PHYSIOLOGICALLY PLAUSIBLE MODEL OF A NOVELTY FILTER IN THE VISUAL SYSTEM THAT SIMULATES THE MCCOLLOUGH-EFFECT

P.V. Maximov, V.V. Maximov
Institute for Information Transmission Problems, RAS
Bolshoi Karetny 19, 127994 Moscow GSP-4, Russia
pmaximov@iitp.ru
Bio-Optics, Bio-Photonics, High Resolution Imaging, Vision and Photoreceptors, oral

To account for the orientation-contingent colour after-effect a new neural model of the novelty filter is proposed that uses chemical synapses with Hebbian plasticity and electrical synapses permitting to propagate control signals to modifiable synapses.

The orientation-contingent colour after-effect, or the McCollough-effect (ME), is generally obtained when after the exposure for a few minutes to alternating coloured gratings (e.g., red vertical and green horizontal ones) subjects perceive similarly oriented achromatic gratings as if they were tinted with complementary hues. A traditional explanation of the ME by adaptation of single detectors selective to colour and orientation suffers from a number of inconsistencies: 1) the ME lasts much longer (from several days up to 3 months) than ordinary adaptation, decay of the effect being completely arrested by night's sleep or occluding the eye for a long time, 2) the strength of the ME practically does not depend on the intensity of adapting light, 3) a set of related pattern-contingent after-effects discovered later required for such an explanation new detectors, specific to other patterns. These properties can be explained, however, in the framework of associative memory and novelty filters. It is supposed that in the course of "adaptation" the colour grid patterns are stored in the memory. When testing with an achromatic grid, the associative image of the grid with the same orientation is recalled from the memory and subtracted from the input image, leaving in the resulting image only novel feature, namely, a complementary colour of the grid – Fig. 1.

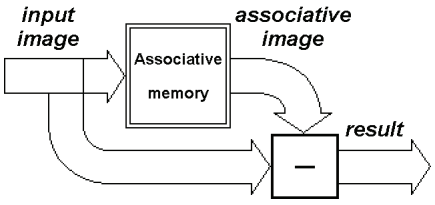

Fig. 1. The novelty filter

In the previous works[1,2] we developed a linear computational model of the filter based on a neural network with modifiable synapses (Fig. 2) which consists of (i) an input layer of two (left and right eyes) square matrix with two analog neurons (red and green) in each pixel, (ii) an isomorphic associative neural layer, each secondary analog neuron being synaptically connected with all neurons of the input layer, except a neuron of the same position, so that excitations y_i of the associative neuron were weighted sums of signals x_j of input neurons:

$$y_i = \sum_{j=1}^{m} w_{ij} \cdot x_j \text{ , where } i = 1,2,\ldots,m, \text{ and } w_{ii} = 0,$$

(iii) an output layer of novelty neurons, each neuron of which calculated the difference between corresponding values of the input and associative layers: $z_i = x_i - y_i$.

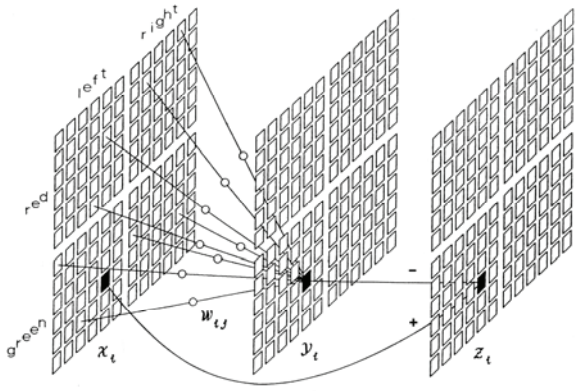

Fig. 2. Computational model of the novelty filter that accounts for the ME.
Red- and green-sensitive neurons of left and right eyes of the same spatial position are distributed to four adjacent areas for clarity.

Modification of synaptic weights conformed to the following learning rule:

$w_{ij}^{t+1} = w_{ij}^{t} + \varepsilon \cdot x_j^t \cdot z_i^t$, where $i \neq j$.

