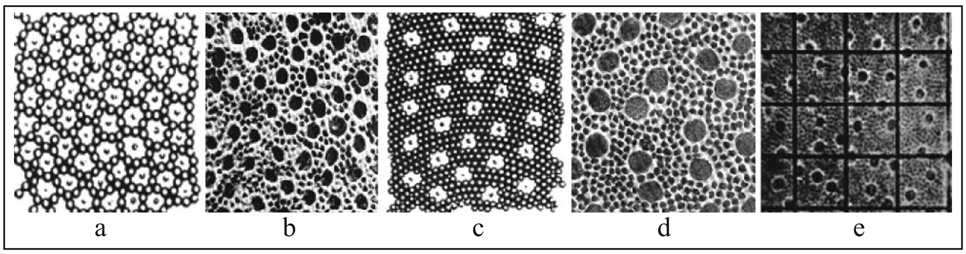

Fig 3. A peripheral cross-section of the five zones of the retina, namely, the parafovea (a) and perifovea (b), and the near-peripheral (c), mid-peripheral (d), and far-peripheral (e) zones.

A more detailed information on the structure and organization of the different zones of the peripheral retina and fovea is presented in [5]. There are seven zones distinguished in the retina, namely, four zones in the fovea (the foveola, fovea, parafovea, and perifovea) and three zones in the peripheral retina (the far-peripheral, mid-peripheral, and near-peripheral) (Fig. 3). The number of rods and cones in each zone, their diameters, the ratio of rods to cones in the different retina zones, and the number of ganglion cells in the zones is stated in Table 1.

The presence of rods that are ingenuously located around the retina cones, the different dimensions of the receptive fields of the retina, and the differently sized rods and cones in the peripheral retina zones aimed at ensuring a high sensitivity within a wide range of light differences allows for a presumption that nature organized the compact excitation/inhibition zones around the nearest bipolar cell receptors more efficiently, i.e., made the cones into the center and made the surrounding rods into the inhibition zone. The bipolar cell outputs create receptive fields for the ganglion cells.

From Table 1 it is evident that the ratio of the diameter squares or areas of the cones to the rods does in a certain way comply with the organization requirements of the excitation/inhibition centers. Moreover, the ratio is maintained in all the peripheral retina zones, and even in the case of a significant change in the number and size of the rods surrounding a cone, it does not change in time and fulfills their correct functioning requirements. It is unlikely that all the retina rods of the receptive field of a bipolar cell in a certain retina zone are absolutely identical in their form and size. Perhaps that is why when growing a cone during the development of a child, the nature would constantly harmonize its area with the total area of the rods by using feedback and by constantly checking that there are no signals from the on- and off-centers in the case of the diffused lighting. Compared to the commutation of the receptor outputs over large distances, this organization is simpler by several orders of magnitude as pinpoint illumination differences are already emphasized, meaning that the information determining the boundaries of objects that are different from the background is extracted. A significant overlapping of the receptive fields of the ganglion cells is hardly necessary with this information processing organization as this duplication is demanding and unhealthy for the retina. There is indeed an overlap of receptive fields belonging to the bipolar cells on which the on- and off-centers are formed, but it may be organized in a less demanding way (without using an additional horizontal cell). The horizontal cell may be one and the same for both of the centers.

TABLE 1. The Size of Receptors (Rods and Cones) and their Distribution between the Zones of the Human Eye Retina

Retina Zone	The Cone Diameter, μm	The Rod Diameter, μm	The Number of Cones $\cdot 10^3$	The Number of Rods $\cdot 10^3$	The Cone/Rod Ratio	The Number of the Ganglion Cells
Fovea						
Foveopla	—	—	28	—	—	—
Fovea	2.8–3.2; 3.0–3.6	1.0 1.2	88	44	4:1 1:1	43
Parafovea	4.3	1.4	68	409	1:6	98
Perifovea	4.8	1.5	316	3164	1:10	169
The Total Amount			500	3617	—	310
The Peripheral Retina						
The Near-Peripheral Zone	5.5	1.6	411	7397	1:18	154
The Mid-Peripheral Zone	6.0	1.7	933	23330	1:26	337
The Far-Peripheral Zone	7.2	1.8	4862	82658	1:17	479
The Total Amount			6206	113385	—	970
The Total Number			6706	117000	—	1280

