JOINT STUDY OF VISUAL PERCEPTION MECHANISM AND COMPUTER VISION SYSTEMS THAT USE COARSE-TO-FINE APPROACH FOR DATA PROCESSING

Anton Sharypanov, Alexandra Antoniouk, Vladimir Kalmykov

***Abstract**: Aspects of visual perception mechanism and pattern recognition methods are examined jointly. Latest results from neurophysiology in studying the visual system of living beings are discussed. Another view on coarse-to-fine approach for technical vision tasks is presented. On the basis of systemological analysis of neurophysiology sources a new hypothesis about visual neuron's functioning is proposed. This hypothesis explains the mechanism and takes into account receptive fields excitatory zones resizing during visual act.*

***Keywords**: coarse-to-fine, neurons of visual system, intercellular processes, pattern recognition, image processing, variable resolution*

***ACM Classification Keywords**: I.4.1 Digitization and Image Capture, I.5.1 Models*

Introduction

State of the art in technical vision, various approaches to pattern recognition, image processing and object detection tasks shows that researchers frequently meet the problem of great computational complexity, particularly – the problem of great time and machine resource consuming. One of the strategies that solves these problems is so-called "coarse-to-fine approach", i.e. the technique of refining initial data that exclude inappropriate objects or irrelevant ranges of the image on earlier stages of processing in order to apply computationally intensive part of the algorithm to reduced volume of data. One of the variants of this approach is to use a coarse-to-fine modification of known algorithm, suitable for solving the particular class of problems.

For example, in [1] proposed a multi-resolution part based model and a corresponding coarse-to-fine inference algorithm which is extremely efficient. The method is based on the observation that matching of each part of the image is the most expensive computational operation in comparison to detection of significant parts and computation of their optimal configuration, so the minimization of number of part-to-image comparisons implies detection acceleration. Starting from matching the lowest resolution part the method selects only the best placement in each image neighborhood. These locally optimal placements are then propagated recursively to the parts at higher resolution. By recursive elimination of unlikely part placements from the search space, the set of possible locations is narrowed so that the computation of only few part-to-image comparisons is performed. This method gives a ten-fold speed-up over the standard dynamic programming approach.

The algorithm in [2] follows the same idea of discarding of large regions from the hypothesis space at an early stage of the recognition however the object detector for each resolution is not the same. Each detector uses inexpensive tests on the image features and narrows down the set of possible variants to a smaller number, which are explored in greater detail by the next detector. Also each detector computes a quantity for each region of hypothesis space; this region is accepted for consideration at the next level of resolution if the quantity exceeds a given threshold. All thresholds are set automatically based on probabilistic measurements.

Coarse-to-fine strategy application to vehicle motion trajectories clustering is discussed in [3]. Raw trajectories are clustered into “coarse clusters”. Each “coarse cluster” consists of trajectories having very similar directions, but with different positional characteristics. The feature for further fine clustering is trajectory resampling point set, and the Euclidean distance is used as the distance measure between two trajectories.

For face recognition, a coarse-to-fine procedure can be implemented by sequential application of different face recognition methods in order to reduce the candidate examples on each step. In [4] the decision process is developed through consecutive stages, such as “one-against-all (OAA) of SVM (support vector machine)”, “one-against-one (OAO) of SVM”, “Eigenface”, and “RANSAC”. The stage 1 “OAA of SVM” and stage 2 “OAO of SVM” uses the discrete cosine transform features extracted from the entire face image. On stage 3 “Eigenface” face images are projected onto a feature space (face space). The face space is defined by the “Eigenfaces”, which are the eigenvectors of the set of faces and based on intensity information of the face image. “RANSAC” is applied in the last stage, in which the epipolar geometry method with space information of the testing image is matched with the two training images, and then the image with the greatest match numbers of and the shortest distance to corresponding feature points is selected.

The task of establishing the correspondence between pixels in two images (finding a markup) with human faces, addressed in [5], is effectively solved by building “cascades” of markups. The resolution is decreased two times in both initial images per cascade and new markup for them is built. After that the starting approximation for initial markup is defined based on the new markup and the field of motion is searched but with less quantity of markings. The algorithm that solve the task utilizing one “cascade” runs eight times faster while preserving accuracy in finding the field of motion for two images.

Examples of dynamic programming (DP) application to recognition problems are numerous and include speech recognition, character recognition, deformable template matching, soft decoding and road tracking. However such problems often lead to enormous state spaces and the computations can be infeasible, even with DP. To overcome this obstacle, in [6] proposed a variation on DP - coarse-to-fine dynamic programming (CFDP). The essential idea of this algorithm is to form a series of coarse approximations to the original DP trellis by aggregating trellis states into “super states”. For each coarse approximation, the optimal path is found using DP with “optimistic” arc costs between the super states. The super states along this optimal path are refined and the process is iterated until a demonstrably globally optimal path is found. In many cases this global optimum is achieved with considerably less computational expenditure than straight DP. This CFDP algorithm is particularly well-suited to DP problems with large state spaces. According to [7], the speed of CFDP depends on the structure of the grouping and the nature of the problem. In the best case, CFDP gives a large computational savings over standard DP; in the worst case, it will actually be slower.

