Vector model for mapping of visual space to subjective 4-D sphere

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2014 J. Phys.: Conf. Ser. 490 012124
(http://iopscience.iop.org/1742-6596/490/1/012124)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 158.129.202.15
This content was downloaded on 20/05/2014 at 11:38

Please note that terms and conditions apply.
Vector model for mapping of visual space to subjective 4-D sphere

Dalius Matuzevicius\textsuperscript{1,2} and Henrikas Vaitkevicius\textsuperscript{1}

\textsuperscript{1}Department of General Psychology, Faculty of Philosophy, Vilnius University, Universiteto str. 9/1, Vilnius, Lithuania
\textsuperscript{2}Department of Electronic Systems, Faculty of Electronics, Vilnius Gediminas Technical University, Naugarduko str. 41-413, Vilnius, Lithuania

E-mail: dalius.matuzevicius@gmail.com, henrikas.vaitkevicius@ff.vu.lt

Abstract. Here we present a mathematical model of binocular vision that maps a visible physical world to a subjective perception of it. The subjective space is a set of 4-D vectors whose components are outputs of four monocular neurons from each of the two eyes. Monocular neurons have one of the four types of concentric receptive fields with Gabor-like weighting coefficients. Next this vector representation of binocular vision is implemented as a pool of neurons where each of them is selective to the object’s particular location in a 3-D visual space. Formally each point of the visual space is being projected onto a 4-D sphere. Proposed model allows determination of subjective distances in depth and direction, provides computational means for determination of Panum’s area and explains diplopia and allelotropia.

1. Introduction
Perception of 3-D visual space through processing of two different projections of the scene onto each retina is addressed in many papers [1–8]. Current widely accepted models of stereo vision interprets the coding of either disparity or visual direction; moreover, no biologically plausible models combine processing of both. Such properties of stereo vision as diplopia, allelotropia, and stereoacuity phenomena still are not computationally explained. Well-known energy model suggested explanation for physiological mechanisms of disparity detectors’ construction [5–8]. However the disparity describes only one spatial coordinate of the object. Mechanisms of direction encoding are not considered, suggesting the opinion about different means of coding disparity and direction.

We present a mathematical model of binocular vision and show that coding mechanisms of spatial coordinates (depth and direction) could be the same. Using proposed model we quantitatively describe perceived depth and direction as a function of disparity and parallax, Panum’s fusional area, diplopia and the exponential decrease of depth sensitivity while object moves away from horopter.

2. Model structure
Schematic diagram of a binocular vision system is presented in Fig. 1 where the neural structure (model) for binocular computations is the key element.

The model has two distinct parts: local linear part (structured array of local analyzers – LAs, Fig. 1b) and the global nonlinear part (global analyzer – GA, Fig. 1c). LAs (detailed in...
Figure 1. Schematic diagram of stereo vision system: (a) positions of stimulus in space with projections onto receptive fields of retinas; (b) array of local analyzers; (c) Structure of global analyzer and transduction of signals from local detectors to global detectors; (d) responses of local and global detectors and identification of Panum’s area when stimulus moves on cyclopean axis; (e) responses of local and global detectors when stimulus moves on horopter.

Fig. 2a) perform linear mapping of each point from the local visual space onto a 4-D sphere. Such procedure assigns local spatial coordinates to the object presented in a particular small area of 3-D visual space. Meanwhile, the global analyzer combines responses of LAs by means of nonlinear operations (here – MAX operation) and provides the global spatial coordinates (horizontal parallax and disparity) of a presented binocular feature.

Local analyzers (LAs) also have two distinguishable parts: monocular and binocular (Fig. 2a). Monocular part is the group of four monocular neurons situated in each eye and being connected to the same crowd of photoreceptors but having different types of Gabor-like receptive fields (RFs) [9]. So monocular outputs is a weighted sum of photoreceptor outputs and can be treated as components of the 4-D vector whose direction encodes position of the light centroid over the current RF. It is noteworthy that LA has corresponding RFs, i.e. its RFs are in the corresponding areas of retinas [10]. RF structures (weight functions) of four monocular neurons are presented in Fig. 2b and are obtained according to Eq. (1). Given the distribution of the light over corresponding RFs $I(\alpha, \gamma)$ and $I(\beta, \gamma)$, Eq. (2) defines outputs of monocular neurons $M_{iL}$ and...
Neurons of binocular part (Fig. 2a) resemble class of “simple” neurons. There are four independent neurons that linearly sum up signals received from monocular neurons (Fig. 2a, Eq. (4)) and a pool of linearly dependent neurons that act as detectors of disparity and parallax (Fig. 2a, Eq. (5)). Four independent neurons can be treated as components of the 4-D vector $B(\phi_c, \Delta_c) = \{B_1, B_2, B_3, B_4\}$ whose direction encodes parallax and disparity of a binocular feature. Simple binocular neurons-detectors receive signals from neurons $B_1 - B_4$ and are tuned to the specific spatial coordinates, i.e. they are selective to the specific orientation of vector $B$, or – to particular location of the object in a visual space spanned by the current LA. When projections of the object are presented on RFs of that LA, only one detector is maximally excited. Finding the row and column numbers of maximally excited detector means identification of local depth and direction respectively. So the mission of local analyzer is to assign local coordinates to the object that is presented in a such part of visual space which is being projected onto left and right RFs of that LA.