However, in the on-center case, it connects with its output to the inhibition zone of an on-bipolar cell, and in the off-center case, it connects to the central zone of an off-bipolar cell. The cone connection location changes respectively (in the case of an on-bipolar cell, it is connected to the fovea, and in the case of an off-bipolar cell, it is connected to the inhibition input). In other words, the on- and off-centers that are organized based on two corresponding bipolar cells are based on one and the same information coming from one cone and a family of rods that are connected by a single horizontal cell. Thus, their receptive fields not only overlap, but completely converge, and that may have been confusing when determining receptive field boundaries by the coloring method. It is assumed that the receptive fields belonging to the neighboring on- and off-centers are not overlapping or converging in a similar manner. The outputs of an on-bipolar and an off-bipolar cells are connected by the amacrine cells to the on- and off-ganglion cells, respectively, which are bound to the one and the same location of a single cone on the retina. Since it is impossible for both centers to work on one and the same information, either the information of only one ganglion cell is transmitted to the LGN or no information is transmitted at all if the ganglion cell threshold is not exceeded. There is information in the scientific literature, for example, in [3], that on- and off-bipolar cells have different receptor molecules that ensure either their di- or hyperpolarization. In computer vision systems, the signs differentiating the brightness of the center and its surrounds at the output of the bipolar element are used to determine the on- or off-centers. Thus, the sign “+” corresponds to the on-center, while the sign “–” corresponds to the off-center [26].

The efficiency of the retina structure is confirmed by the use of two ganglion cells with the on- and off-centers in one location of the retina with a low impulse frequency. Moreover, instead of using a more demanding double impulse range [3], one cell is discharged in the case of the stimulus strength decrease, while the other is discharged in the case of its increase.

However, the question about the organization of the receptive fields of the bipolar cells with an area-vise large cone and many rods around it remains. In particular, photographs (see Figs. 2 and 3) show a cross-section of rods and cones of the retina, but they are illuminated not from the side of the end of the receptors as is the case with the videosensors of traditional videocameras, but from the opposite side. In other words, the light passes through the network of billions of neurons and connections between them, being reflected and refracted multiple time, and illuminates the side surface of the receptors at an acute angle, which is much larger than the areas of the cross-sections. It is incredibly hard to calculate the area of the complex side surfaces of the rods and cones; therefore, we hope that the ratio of the areas of their side surfaces satisfies the principal conditions of the organization of the concentric excitation/inhibition fields as well. Moreover, it is known that cones have a much lower brightness sensitivity compared to the rods. Thus, the nature may have harmonized these sensitivities by increasing the cone surface. There are no studies concerning this question.

In addition to solving the problem of the harmonization of the excitation and inhibition zone areas for the correct functioning of the on- and off-centers, the organization of the receptive fields of the bipolar cells based on the cones and their surrounds in the form of rods performs another function, namely, determining the balance between the gray and the colorful, which are the opposing components in videoinformation. In other words, when the on- and off-centers are illuminated, the bipolar cell on which they are located determines what kind of light it is, i.e., whether it is colored in some shades of gray or in one of the primary (RGB) colors. If the signal sum from the inhibition zone of a bipolar cell exceeds the signal from the corresponding cone, then the off-center is activated, which testifies that the center is illuminated by the gray light. When the signal from the excitation zone dominates, the on-center is activated, i.e., the light color corresponds to the cone type. It can be assumed that the nature gave the cone the corresponding surface area and sensitivity to the illumination level to ensure the functioning of the on- and off-centers. If the cone type does not correspond to the illumination type and level, a bipolar cell with an off-center is activated, which constitutes gray, or a neighboring bipolar cell with a cone corresponding to the illumination color may be activated. It is worth remembering that there are no colors in nature, but there are different light wave lengths, which are interpreted as color by the human visual system. The light length perception may be completely different in the case of other creatures.

It is known that the rod-based vision has a much higher sensitivity when it comes to perceiving the light in a broad spectral range compared to the cone-based one, although it is unknown whether this is true for all the retina regions as the area of the side surface of cones on the peripheral retina increases much more than that of the rods. Moreover, in the case of very long (exceeding 650 nm) waves, the perception threshold of the photopic (cone-based) vision lowers in comparison to the perception threshold of the rod-based vision. In other words, if a weak light with the wavelength of over 650 nm is sufficient to be seen, it is perceived as colorful (particularly red) [4]. In addition to the light intensity, the light perception threshold is influenced by the surface area of the cone that is stimulated, stimulation time, and light wave length, as well as by the retina regions that are stimulated by the light as different retina regions have different areas of cones and rods and different numbers of rods that create bipolar excitation/inhibition centers. Thus, when constructing a light perception model, all these components have to be considered. Here, some results can be obtained that are of interest for the neurophysiologists and many other experts.