The goal of coarse-to-fine approach in mentioned cases is to reduce intensive processing only to some regions of starting image or to some parts of initial dataset containing the information that seems to be useful and no mater what the particular coarse-to-fine mechanism is used.

But many of the image recognition problems that result in NP-complete tasks or even can not be addressed with traditional methods are solved by human visual system in less than no moment. It definitely looks like one can benefit from studying the processes taking place there to reach the level of performance comparable to that of the visual system. In previous decades researchers have already tried to examine some aspects of image processing in conjunction with contemporary results in neurophysiology.

Investigations on changes in spatial sizes of direction-selective cells in the primate visual system [8] has inspired Battiti et al. to study how integrating motion information across different spatial scales could help improving the estimate of the optical flow. An adaptive multiscale method, where the discretization scale is chosen locally according to an estimate of the relative error in the velocity estimation, based on image properties was proposed in [9]. This coarse-to-fine method provided substantially better estimates of optical flow than did conventional algorithms, while adding little computational cost.

It will be shown in the following paragraphs that studying the processes taking place before V1 visual cortex seems to be useful for developing new efficient methods in technical vision systems construction.

On mechanism of visual perception

Objects of concern are retinal ganglion cells and LGN neurons found in visual system (visual neurons) and their receptive fields. It was discovered [10, 11] that the sizes of receptive fields' excitatory zones change during the visual act, which eventually mean dynamical changes in visual system's resolution. It is known that the receptive field of a visual neuron consists of many receptors that send signals to it through one or more synapses [12]. Receptive field is circular in simplest case. The functioning of such visual neuron is studied for its action potentials in response to stimulus – a circle on image, contrasting relative to the background, which is projected onto the retina at a given time interval.

Human visual system operates in a sequence of visual acts that last for ~150ms each. After that period a saccade (oculomotorius muscles' twitching) happens, the image on retina shifts and the next visual act begins.

The fact that the size of receptive fields' excitatory zones does not stay stationary during visual act was established in the course of research work in [10]. Excitatory zone diameter is the minimum diameter of stimulus at which the maximum number of neuron responses is achieved. In order to get quantitative description of changes in receptive field's excitatory zone the time slices method was proposed. This method is based on the assumption that if the diameter of excitatory zone changes during visual act, then the maximum number of spikes from this neuron matches the time interval when diameter of excitatory zone meets the stimulus diameter or their diameter values varies insignificantly.

The time slices method for obtaining the diameter of excitatory zone as a function of time is used as follows. Post-stimulus histograms (PSH) that represent the responses of neuron on circular stimuli of different sizes were divided into a series of sequential temporal intervals (7.5 or 15ms each) and the number of spikes in each interval was determined (Figure 1 [10]). The time slices method also allows determining the serial number of temporal interval with maximal neuron responses. The bigger the size of stimulus, the earlier maximal response is achieved, so one can conclude that the diameter of round receptive field excitatory zone of a neuron changes in the course of visual act, namely it shrinks from maximal to minimal, up to 1-2 receptors in the case of ganglion cell. Thus there exists maximal resolution for visual system defined as the number of receptors in the field of view center and the variable resolution that changes during visual act. Variable resolution is determined by the size of neuron's receptive field excitatory zone.

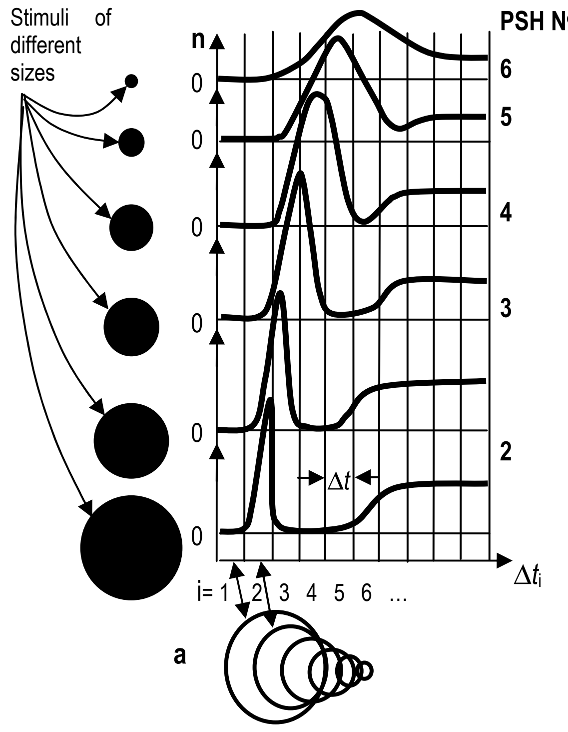


Figure 1 [10]. Neuron responses on stimuli of 6 different sizes
n axes represents the number of spikes in corresponding time slice Δt_i.
Maximum number of spikes corresponds to interval where the size of stimulus meets the excitatory zone size.
a – decreasing of receptive field's excitatory zone area during visual act.

Temporal response pattern and dynamics of receptive field structure in single lateral geniculate nucleus (LGN) neuron of a cat using static spot stimuli flashed on the receptive field for 400–500 ms were studied in [11]. Spatial receptive field parameters from spatial summation curves, determined for successive 5-ms intervals throughout the stimulus period, were estimated (Figure 2 [11]). Thereby it was possible to study dynamics of the response properties during periods similar to those in natural fixations, and with a method that did not presuppose a linear system. The results showed pronounced changes in the receptive field structure during the spot presentation.