The function of the model was examined by simulation. After a few presentations of coloured gratings the model displayed the ME that was slowly destroyed by subsequent presentations of random pictures. With a sufficiently large receptor matrix the effect lasted a thousand times longer than the period of adaptation. At the same time the effect was rapidly erased by presentation of achromatic grids of the same orientations and spatial frequencies as adapting ones. Continuous darkness did not change the strength of the effect. Like real ME, the model did not display an interocular transfer. The model can account for different pattern-contingent colour after-effects without assuming any predetermined specific detectors. Such detectors were constructed in the course of adaptation to specific stimuli (gratings).

Unfortunately, while the proposed scheme demonstrated fundamental possibility of the ME in the network with modifiable synapses, it was physiologically implausible[3]. Unlike the Hebbian synapses whose plasticity was controlled by membrane potentials of presynaptic and postsynaptic neurons only (Fig. 3, A), there was a third neuron in that model, involved in the control of modification of synaptic weight, besides the two neurons immediately forming the synaptic contact (Fig. 3, B). It was unclear what kind of neurophysiological mechanism could organize the information flow from that remote neuron to the synapse.

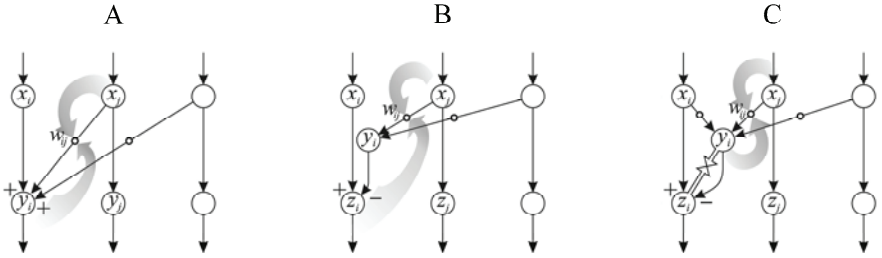

Fig. 3. Wiring diagrams of synaptic weights modification.

(A) the Hebbian network, (B) the physiologically implausible novelty filter and (C) the novelty filter with electrical synapses. Neurons are shown by large circles: *x* – input layer, *y* – associative layer, *z* – output layer (novelty neurons). Thin black arrows represent conventional synaptic contacts between neurons. The electrical synapse is shown with open arrows directed towards each other. Large gray arrows indicate paths of signals that control the synaptic weights.

The problem can be simplified, if some fraction of membrane potential of the output novelty neuron would be sent directly to the corresponding associative neuron in order that the latter could control synaptic weights of its synapses (e.g., by using the Hebbian scheme). One of the possible ways to do that is to use so called *electrical synapses* between neurons (Fig. 3, C). At these synapses there are no synaptic gap and chemical transmitters. Instead, there are gap junctions, direct channels between neurons that establish continuity between the cytoplasm of adjacent cells and provide spread of membrane potential from one cell to another. Such synapses are quite common in the nervous system, including the retina. In this paper we suggest a novelty filter based on this idea.

Like the previous model, the model with electrical synapses consists of three isomorphic neural layers of analog neurons. As distinct from that model each secondary neuron is synaptically connected only with a restricted set of nearby input neurons, including the neuron with the same location, its total input signal being $\tilde{y}_i = \sum_{j \in D_i} w_{ij} x_j$, where $i=1,2...,m$, w_{ij} are synaptic weights and D_i denote a subset of input neurons, synaptically connected to *i*-th associative neuron. Neurons of the third layer are connected to the corresponding neurons of the first two layers – with excitatory synapses to the input layer neurons and inhibitory ones to the associative layer neurons. Hence the output neurons calculate the difference between the values of corresponding neurons of input and associative layers: $z_i = x_i - y_i$. Besides that the output neurons were connected with corresponding associative neurons by means of an electrical synapse. Resultant signal of the associative neuron (its membrane potential) is $y_i = p \cdot z_i + (1-p) \cdot \tilde{y}_i$, where p is a constant determining the efficacies of electrical synapses. Results of computer simulations presented later were obtained with a value of $p = 0.2$.