The information on a very large number of cones (more than 800 thousand) and rods (more than 12 million) in the farthest peripheral retina zones with a very small number of ganglion cells (around 10 thousand) is presented in [5, Table 2.1]. If the concentric fields are indeed organized based on the cones, then information from 80 bipolar cells has to be transmitted through each ganglion cell, which is several times more compared to the pressure exuded on other ganglion cells. That is why nature would hardly allow 1% of the ganglion cells to inhibit the work of the other 99%. Here, a situation may take place that is reported on in [3–5, 27], for example, which postulates that the signals coming from the peripheral retina rods are being summed in order to increase the sensitivity under the conditions of insufficient lighting. In this case, summing the signals coming from the bipolar cells using the amacrine cells may coarsen the spatial information in exchange for the brightness sensitivity.

It may seem like the organizational problem of the concentric receptive fields is solved to some extent. However, there arises a question on the possibility to reconstruct the receptive fields, which is stated in [4, 8–11, 24, 27], for example. Many unsolved problems of the oculomotor system organization are written off to the plasticity of the neuron organization and its adaptability to the illumination and contrast levels. However, when reconstructing the receptor fields, there arises a disbalance in the organization of the excitation/inhibition centers as the areas of cones, which are the excitation centers, cannot be changed. They are encoded into the human DNA and are formed throughout the developmental period of a child. This is not the only question about the reconstruction hypothesis of the receptive fields of the retina.

When it comes to the statement regarding the high plasticity of the neural network aimed at adapting to the illumination level (by summing the retina rods under the condition of the insufficient light) and to the contrast (by growing the rod cell layers around the central one in the case of a weak contrast) [7–10], it can be noted that it is hardly expedient. Since different peripheral retina zones have different dimensions of the receptive fields and different light perception sensitivities, the gaze relocation using microsaccades into a more sensitive zone ensures the functioning of the vision system in a broad illumination range. Moreover, the concentric organization of the receptive fields significantly neutralizes changes in the illumination of the whole image when extracting informative features [5]. Thus, the hypothesis on the receptive fields being the tracking system that harmonizes the field dimensions with the present illumination level [8, 10] is hardly adequate as the differences in the image illumination and not the absolute illumination are what is important for the brain. A slow adaptation to the illumination level is performed without reconstructing the receptive fields, but through electrophysiological processes.

The idea of the structural and functional reconstruction of the receptive fields, which is embedded in the dynamic process models [8, 10], constitutes the formation of the maximum feasible excitation zone of the receptive fields for each ganglion cell in the retina. In what follows, the formation of a much larger inhibition zone takes place with the aim of finding the maximum large light or dark spot differing from the background as an informative feature. After that, the excitation zone is gradually reduced. That allows us to highlight smaller spots in the center of the receptive field of a ganglion cell. However, this organization of the informative feature extraction does not allow for the extraction of lesser features, which are situated on the different sides of the center of the receptive field of a ganglion cell. It is unlikely that the luxury of organizing large excitation and inhibition centers with a large number of connections with receptors and with long lines for the bipolar cells by using the interplexiform (horizontal) cells, as well as with a considerable overlapping of the receptive fields of the ganglion cells could be allowed by nature. Moreover, it is unknown what governs these processes. Under the condition of a significant overlap of the receptive fields of the ganglion cells, a large percentage of the neighboring receptive fields is repeated, which is unbelievably demanding. Additional problems arise if a spot is situated at the intersection of the receptive fields of a couple of ganglion cells as then it is unclear how to combine them and how to transmit information about them.

An additional argument against the necessity to reconstruct the receptive fields under the conditions of insufficient illumination or contrast is the existence of a vast number of tools and mechanisms embedded in the structure and organization of the human eye retina and the oculomotor system, which allow for the brightness perception range of 10–12 orders of magnitude, which is inaccessible for the technical means. These are the following tools and mechanisms:

— controlling the pupil of the eye in accordance with the illumination level and changing the form of the eye lens in order to focus on the chosen object;

— using two receptor types (cones and rods) with a different sensitivity;

— increasing the diameter of cones and rods and, consequently, their surface areas from the fovea to the far retinal periphery, which increases their sensitivity by several orders of magnitude;

— using the concentric organization of the receptive fields, which allows for a significant increase of the sensitivity to the extracting the pinpoint informative features;

— the existing zones with the heightened sensitivity to the contrast (the far- and mid-peripheral zones) in the peripheral retina, which make sure that an insignificant eye movement (microsaccades) ensures that the determined fragment finds its way into this heightened sensitivity zone;

— summing the signals from the bipolar cells in the farthest peripheral retina zone, which increases the brightness sensitivity in exchange for the spatial resolution (needs confirmation).