Initially, the neurons had wide receptive field centers. The center rapidly shrank to a minimum that occurred on average ~70 ms after stimulus onset whereupon the center widened slightly [11]. Thus the maximum spatial resolution occurred in a brief time window after onset of stimulation. In parallel, the center-surround antagonism increased. The changes in spatial resolution did not follow the changes of firing rate. The initial strong burst of action potentials appeared earlier than the maximal spatial resolution. The authors state that these results are consistent with the hypotheses that the firing pattern of the neurons during brief static stimulation initially mediates a strong but spatially coarse message to cortex that gradually changes into a weaker, but spatially more precise message.

For all the nonlagged cells, there was a pronounced change in the selectivity for spot size during the spot-on period. At the beginning of the period, the neurons responded well to a broad range of spot sizes, but subsequently, the response was restricted to a gradually narrower range of the smaller spots.

The quantitative estimates of the receptive field center width revealed pronounced changes during the stimulus presentation in all the nonlagged cells, in particular during the first 150 ms after spot onset. The receptive field center was initially wide, but rapidly shrank to a minimum. In the majority of neurons, the center thereafter widened again such that the minimal size occurred only briefly.

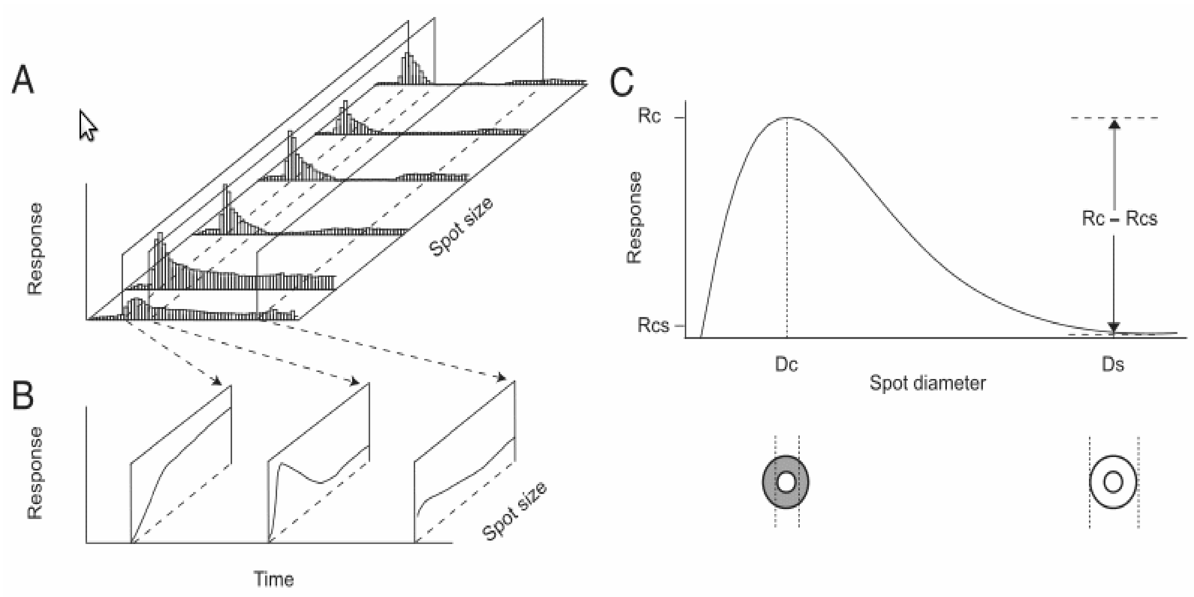


Figure 2 [11]. Schematic illustration of method for studies of dynamics of receptive field properties

A: peristimulus time histogram (PSTH) with 5-ms bin width was determined for response to each of a series circular spot stimuli presented on receptive field for 400-500 ms. Spots differed in width from smaller than receptive field center to larger than whole receptive field. PSTHs for different spot sizes were used (shown by 6 schematic PSTHs). Time slices through whole set of histograms were made for each 5-ms bin, and set of response vs. spot width values in each time slice was used to make a spot width tuning curve for respective time after stimulus onset , shown by 3 schematic tuning curves in B. C: in each tuning curve, width of spot that gave maximum response (R_c) was taken as estimate of width of receptive field center (D_c), width of smallest spot that gave minimum response (R_{cs}) was taken as estimate of width of receptive field surround (D_s), and reduction of response from maximum to minimum divided by response maximum was taken as estimate of center-surround antagonism at respective time.

Contrary to that the lagged cells showed no clear dynamic changes in the spatial structure of the receptive field. When the response appeared, the neuron already had a small receptive field center [11].

Also the hypothesis that a small stimulus spot will initially activate many neurons, most of them only transiently was checked (Figure 3 [11]). The conditions when the stimulus falls outside of the receptive fields centers during their shrinkage were modeled. For all three positions, there was an initial transient response, but the subsequent sustained response that occurs to a small spot in the minimum receptive field center was lacking in these cases.