We use a classical Hebbian learning rule in the model:

$\Delta w_{ij} = \varepsilon \cdot x_j \cdot y_i$. (*)

As follows from the latter expression, Δw_{ij} cannot be negative with nonnegative values of x_j and y_j, and hence the connections of any secondary neuron with input neurons will not decrease in the course of adaptation. In particular, this means that any erroneous connection between some associative neuron and some input neuron of the opposite color type cannot disappear during further adaptation. Therefore, as distinct from our previous model, we use the input stimuli with values from -0.5 to 0.5 (the value of -0.5 corresponding to low intensity and the value of 0.5 corresponding to high intensity in each channel). Another feature of our model was vetoed sign change of synaptic weights[4]. Actually an excitatory synapse cannot turn into an inhibitory one during modification of weights. Therefore inhibitory connections were prohibited in the model, and the learning rule was

$$\widetilde{w}_{ij}^{t+1} = \begin{cases} w_{ij}^t + \Delta w_{ij}, & when \quad w_{ij}^t + \Delta w_{ij} > 0 \\ 0, & when \quad w_{ij}^t + \Delta w_{ij} \le 0 \end{cases}.$$

The next problem that we had faced during our transition to this model was the fact that this basic Hebbian scheme (*) led to unrealistic growth of the synaptic weights. To overcome this inconvenience we used an ordinary postsynaptic normalization:

$$w_{ij} = \frac{\widetilde{w}_{ij}}{\sum_{j \in D_i} \widetilde{w}_{ij}}.$$

Results of simulation are presented at Fig. 4. In the model, the size of the retina was 16×16 pixels and sizes of the areas of input neurons connected to each associative neuron were 9×9 pixels. The strength of the effect was evaluated by presentation of vertical white grids to the input layer of the model. The value of the effect was calculated as the difference between mean excitations of novelty neurons z, corresponding to white pixels of input image, in green and red channels. We made the same series of computer experiments with this model that was done with the previous model of the novelty filter[2]. During the first 1000 cycles of iterations random stimuli were presented to the model. After that, to obtain the ME, red vertical and green horizontal grids were presented alternately, and next, random stimuli were presented again. It can be seen in fig. 4 that adaptation (branch *a* of the curve) goes much faster than decay (branch *b*). At the same time the effect is rapidly erased by presentation of alternately oriented achromatic grids (branch *c*). As expected, the persistence of the ME (the ratio of speeds of adaptation and desadaptation) is proportional to the number of plastic synapses in the receptive fields of associative neurons. The rate of learning depends on the intensity of the adapting stimuli, but the final strength of the effect does not – dotted curve in fig. 4 represents the course of adaptation to twice less bright grids than that for the branch *a*. It is necessary to notice that the fig. 4 presented below almost repeats the fig. 4,*a*, published in the article[2], dedicated to the initial (physiologically implausible) model of the novelty filter. The only difference is a much slower decay during desadaptation (branch *b*), due to the greater number of plastic synapses in the receptive fields of the associative neurons in this case (area of 9×9 pixels) in comparison with the previous model (6×6 pixels).

Thus, the results of simulation show that physiologically plausible scheme of novelty filter with electrical synapses does work and presents the same phenomena as our initial novelty filter does[1,2].

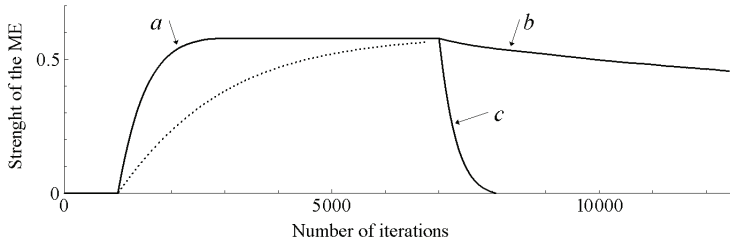

Fig. 4. Time courses of the McCollough-effect in the model during (*a*) an adaptation to alternating colour grids and (*b*) a subsequent desadaptation with random stimuli or (*c*) erasing the effect with black-and-white test grids.

As concerns the possible localization of such a mechanism in the visual system and specific types of neurons involved in its organization the question remains still open. One may suppose any level from the retinal inner plexiform layer to first stages of visual processing in the striate cortex, where signals from two eyes converge. In any case the direct way (from input neurons to novelty neurons) must be organized with

neurons that possess small receptive fields, sufficient to provide for a good spatial resolution of the grids, while some interneurons with wide dendritic trees serve as associative neurons. In the retina, it is naturally to suppose bipolar cells as the input neurons, the output retinal neurons (ganglion cell) as novelty neurons, and amacrine sells as associative neurons.