Thus, the vision system has enough tools and mechanisms to ensure a high sensitivity to illumination and contrast without the reconstruction of the receptive fields.

THE RIGID ORGANIZATION OF THE RECEPTIVE FIELDS TAKING INTO ACCOUNT THEIR BINDING TO THE LOCATION ON THE RETINA

The brain possesses information on the organization of the receptors of the peripheral retina (the smaller receptor fields are situated closer to the fovea, the smallest ones are situated in the fovea zone, and the large ones are situated at the retina periphery, and the peripheral retina may have additional specialized information processing zones). All of them are topometrically connected to the corresponding ganglion cells, as well as to the sensory and motor areas of the higher brain levels through the LGN.

A number of patterns was discovered in the visual system structure, which substantially ordered its organization. The order is present both in the organization of connections in the visual system as a whole, as well as in the determined spatial organization of connections of single neurons. Here, a couple of stages in the structure order can be distinguished.

1. The structural order of the retina and of the visual system as a whole is reflected in the determined and quite rigid system of connections between certain system components, which allows us to highlight a couple of signal processing levels (the concentric organization of the receptive fields of the bipolar cells based on cones as the excitation zone and on a family of rods as the inhibition zone, as well as the connection of bipolar cell families with the corresponding ganglion cells under the control of the amacrine cells).

2. The projective and topical connection order, which remains at all the vision system levels, allows us to bind the results of the information processing at each level to their location on the retina.

The development of the cellular elements of the retina and their connections are rigid in their genetical determination and hardly dependent on the external influence. The synoptic links between retina cells of an adult human remain pretty conservative.

Since the ganglion cells of the peripheral retina are bound to the location of the bipolar cells and topometrically transmit the information to the LGN, the visual cortex, and the higher sensory and motor brain levels as a result, a conclusion can be reached that each ganglion cell has to have opportunities to extract informative features on the contrast differences from the bipolar cells of its receptive field simultaneously with other ganglion cells by transmitting these features consecutively in time. Such opportunities can be provided by the amacrine cells, for example, which have a very extensive horizontal and vertical network of processes connecting the bipolar and ganglion cells for communication and control of their functions.

The amacrine cells may consecutively, according to the order determined in advance by the brain, connect the bipolar cell outputs to the ganglion cell inputs of the corresponding receptive fields, which transmit this information to the LGN. Here, it is possible that the amacrine cells determine the (on- or off-) bipolar cell whose information has to be transmitted or, alternatively, when a zero value has to be transmitted if the bipolar cell has not detected neither the increasing nor the decreasing illumination difference at the chosen point. During the transmission process, the priority mechanism or the information transmission only from the bipolar cells exceeding the ganglion cell threshold are hardly utilized as it necessitates the transmission of the addresses of the corresponding bipolar cells. In other words, the information from the bipolar cells on the magnitude of differences at the corresponding points of the receptive field of a ganglion cell is transmitted consecutively and in the determined order, and all the ganglion cells are transmitting their information simultaneously. The transmission of the illumination difference values is partially performed via pulse frequency spikes or by using the time interval between the impulses during spikes with a precise binding of the corresponding bipolar cells on the retina.

A ganglion cell is connected through amacrine cells with many bipolar cells where its on- and off-centers are organized. Moreover, amacrine cells, which work according to the program inherent to them, connect the corresponding bipolar cell outputs and can successively transmit the analysis results of the pinpoint illumination intensity differences to the inputs of the ganglion cell that are within its receptive field. The transmission of information from the bipolar cells to the ganglion cells is performed consequently over time and with a clear synchronicity with the higher brain levels that know that sequence. The sequence has not been studied yet. Since the rods perform a more passive role compared to the cones and their information is not transmitted to the higher brain levels, it is presumed that the visual acuity should be determined not according to the number of rods (≈120 mil), but according to the number of cones (≈8 mil) taking into account, of course, the high cone density in the fovea and their low density at the periphery.