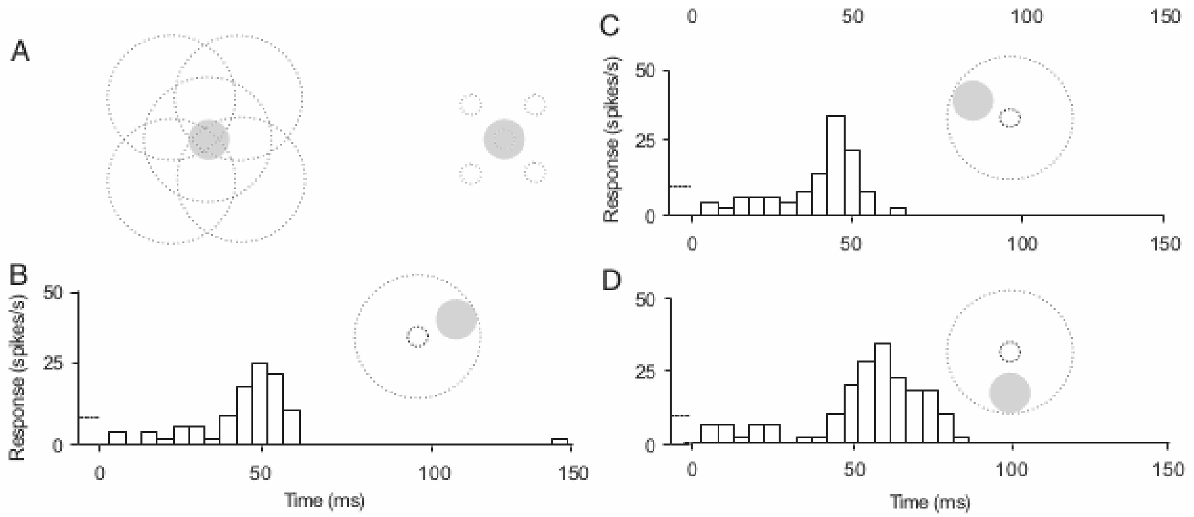


Figure 3 [11]. Initially wide RFCs suggest that number of activated neurons decrease during stimulation

A: schematic illustration of hypothesis. B–D: eccentric stimulus spots presented outside minimum field center (small center circle) but inside maximum field center (large circle) gave a fast initial response that disappeared as center shrank toward minimum. Only responses over the 1st 150 ms of a 500-ms stimulation period are plotted. Dotted line before stimulus onset shows level of spontaneous activity. Nonlagged cell. Bin width in histograms, 5 ms. Number of spot presentations at each location (interleaved) was 100.

The conducted experiments with combined S-potential and action potential recordings showed that the initial fast shrinkage of the receptive field center was present already in the retinal input to LGN neurons. The degree of shrinkage was similar for the retinal input and the LGN neuron, and apart from the faster shrinkage in the LGN neurons, the temporal pattern of the shrinkage was also similar.

However the authors mention that mechanisms for generation of center width dynamics in retinal ganglion neurons are unclear, so further studying of intracellular processes is needed.

The fact that an optimal visual stimulus flashed on receptive field center in retinal ganglion cells typically evokes a strong transient response followed by weaker sustained firing and so happens in thalamocortical (TC) neurons of LGN in a state-dependent manner is well-known and described in literature, but the mechanisms by which transient firing changes to sustained are less well known. One of the results obtained in [13] shows that constant frequency pulse train stimulation of retinal afferents causes depolarization through temporal summation of excitatory post-synaptic potentials (EPSPs) in TC neurons (Figure 4 [13]) which occurs no matter what the holding potential of the cell membrane was set but the value of the holding potential influenced the ability of TC neurons to generate spikes in later part of the train.

In this study the holding potential of a cell was adjusted to different steady state values by direct current injection through recording electrode at the beginning of each experiment and didn't change during particular measurement. The authors suggest that regulation of the sustained response through the level of the membrane potential is a key mechanism for regulation of the strength of input to cortex depending on states like arousal, attention etc.

Another aspect of coarse-to-fine approach in computer vision

Researches in the field of visual perception physiology and creation of information technologies for automatic processing of visual information (technical or computer vision in other words) are fairly interconnected domains of human activity. Indeed, the subject in both disciplines is the study of visual perception. For physiology of vision the subject is the visual perception of humans and animals, while one of the subjects for computer sciences is creation and testing of technical vision means. The progress in one of these domains would initiate the progress in the other.