Our scheme of the novelty filter with electrical synapses (Fig.3,C) contains an inhibitory synapse between associative neurons and novelty neurons. C.C.D. Shute[5] suggested that the ME could reflect a mechanism which should be inhibited by GABAergic neurons. This means that associative interneurons release inhibitory neurotransmitter GABA. The classical inhibitory neurotransmitter, gamma aminobutyric acid (GABA) occurs in many different varieties of amacrine cells in most vertebrate retinas[6]. Electrical synapses are also frequent in the retina[7]. In most cases gap junctions make contact between neurons of the same type. But sometimes electrical synapses exist also between neurons of different types, including contacts between some types of amacrine and ganglion cells. In particular, R. Jacoby et al. have described two types of amacrine cells that make gap junctions with so called parasol ganglion cells in the primate retina[8]. One type of these amacrine cells was GABAergic and made both inhibitory chemical synapses and gap junctions with the parasol ganglion cell dendrites. Unfortunately, this harmonious morphological scheme cannot account for the ME. The parasol ganglion cells cannot serve for the visual pathway of high resolution. They have rather large receptive fields, project to the magnocellular layers of the lateral geniculate nucleus and are particularly important for the perception of motion, but not the visual tasks, requiring high spatial resolution[8].

As to a fine morphological structure of synaptic interactions at the subsequent levels of the visual system the information is rather poor. The massive presence of gap junction couplings among interneurons of the striate cortex was recently discovered[9]. Gap junctions are common between cortical GABAergic interneurons[10], but little is known about the electrical synapses between these interneurons and neurons of the direct pathway, responsible for high spatial resolution.

This work was supported by the program of the RAS Branch of Biological Sciences
“Physiological mechanisms for regulation of internal environment and control of behavior in living systems”.

[1] V.V. Maximov, P.V. Maximov 1995 Visual associative memory can account for orientation-contingent colour after-effect // *Göttingen Neurobiology Report (edited by N.Elsner & R.Menzel)*, Stuttgart: Thieme, P. 899
(an electronic version of the poster is available at the site: http://www.iitp.ru/projects/posters/me)

[2] V.V. Maximov, P.V. Maximov 2004 Visual associative memory and orientation-contingent colour after-effect // *Biofizika*, **49**, P. 920-927 (in Russian)

[3] P.V. Maximov 2006 Neurones with modifiable synapses are able to establish colour-specific connections in the retina // *Proceedings of the ICO Topical Meeting on Optoinformatics / Information Photonics’2006 (edited by M.L. Calvo, A.V. Pavlov, J. Jahns),* St Petersburg: ITMO, P. 252-254.

[4] P.V. Maximov, V.V. Maximov 2004 Visual associative memory with omitted inhibitory synaptic connections: Is it still able to improve the quality of distorted images? // *III International Optical Congress ‘Optics – XXI Century’, Topical Meeting on Optoinformatics – Book of Abstracts*, Saint-Petersburg: ITMO, P. 32-33.

[5] C.C.D. Shute 1979 *The McCollough effect: An indicator of central neurotransmitter activity*. London: Cambridge Univ. Press.

[6] H. Kolb, E. Fernandez, R. Nelson *The organization of the retina and visual system*. http://www.webvision.med.utah.edu/

[7] D.I. Vaney 1994 Patterns of neuronal coupling in the retina // *Prog. Ret. Eye Res.*, **13**, P. 301–355.

[8] R. Jacoby, D. Stafford, N. Kouyama, D. Marshak 1996 Synaptic inputs to ON parasol ganglion cells in the primate retina // *J. Neurosci.*, **16**, P. 8041–8056.

[9] T. Fukuda 2007 Structural organization of the gap junction network in the cerebral cortex // Neuroscientist, **13**, P. 199 - 207.

[10] T. Fukuda, T. Kosaka, W. Singer, R.A.W. Galuske 2006 Gap junctions among dendrites of cortical GABAergic neurons establish a dense and widespread intercolumnar network // J. Neurosci., **26**, P. 3434-3443.