It seems as if it would be easy to distinguish straight line sections (under the concentric organization of the receptive fields, the line start and end are easily distinguished compared to the intermediate points on the line), their orientation, etc. [4, 7, 8, 16] in the case of the above organization of the bipolar and ganglion cells through amacrine cells. However, according to this hypothesis, a lot of additional information on the addresses of the start and end of the line or its length with it being bound to the location, a multitude of different variants of its orientation, etc. have to be transmitted. Thus, this hypothesis is hardly feasible. It is more likely that the results from the bipolar cells, which either passed or did not pass the ganglion cell threshold, are consequently transmitted using the amacrine cells. Here, the value of the brightness difference is transmitted by the corresponding frequency or by the time interval between the spike impulses. This is the ultraprecise and the most reliable information on the contrast between the image elements and on their location on the retina for the future determination of boundaries, straight line sections, their orientations, etc. This may take place in the LGN and the higher brain layers.

The refraction period between the action potentials (spikes) of a ganglion cell in the case of crossing its neural threshold at the output of the bipolar cells is around 1 ms [4]. In other words, this cell can transmit information, for example, from its 30 bipolar cells, in 30 ms and not significantly impact the functioning of the general information perception of the human visual system. Moreover, the existence of the consecutive processes of information transmission from the retina may make one think that the process of the spatial information perception is not continuous. However, the process is discretized in time, and the discretization time is a couple of tens of frames per second (obviously depending on the maximum number of bipolar cells in the receptive field of the ganglion cell that has to transmit their information). In the case of the peripheral retina, this process certainly has to be clearly synchronized for all the ganglion cells.

The synchronization process of the information transfer may be different for the fovea, which has significantly smaller receptive fields and a small number of bipolar cells, respectively, as it fulfils completely different functions and its information practically never overlaps with the information from the peripheral retina. In the case of the fovea, the distances and the saccades are shorter, and they can be performed with a higher frequency for the quick estimation and recognition of objects in the focus zone.

There is not enough information on the number and location of the different types of cones in the fovea, as well as on their organization. Based on the cone to rod ratio (Table 1), the following conclusions can be drawn:

— in the foveola and fovea zones, the rods are either absent or their number is insufficient to organize concentric excitation/inhibition zones; the rods may be performing the “supporting” function in the case of the insufficient illumination for the cones, i.e., a rod will reflect a shade of grey so there is no black spot;

— in the parafovea and perifovea zones where the rod to cone ratio is increased to ten, the organization of the concentric excitation/inhibition zones is possible in order to extract the informative features in color or in the shades of grey similarly to the peripheral retina.

There are many confirmations of the existence of the models corresponding to the Young–Helmholtz trichromatic theory and the second Opponent Process Theory of Hering in the human vision system in [3–6]. However, the absence of information on the organization of cones in the fovea does not allow us to propose any hypothesis on the possibility of their implementation in the retina. The theories may be implemented at the higher brain levels.

A couple of organizational models of the trichromatic theory are developed for the sensors of the computer vision system. The most important model for the sensor construction is the Bayer color filter where three primary colors are located in the sensor matrix as follows:

$$\begin{matrix} r & g & r & g & r & g & r & g & r & g \\ g & b & g & b & g & b & g & b & g & b \\ r & g & r & g & r & g & r & g & r & g \\ g & b & g & b & g & b & g & b & g & b \end{matrix}.$$

Taking into account the heightened sensitivity of the human eye to green, its amount in the sensors is increased by three. The three primary colors for the on-center [24] can be extracted from the sensor using the Laplace masks of the size 3×3. To extract the (exciting) blue color B using the inhibitory (opponent) colors, we use r and g; for the red color R, the colors b and g are inhibitory; therefore, to extract them, the 8-connected mask is used

−1	−1	−1
−1	8	−1
−1	−1	−1

For the green color, G, r, and b are inhibitory colors; however, here, the 4-connected Laplace mask is used

	−1	
−1	4	−1
	−1	

In the case of the off-center, the mask signs are reversed, which completely corresponds to the trichromatic theory of color vision. The yellow color Y is additionally extracted in accordance with the formula $Y = (r+g)/2 - |r-g|/2 - b$ [24].

In what follows, pairs of opponent colors RG and BY are considered, for example, where R is the center and G is the surround, and where B is the center and Y is the surround or vice versa. In other words, both theories are functioning and used in computer vision systems.

The logarithmic light brightness perception by the human vision system hardly influences the above oculomotor movement system as, in the case of small (up to three orders of magnitude) illumination differences, the light is approximately perceived as linear for the general light perception range of 10–12 orders of magnitude.