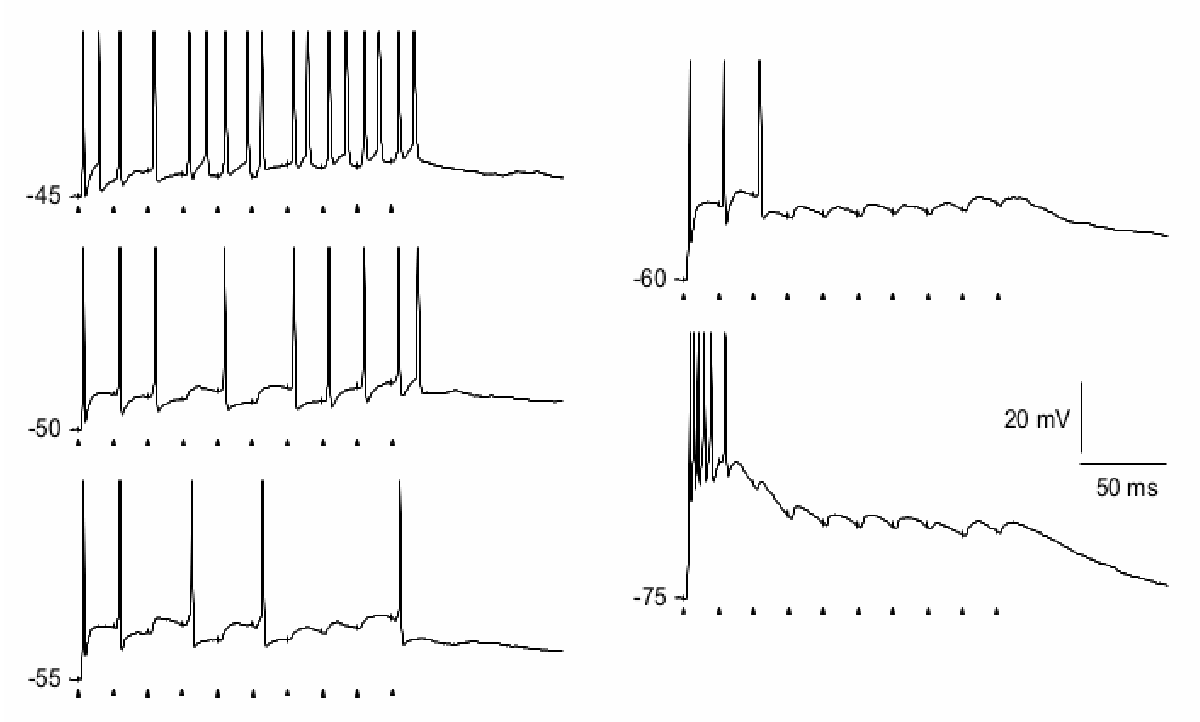


Figure 4 [13]. Firing pattern of a TC neuron at different holding potentials evoked by pulse train stimulation of retinal afferents

The frequency of pulse train was 50 Hz. Holding potentials are indicated to the left of the trace. Timing of stimulus pulses is indicated by arrow-heads below the traces, and can also be deduced from the stimulation artifacts on the traces (truncated). Spike amplitudes were truncated at 0 mV. At -75mV, as expected, the pattern was dominated by a short-latency low-threshold calcium potential, and the elicited action potentials lacked precise timing with respect to the single pulses in the stimulus train.

Unfortunately mutual results' interchange between these disciplines doesn't happen, maybe due to weak interaction among specialists representing different sciences and different methods that are typically used by them.

In the field of image recognition and image processing it wasn't paid much attention to prototyping computer vision from neurophysiological aspects of visual perception (see, for example [14, 15]) or simplified models were considered [16].

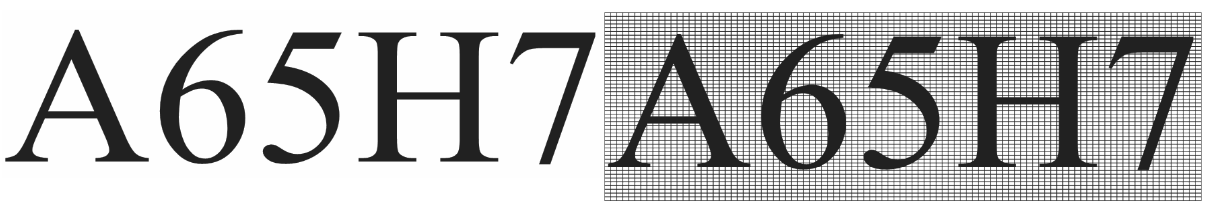
It is used to suggest [15] that initial image for processing is presented in analogous form. As a rule the image is bounded by rectangle – field of view – with dimensions that are suitable for processing. The initial image should always fit this rectangle and fill it if possible.

The first operation one do with image being processing is field of view discretization and brightness quantization for colors in image palette. The following parameters of computer vision system are chosen according to practical considerations: resolution – the number of discrete, usually squared, elements of image – pixels – that fit in one measurement unit (inch, centimeter); the size of image in pixels; the set of values of brightness that each pixel can take value from for each of basic colors (grayscale, RGB, etc.).

It is suggested that the resolution of computer vision system is best suited for specified class of images being processed. This will not lead to unnecessary details being arising with extremely high resolution (like object's contour distortion) and to essential details of the objects being disappearing when the resolution is extremely low. The form of quantization function and the number of brightness levels for colors should conform to images being processing in terms of ability to display the essential details of the object in exactly the same way.

Thus the image can be presented as a two-dimensional array $V(N, M)$ having width N and height M. Each element $v(n, m)$ of this array corresponds to either the brightness of pixel with coordinates n, m for grayscale image or to brightness values of basic colors (e.g. red, green, blue) for color image. This array can be treated as a matrix or a vector depending of what the mathematical methods are chosen.