Global analyzer (GA) (Fig. 1c) is the internal map of external visual field. Here global spatial coordinates of the featured object are derived from the local ones. Neurons of GA resemble class of “complex” neurons. Neurons in GA are ordered in such way that neurons located in the middle column are responding to zeroth direction (when stimulus is on horopter), and neurons in the middle row are selective to zeroth depth (when stimulus is on cyclopean axis). Side neurons are signalling left or right direction, and top/bottom -row neurons are tuned to crossed/uncrossed disparities. Each GA neuron receives signals from neurons of several adjacent LAs and passes the maximal one, i.e. connections are created only between LAs and GA neurons which are selective to the same location in visual space. Depth and direction of the object in visual space are identified by the row and column numbers of a maximally excited global neuron-detector. Responses of LAs and GA detectors when stimulus is moving on horopter or cyclopean axis are presented in Fig. 1d,e.
3. Results and discussion
The proposed neural model of binocular vision is able to identify direction (parallax) and depth (disparity) of the object, and identified direction and depth are linear functions of physical parallax (Fig. 3a) and disparity (Fig. 3b) of that object. In these graphs numbers of maximally excited detectors were obtained from GA responses in Fig. 1d,e. Identified direction and depth are constant while stimulus is located on horopter and cyclopean axis respectively.

Events in GA when stimulus is moving along the horopter and cyclopean axis are significantly different. Only one detector is maximally excited while the stimulus is moving along the horopter. When stimulus gets far from horopter two detectors encoding different parallaxes appear maximally excited, and here diplopia occurs. Area in visual space in which presented stimulus causes only one peak in GA may be called as the Panum’s region of the model.

Identified parallax remains constant until stimulus is presented within Panum’s area and the identified depth is linear function of the physical disparity. That is in agreement with experimental data [11]. Moreover the identified parallaxes of diplopic stimuli differ from parallaxes of the same monocular stimuli, when each of them is presented alone, i.e. there is resemblance to the phenomena of allelotropia [12].

It is known that the accuracy of depth perception is decreasing exponentially while the stimulus is moving away from the horopter [13, 14]. Our model demonstrated such pattern (Fig. 3c) when the noise was introduced, and this argues that there is no necessity to have two (course and fine) systems for disparity determination in order to explain exponential decrease.

Acknowledgments
B. Postdoctoral fellowship is being funded by European Union Structural Funds project "Postdoctoral Fellowship Implementation in Lithuania”

References
[1] Barlow H B, Blakemore and Pettigrew J D 1967 J. Physiol. 193 327–42
[2] Pettigrew J D, Nikara T and Bishop P O 1968 Exp. Brain. Res. 6 353-72
[3] Poggio G F, Gonzalez F and Krause F 1988 J. Neurosci. 8 4531–50
[4] Marr D and Poggio T 1979 Proc. R. Soc. Lond. B 204 301–28
[5] Freeman R D and Ohzawa I 1990 Vision Res. 30 1661–76
[6] Ohzawa I 1998 Curr. Opin. Neurobiol. 8 509–15
[7] Ohzawa I, DeAngelis G C and Freeman R D 1997 J. Neurophysiol. 78 3279–909
[8] Fleet D, Wagner H and Heeger D J 1996 Vision Res. 36 1839–57
[9] Fomin S V, Sokolov E N and Vaitkevicius H H 1979 Artificial sensory system: a problem of sensory simulating (Nauka: Moscow) p 189
[10] Kulikowski J J and Vidyasagar T R 1986 Exp. Brain Res. 64 5–18
[11] Lawson R B, Cowan E, Gibbs T D and Whitmore C G 1974 J. Exp. Psychol. 103 1142–6
[12] Rose D and Blake R 1988 J. Opt. Soc. Am. A 5 1512–21
[13] Farell B, Li S and McKee S P 2004 J. Vision 4 488–99
[14] McKee S P, Verghese P and Farell B 2005 J. Vision 5 783-92