OPERATION MODES OF THE HUMAN VISION SYSTEM AND THEIR SPECIAL FEATURES

The videoinformation perception process largely depends on the goal determined by the brain. A human perceives the world around them through the retina working together with the eye movements. When it comes to the videoinformation perception process, the following problems (operation modes) of the vision and the oculomotor movement systems can be distinguished:

— a passive viewing of the surrounding world;

— searching for an object in an image;

— following the object in the videocontinuity;

— a detailed consideration, recognition, classification, and measuring of the object;

— analyzing the scene.

The implementation of these problems is controlled by the eye movements that are either reflexive or governed by the brain. The following different eye movement types are known:

— saccades or different eye movements that turn the gaze from one object to another. Large saccades ensure that the gaze turns by $20°$ and more, minisaccades ensure the turn by $20°$ to $3°$, and microsaccades ensure the turn up to $3°$. Saccades are mostly reflexive movements, but they are planned in advance as a result of the acquired experience of the nervous system in object searching;

— tracking eye movements take place when the object is moving against an unmoving background. They are slow; thus, in this case, the stimulus is not the object itself, but its location, movement direction, and movement velocity. A relatively stable location of the object on the retina and its sharper perception is ensured through the synchronization of the eye movements with the object;

— tremor-like eye micromovements arise when the onlooker consciously focuses their gaze on an object, which ensures a constant excitation based on the image illumination differences as images disappear with time otherwise;

— other (vestibulo-ocular and vergence) eye movements, which ensure that the gaze stays fixed on the object during a turn, as well as the movement coordination of both eyes, are not considered in this paper.

Although there is not enough information for a detailed description of the different operational modes of the vision system as it is incredibly difficult to study dynamic processes, the organization and functioning of the retina is considered with taking into account the special features of these modes, because the retina functions differently in them.

The visual system operation algorithm has its own features in each operation mode, and the brain uses a large toolkit of means, techniques, and methods of adaptation and influence on the perception processes through the corresponding eye muscles for its implementation. These are the peripheral retina and the fovea with the receptors (the rods and the cones) of a different spatial resolution, sensitivity, and possibilities to perceive light in different wavelength ranges (achromatic and colored visions), the neural network extracting informative features from an image, and the higher brain layers.

The retina obtains dispersed points of an image (a mosaic image) through the rods and cones. According to all the data, this process is continuous. However, a visible image is created as a result of the multistage information processing by the layers of the retina neurons and the higher brain layers. In other words, the image awareness takes place in tens of ms, and the process is discrete both in time (the ganglion cells consecutively transmit the information from their bipolar cells) and space (a restricted number of bipolar cells), as well as in the signals on the contrast magnitude (the existence of ganglion cell thresholds).

When passively (without a set goal) observing one's surroundings from a window of a moving vehicle, for example, the information does not fully enter our consciousness. However, when an object or a situation that may capture one's attention enters one's field of vision, the turn of one's head or eyes is performed by a reflexive command issued by the brain, which inhibits the stimulus moving across the retina and ensures the possibility of a detailed consideration of the above object or situation.

The search for an object in an image is performed by the peripheral retina. However, statements that an excitation in some location on the retina lead to the gaze redirection by saccades into the fovea to analyze the information hardly explain this situation as it is unclear what exactly is an excitation (there are many of them on the retina), if there are any priorities, etc. It is unlikely that there is a need to redirect one's gaze into the fovea while searching for an object as the fovea is designed for a detailed object consideration. The extraction of the informative features of all the objects that are the center of attention takes place on the peripheral retina simultaneously due to the corresponding organization of the bipolar and ganglion cells using the horizontal and amacrine cells. The enhanced informative features are transmitted onto the higher brain levels from the LGN where they are contrasted against the images of objects in the memory to extract what is searched for, presumably according to the priority in accordance with the acquired experience of the viewer. To clarify the information on the searched for object or its detailed consideration, the corresponding oculomotor muscles are instructed to redirect the gaze to the selected object for it to find its way in the fovea [24].

The tracking of an object in the videocontinuity can be performed using the rigid features from the peripheral retina without engaging the fovea. Moreover, due to the tracking eye movements, the actual location of the object can be determined as well.

The object recognition is performed in the fovea based on the more detailed spatial information extracted by the retina of the fovea, as well as on the higher brain levels based on the additional colored information.

Although a much smaller information amount could be read and processes in the above operation modes, it may be necessary to constantly transmit all the current information from the retina to be processed by the brain as it is necessary to safeguard the eye and the person [29].

In addition to distinguishing objects in a two-dimensional image, the information on their location in a three-dimensional space has to be extracted to analyze a scene. This is achieved by using information from both eyes and by coordinating their movements. However, this problem is not considered in this article.