A lot of methods and algorithms for dealing with image which is presented as a matrix or a vector are developed and applied successfully. Most of them are successfully used for grayscale image processing if discretization and quantization are chosen adequately for that class of images. At the same time the following should be mentioned. For example if the image is processed by means of statistical recognition methods, the brightness values of all pixels in the image are used for determining the measure of similarity of two images $f(V_1, V_2)$. But the pixels in the image usually belong to either object or background. In this case the result of recognition not only depends on pixels brightness of the object but also from brightness values of background pixels that in major cases is not acceptable. Consider an image consisting of a line of arbitrary text on a one-color background and another one where the same line of text is placed over an arbitrary grating (Figure 5). The text on (Figure 5a) can be processed by both statistical and structural methods of recognition.



a) Neutral background; b) Square grating in background

Figure 5. Examples of image with arbitrary text

The text on (Figure 5b) is a far more complex task for recognition. If statistical methods for similarity computing with reference image are applied, the result will be distorted due to pixels belonging to background, but representing an arbitrary placed grating. It is not guaranteed that mutual placement of grating lines and text in the field of view of the image will be the same for arbitrary imposition and subsequent digitization. If attempting to use structural methods of recognition for images like (Figure 5b) it will not be possible to get the contours of the objects. In this case the contours of grating cells will be selected instead of object contours.

At present time the attention of specialists is drawn to tasks dealing with recognition of textures. Grating being placed over text on image is a particular case of them. But methods for texture recognition generally far beyond the complexity of text recognition, face recognition etc. That fact hampers the application of such methods for mentioned tasks significantly or even eliminates the possibility to use them. Similar situation with some differences happens when the grating being placed over the text is of same color as background (Figure 6). For statistical methods computed similarity with reference image will be distorted due to pixels belonging to objects but representing the arbitrary placed grating. The mutual placement of text and grating, again, may be different from image to image after digitization. Attempting to apply the structural methods to images on Figure 6 will give the same result as for image on Figure 5b – the contours of grating cells.

At the same time visual perception cope with similar tasks insensibly, seemingly on subconscious level.

The method of time slices for examining the resizing of receptive field’s excitatory zone (Figure 1 [10], Figure 2 [11]) shows that the time of getting the maximum number of spikes on the neuron’s output counting from the beginning of visual act depends on the size of stimulus. It is possible to suggest that making a decision versus image in visual system happens when the maximum number of spikes on the neuron’s output is reached, in other words – decision can be made for images with different resolution. It is naturally to conjecture that the best resolution for decision making is selected in visual system in the sense of image processing, when the unnecessary details don’t arise and the essential parts of objects don’t disappear.

a) b)

Figure 6. Grating of the same color as background being placed over the text. Images on (a) and (b) have different width of grating lines

It is possible to suggest that it is the processing of observed low resolution images at the beginning of visual act that makes possible consistent visual perception of symbols on different background. Therefore if some optical character recognition (OCR) program would successfully process the image on Figure 5a digitized with low resolution, then it should also successfully process the images from Figure 5b, Figure 6a and Figure 6b digitized with the same value of resolution. A simple experiment with a well-known OCR program FineReader can be conducted in order to check this statement.

The initial dataset consists of four images (Figures 5a, b; 6a, b) 900x280 pixels each, having resolution of 72x72 pixels per inch. The meaningless combination of symbols “A65H7” was chosen intentionally, in order to eliminate the influence of dictionaries on the result of recognition.

The text on Figure 5a was recognized successfully. Processing of the rest three images gave denial of recognition because of inability to find (determine) an object on image. On Figure 7 the same four images as Figures 5 and 6 but digitized with six times lower resolution are presented. Then the text string was recognized successfully on all of them.

a) b)

c) d)

Figure 7. Images having 6 times lower resolution:

a) and b) are from Figure 5 (a, b); and c) and d) are from Figure 6 (a, b)

This experiment clearly shows that the concept of optimal resolution in the sense of processing results may be applied not only to whole image but to each object on it. In this case coarse-to-fine approach was used for different purpose: not to decrease the number of calculation-intensive operations but to solve the problem that can not be addressed at all by traditional methods.

Hypothesis about visual neuron’s functioning at the time of action potential generation

Proposed model of visual neuron's functioning at the time of action potential generation is developed on the basis of systemological analysis of known ideas [17] about neurons' functioning and results of other researchers, presented above. It is an attempt to explain the mechanism and to take into account receptive fields excitatory zones resizing during visual act.

It is known that maintenance of stable resting potential of cell membrane is achieved due to chemical and electrical gradients equilibrium of potassium and chlorine ions. Being changed randomly, the value of membrane potential is restored in the course of potassium and chlorine ions’ transfer through membrane channels. To be exact, chlorine ions transport inward the cell and potassium ions transport outward the cell during membrane depolarization. Considering the neuron as a system, the transport of chlorine ions inward the cell should take place not only in dormancy state but especially during depolarization at the time of action potential generation. Total restoration of chlorine ions’ chemical gradient presumably happens at the end of visual act.

Thus the key concept in our approach is to take into consideration the influence of cell membrane chlorine conductance at the time of action potential generation in the context of whole visual act.