Thus, the set of eye movements in each mode is not chaotic, but stipulated by the set goal taking into account the object characteristics, the experience of the person, and other factors.

It is obvious that not all of the questions concerning the retinal organization are considered in this article. There are many of them left and there is not enough information for their consideration. In particular, the following questions are not clear yet:

— are there concentric fields of excitation/inhibition on the ganglion cells?

— are the amacrine cells able to perform signal summing by connecting the outputs of a couple of bipolar cells to a single ganglion cell input to increase the sensitivity in exchange for spatial capacity or can they connect the output of a single bipolar cell to a couple of ganglion cell inputs for the rise of sensitivity and a possible organization of the concentric fields of the ganglion cells on uniform elements (bipolar cells)?

— is the whole vision system information transmitted to the brain to be analyzed in all the modes?

— how does the sensitivity of the rods and cones increase when their dimensions increase on the peripheral retina (here, the diameter squares of their cross-sections are insufficient, and the areas of their side surfaces have to be accounted for)?

CONCLUSIONS FOR DISCUSSION

A detailed analysis of the issues concerning the organization of the human eye retina from the viewpoint of cybernetics is performed in the article. In particular, the following was achieved:

— the conditions of the correct functioning of the concentric receptive fields are determined and a hypothesis on the organization of on- and off-centers based on the bipolar cells with a cone as the central excitation zone and the surrounding rods as the inhibition zone with the use of the horizontal cells for the on-center, as well as vice versa for the off-center, for both of which the functioning conditions of the concentric receptive fields are satisfied and a significant (by tens of times) reduction of the information volumes is ensured, is formulated;

— the conditions of extracting informative features based on the on- and off-centers in the form of pinpoint values of brightness differences between the neighboring elements are justified, as well as the possibility of the organization of both of the centers based on a single cone and a single horizontal cell with a precise binding to the location on the retina;

— the rigid organization of the receptive fields of the ganglion cells and the limited possibilities of its reconstruction are justified;

— a hypothesis concerning the organization of the transmission of the brightness difference values from the bipolar cells through the ganglion cells under the control of the amacrine cells with a binding to the location on the retina is proposed;

— the fact that it may be necessary to constantly transmit all the current information from the retina and for it to be processed by the brain in order to safeguard the eye and the person is noted;

— a hypothesis concerning the restriction of the visual acuity not by the retina rods, but by its cones while taking into account their location density on the retina is proposed;

— a hypothesis concerning the continuous image perception by the retina receptors and the discrete information perception (in time and space, as well as according to the contrast magnitude) by the brain is proposed;

— features of the functioning of the vision system in different modes, which are ensured by the different eye movements and by the set goal, are considered.

Closing words

Dear colleagues! Not all of my hypotheses and results may be confirmed (as this is a very complex subject). Therefore, I will be thankful for all the critical remarks serving the goal of establishing the truth.

REFERENCES

1. V. M. Glushkov, Cybernetics. Issues of Theory and Practice [in Russian], Nauka, Moscow (1986).
2. V. P. Boyun, Dynamic Information Theory. Fundamentals and Applications [in Russian], V. M. Glushkov Institute of Cybernetics, Kyiv (2001).
3. D. H. Hubel, Eye, Brain, and Vision, W. H. Freeman & Co. (1988).
4. H. R. Schiffman, Sensation and Perception: An Integrated Approach, 5th ed., John Wiley & Sons (2001).
5. N. Lauinger, The Human Eye: An Intelligent Optical Sensor (The Inverted Retina: A Diffractive-Optical Correlator), IFSA Publishing, S. L., Toronto (2014). URL: http://www.sensorsportal.com/HTML/BOOKSTORE/Human_Eye.htm.
6. H. Kolb, “How the retina works: Much of the construction of an image takes place in the retina itself through the use of specialized neural circuits,” American Scientist, Vol. 91, No. 1, 28–35 (2003).
7. A. Ya. Supin, Neuronal Mechanisms of Visual Analysis [in Russian], Nauka, Moscow (1974).
8. I. A. Shevelev, Neurons of the Visual Cortex: Adaptability and Dynamics of Receptive Fields [in Russian], Nauka, Moscow (1984).
9. F. Werblin and B. Roska, “The movies in our eyes,” Sci. Am., Vol. 296, No. 4, 72–79. (2007). https://doi.org/10.1038/scientificamerican0407-72.
10. N. F. Podvigin, F. N. Makarov, and Yu. E. Shelepin, Elements of the Structural-Functional Organization of the Visual-Oculomotor System [in Russian], Nauka, Leningrad (1986).
11. V. P. Boyun, “Human visual analyzer as a prototype for designing a family of problem-oriented machine vision systems,” in: Proc. Intern. Sci. and Techn. Conf. “Artificial Intelligence. Intelligent Systems (AI-2010),” Vol. 1, IPII “Nauka i Osvita,” Donetsk (2010), pp. 21–26.
12. A. I. Kubarko and N. P. Kubarko, Vision (Neurophysiological and Neuro-Ophthalmological Aspects), Vol. 1, Sensory Functions of Vision and Its Disorders in CNS Disease [in Russian], BGMU, Minsk (2007).
13. Yu. G. Kratin, N. A. Zubkova, V. V. Lavrov, T. S. Sotnichenko, and K. P. Fedorova, The Visual Pathways and the Activation System of the Brain [in Russian], Nauka, Leningrad (1982).

14. S. Shah and M. D. Levine, “Visual information processing in primate cone pathways. I. A model,” IEEE Trans. Syst. Man Cybern., Part B (Cybern.), Vol. 26, No. 2, 259–274 (1996). https://doi.org/10.1109/3477.485837.
15. A. Benoit, A. Caplier, B. Durette, and J. Herault, “Using Human Visual System modeling for bio-inspired low level image processing,” Comput. Vis. Image Underst., Vol. 114, Iss. 7, 758–773 (2010). https://doi.org/10.1016/j.cviu.2010.01.011.
16. J. R. Anderson, Cognitive Psychology and Its Implications, 5th ed., Worth Publishers, New York (2000).
17. Yu. E. Shelepin, V. M. Bondarko, and M. V. Danilova, “Foveola construction and visual system pyramidal organization model,” Sensory Systems, Vol. 9, No. 1, 87–97 (1995).
18. P. J. Burt, “Smart sensing within a pyramid vision machine,” Proc. IEEE, Vol. 76, No. 8, 175–185 (1988).
19. V. P. Boyun, “Intelligent selective perception of visual information. Informational aspects,” Artificial Intelligence, No. 3, 16–24 (2011).
20. V. Boyun, “Intelligent selective perception of visual information in vision systems,” in: Proc. 6th IEEE Intern. Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Application (IDAACS’2011) (Prague, Czech Republic, September 15–17, 2011), Vol. 1 (2011), pp. 412–416.
21. V. Boyun, “Bioinspired approaches to the selection and processing of video information,” in: Proc. 2018 IEEE Second Intern. Conf. on Data Stream Mining & Processing (DSMP) (Lviv, Ukraine, August 21–25, 2018), IEEE (2018), pp. 498–502. https://doi.org/ 10.1109/DSMP.2018.8478541.
22. R. C. Gonzalez and R. E. Woods, Digital Image Processing, 2nd ed., Prentice Hall (2002).
23. D. Forsyth and J. Ponce, Computer Vision: A Modern Approach, Prentice Hall (2003).
24. V. Boyun, “The principles of organizing the search for an object in an image, tracking an object and the selection of informative features based on the visual perception of a person,” in: S. Babichev, D. Peleshko, and O. Vynokurova (eds.), Data Stream Mining and Processing, DSMP 2020; Commun. Comput. Inf. Sci., Vol. 1158, Springer, Cham (2020), pp. 22–23. URL: https://link.springer.com/chapter/10.1007/978-3-030-61656-4_2.
25. D. A. Marr, Vision: Computational Investigation into Human Representation and Processing of Visual Information, W. H. Freeman and Company, New York (1982).
26. V. Boyun, “Directions of development of intelligent real time video systems,” Appl. Theor. Comput. Technol., [S.l.], Vol. 2, No. 3, 48–66 (2017). https://doi.org/10.22496/atct.v2i3.65.
27. V. P. Boyun, L. O. Voznenko, and I. F. Malkush, “Principles of organization of the human eye retina and their use in computer vision systems,” Cybern. Syst. Analysis, Vol. 55, No. 5, 701–713 (2019). https://doi.org/10.1007/s10559-019-00181-0.
28. V. P. Boyun, “Real-time video systems with elements of artificial intelligence: Transcript of the report at the meeting of the Presidium of NAS of Ukraine, October 13, 2021,” Visn. Nac. Acad. Nauk Ukr., No. 12, 77–84 (2021). https://doi.org/10.15407/visn2021.12.077.
29. A. P. Kirpichnikov, “The eye is an intellectual sensor,” Datchiki i Sistemy, No. 12, 63–69 (2009).