Under the influence of excitatory receptors post-synaptic potential and also with no impact of inhibitory receptors the potential-dependent sodium channels start to open. Sodium flow increases and depolarization increases too, resulting in opening more sodium channels, so the intracellular membrane potential increases up to the value of sodium equilibrium potential.

While sodium flow increments and depolarization increases, chlorine ions enter cell through membrane pores due to concentration gradient while counteractive electrical potential decreases. Furthermore, a point is reached (perhaps the point of maximum potential value) when potential-dependent chlorine channels that also let in chlorine ions cannot but open. Reaching depolarization maximum starts the drift of sodium and potassium ions outward the cell. That returns the cell membrane potential to its initial value but due to chlorine ions ingress the resting potential shifts for some value toward hyperpolarization (Figure 8) at the time of each following pulse generation.

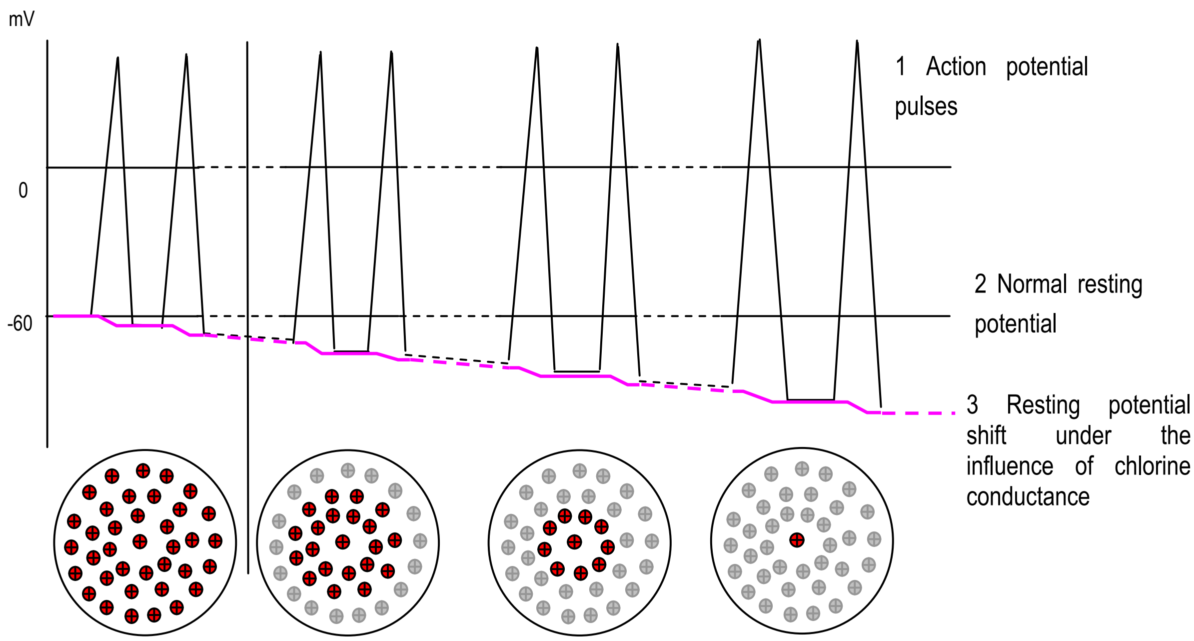


Figure 8. Changes in receptive field excitatory zone

It is known that excitability of different membrane areas in big (afferent and efferent) neurons is not distributed evenly [18]. There exists a low-threshold zone in the area of neuron starting segment (axon hillock and unmyelinated axon initial segment) where membrane possesses several times higher excitation comparing to other cell areas. The opening threshold of potential-dependent sodium channels increases as the distance from axon initial segment grows [19].

So after the *n*-th action potential pulse the resting potential shifts toward hyperpolarization. At the beginning of next (*n+1*-th) pulse this results in potential-dependent sodium channels that have maximal opening threshold (outermost from axon hillock) stops to open under the post-synaptic potential impact of excitatory receptors, located in the same area as this sodium channels. This means that potential of opening threshold for mentioned sodium channels is greater than initial resting potential plus shift value. In other words a number of potential-dependent sodium channels fail to participate in charge accumulation for the *n+1*-th pulse generation.

It is also possible to assume that the distance from excitatory receptor to axon hillock matches distance in the field of view from the point of given receptor to receptive field center. So the “nonparticipation” of some receptors in pulse generation matches lessening of receptive field excitatory zone. On the other hand the less the potential-dependent sodium channels opened in the time of next pulse generation, the greater the time needed for charge accumulation that sufficient to form this pulse. It seems like the reason of decrementing pulse generation frequency over time. This assumption is consistent with experiment on action potentials generation for different holding potentials (artificially set resting potentials) of a cell membrane found in [13].

Even this mechanism is not examined experimentally its presence for excitation zone decreasing of visual system neurons' at the time of visual act fully complies to, explains and confirms the results obtained in [10, 11].

So, during the optical perception, namely one visual act, there exists image data with different resolution for image being viewed and the resolution changes serially from lowest to highest possible value up to the end of visual act. This means that technical vision systems will also have different image elements over time – the pixels of variable size that are changing from maximal to minimum possible size.

Conclusion

As for now we can state that a coarse-to-fine approach is used spontaneously by researchers in different fields of technical sciences. At the same time some studies are carried out in the domain of neurophysiology, showing the presence of such mechanism in living beings' visual system. A hypothesis explaining the functioning of that mechanism in retinal ganglion cells and LGN neurons was presented in this paper.

This hypothesis approval perhaps will enable systematization of coarse-to-fine approach in the field of technical vision; developing of general recommendations and best practices for its application to recognition tasks that can not be resolved at all by traditional methods.

For neurophysiology proof of this hypothesis will mean the possibility and necessity for combined consideration of visual system neurons' intracellular processes and intercellular interaction, refinement of visual neuron's functioning model at the time of action potential generation and explanation of coarse-to-fine mechanism in visual system of living beings.

Bibliography

[1] M. Pedersoli, A. Vedaldi, J. Gonz`alez. A Coarse-to-fine approach for fast deformable object detection. In CVPR, june 2011

[2] P. Moreels, P. Perona. Probabilistic Coarse-To-Fine Object Recognition. Technical report, 2005. California Institute of Technology

[3] Xi Li, Weiming Hu, Wei Hu. A Coarse-to-Fine Strategy for Vehicle Motion Trajectory Clustering. ICPR, 2006

[4] Jiann-Der Lee and Chen-Hui Kuo (2009). A Multi-Stage Classifier for Face Recognition Undertaken by Coarse-to-fine Strategy, State of the Art in Face Recognition, Julio Ponce and Adem Karahoca (Ed.), ISBN: 978-3-902613-42-4, InTech, Available from: http://www.intechopen.com/books/state_of_the_art_in_face_recognition/a_multistage_classifier_for_face_recognition_undertaken_by_coarse-to-fine_strategy

[5] Tyshchenko M. A. 3D reconstruction of human face in person identification problems. PhD thesis. International Research and Training Center for Information Technologies and Systems, Kyiv, 2012

[6] Christopher Raphael. Coarse-to-Fine Dynamic Programming. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001. vol. 23. p. 1379-1390.

[7] Brian O. Lucena. Dynamic Programming, Tree-width and Computation on Graphical Models. PhD thesis. Division of Applied Mathematics. Brown University 2002

[8] Maunsell, J.H.R. And Van Essen, D.C. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. J. Neurophysiol. 49: p. 1127-1147, 1983.

[9] R. Battiti, E. Amaldi, C. Koch. Computing Optical Flow Across Multiple Scales: An Adaptive Coarse-to-Fine Strategy. International Journal of Computer Vision, 6:2, p. 133-145, 1991

[10] Н.Ф. Подвигин Динамические свойства нейронных структур зрительной системы Ленинград: Наука, 1979, 158 с.

[11] Ruksenas O, Bulatov A, Heggelund P. Dynamics of Spatial Resolution of Single Units in the Lateral Geniculate Nucleus of Cat During Brief Visual Stimulation. J Neurophysiol 97:1445-1456, 2007.

[12] David H. Hubel. Eye, brain, and vision New York: Scientific American Library: Distributed by W.H. Freeman, 1988 240 p.

[13] S. Augustinaite and P. Heggelund. Changes in firing pattern of lateral geniculate neurons caused by membrane potential dependent modulation of retinal input through NMDA receptors. J Physiol 582.1 pp 297–315, 2007

[14] М.Шлезингер, В.Главач Десять лекций по статистическому и структурному распознаванию. – Київ: Наукова думка, 2004. – 535с.

[15] Т.Павлидис Алгоритмы машинной графики и обработки изображений. М Радио и связь 1986г. 400с

[16] У.Прэтт Цифровая обработка изображений: Пер. с англ.—М.: Мир, 1982.— Кн.1—312 с, ил.

[17] Nicholls J. G., Martin A. R., Wallace B. G., Fuchs P. A. 2001. From Neuron to Brain. 4th Ed. Sinauer Associates, Inc. Sunderland, MA USA.

[18] Wollner, D.A., and Catterall, W.A. 1986. Localization of sodium channels in axon hillocks and initial segments of retinal ganglion cells. Proc. Natl. Acad. Sci. USA 83, 8424–8428.

[19] Fried S. I., Lasker A. C. W., Desai N. J., Eddington D. K., Rizzo J. F. 3rd. Axonal sodium-channel bands shape the response to electric stimulation in retinal ganglion cells. J. Neurophysiol. 101: 1972–1987, 2009.

Authors' Information

Anton Sharypanov *– leading software engineer, master of science in flexible cybernated systems and robotechnics, Institute of cybernetics, prosp. akad. Glushkova 40, 03680, Kiev, Ukraine;*
e-mail: _sha_@ukr.net

Alexandra Antoniouk *– senior researcher, PhD (candidate of sciences in mathematics and physics), Institute of Mathematics NAS Ukraine, Tereshchenkivska, 3, 01 601, Kyiv, Ukraine;*
e-mail: antoniouk@imath.kiev.ua

Vladimir Kalmykov *– senior researcher, candidate of engineering sciences, Institute of problems of mathematical machines and systems, prosp. akad. Glushkova 42, 03680, Kiev 187, Ukraine;*
e-mail: vl.kalmykov@gmail.com, kvg@immsp.kiev.ua