Vision Research 105 (2014) 199–203

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Coding depth perception from image defocus

Hans Supèr [a,b,c,*], August Romeo [a]

[a] Dept. of Basic Psychology, Faculty of Psychology, University of Barcelona (UB), Spain
[b] Institute for Brain, Cognition and Behavior (IR3C), Spain
[c] Catalan Institution for Research and Advanced Studies (ICREA), Spain

ARTICLE INFO

Article history:
Received 4 July 2014
Received in revised form 3 October 2014
Available online 31 October 2014

Keywords:
Depth from defocus
Chromatic aberration
Spiking model

ABSTRACT

As a result of the spider experiments in Nagata et al. (2012), it was hypothesized that the depth perception mechanisms of these animals should be based on how much images are defocused. In the present paper, assuming that relative chromatic aberrations or blur radii values are known, we develop a formulation relating the values of these cues to the actual depth distance. Taking into account the form of the resulting signals, we propose the use of latency coding from a spiking neuron obeying Izhikevich's 'simple model'. If spider jumps can be viewed as approximately parabolic, some estimates allow for a sensory-motor relation between the time to the first spike and the magnitude of the initial velocity of the jump.

1. Introduction

The visual system of spiders includes layers which receive defocused images containing depth information. A remarkable behavioural experiment reported in Nagata et al. (2012) has shown that depth perception is affected by light wavelength through the variable amount of defocus related to the chromatic aberration. This observation has resulted in a specific mechanism proposed by the authors of that work. Although it is not immediate to establish direct relations between this model and the responses in humans or other vertebrates, there are a number of aspects more or less evidently connected with its working.

The question of ocular control determined by chromatic defocus has been object of study in humans and guinea pigs in Kotulak, Morse, and Billock (1995) and Qian et al. (2013) and Refs. therein. Accommodative gain in humans depends on the chromatic bandwith of the stimulus under dynamic conditions (the present work offers a simplified approach in which bandwith issues are not considered, but they are undoubtedly present in the real world).

In vision systems with two sensor classes, each with different wavelength sensitivity, chromatic aberrations can be exploited. For such a category, which includes the human eye with their L and S cones, chromatic aberrations provide a signed cue to defocus (Fincham, 1951; Flitcroft, 1990). This is possible because the aberration introduces a sign-dependent tendency for one sensor class to have greater amplitudes than the other. At the same time, monochromatic aberrations can provide an odd-error cue to focus direction (Wilson, Decker, & Roorda, 2002), as blur shape may lead to the detection of differences in the appearance of the point spread function (PSF) between myopic and hyperopic defocus. For humans, adaptations to blur changes are important because adjustments may play a role in tuning the match between cortical responses and the spatial structure of images (Webster, Georgeson, & Webster, 2002).

Perception of depth from defocus (DFD) is a well established concept, and the idea of using changes in focus settings has been in circulation for a long time. Computationally oriented outlooks are offered by e.g. (Chaudhuri & Rajagopalan, 1999) or (Schechner & Kiryati, 2000), while recent psychological or biological considerations are better illustrated by papers like Mather and Smith (2000), Held, Cooper and Banks (2012), and Read (2012), and bibliography therein. Regarded as a source of defocus estimation, blur seemed to be just a complementary cue, the pre-eminent one being stereoscopic disparity. However, blur is not so weak as previously guessed. While disparity is more precise near fixation blur is more precise away from the focusing plane (Held, Cooper, & Banks, 2012; Read, 2012).

In the calculation of DFD, blur differences between images are employed as a cue for distance estimation. Accurate blur evaluation from natural scenes is in general a difficult problem which shall not be addressed in the present work. Instead, we will assume that the values of the relevant variables have already been extracted from the visual input. Even such a simple method as measuring the 'edge bleeding' length is applicable only in cases where the image is binary and contained in a single plane perpendicular to the line of gaze. Procedures based on transfer functions usually require some amount of frequency-based reasoning. In fact, real data always depend on the present space frequencies, and that

---

[*] Corresponding author.

http://dx.doi.org/10.1016/j.visres.2014.10.022

sort of limitation would also constrain the information from the spider eye. After the classical papers (Pentland, 1987; Subbarao, 1988), many techniques have been devised for the calculation of blur parameters, tackling major issues such as the presence of noise (Chaudhuri & Rajagopalan, 1999), space invariance (Chaudhuri & Rajagopalan, 1999; Jin & Favaro, 2002), the geometric model (Favaro & Soatto, 2005) or the nature of the employed filters (Watanabe & Nayar, 1998; Burge & Geisler, 2011).

A study of responses to perturbations in Schechner and Kiryati (2000) showed that for the case of two-dimensional DFD using a circularly symmetric lens-aperture the *aperture problem* does not appear. From that viewpoint, DFD allows for more robust estimations (however, note that the comparison involved stereo systems made of ideal cameras, not biological pupils).

In the present paper we wish to make an explicit formulation of the idea in Nagata et al. (2012). Supposing that the problem of measuring the chromatic aberration or the blur radii from the images themselves has already been solved at some previous stage, we concentrate on the question of setting up some neural coding for the obtained magnitudes, and suggest the use of spike latency. The proposal is illustrated by a numerical simulation of a spiking neuron hypothetically doing this job.

2. Methods

Efficient visual systems should enable animals to obtain food and avoid predators. Jumping spiders try to jump accurately enough to catch their preys. They are equipped with two pairs of principal eyes (PEs) and anterior lateral eyes (ALEs). Even if ALEs are occluded jumps can be precise, showing that PEs suffice for absoulte depth perception. The remarkable point is that PEs do not have overlapping fields, thus ruling out stereo effects. Moreover, these eyes have no focal adjustment mechanism, and no moves capable of generating motion parallax have been observed either. Thus, the only remaining explanation lies in the use of blur differences from the different wavelength sensitive parts of the PE retinas (Nagata et al., 2012 and Refs. therein).

A basic way of modelling the function of these PEs is to consider a lens with two focal lengths depending on the wavelength values. Like in Nagata et al. (2012), the starting point is the thin lens equation written in the form

$$\frac{1}{d}+\frac{1}{v}=\frac{1}{F} \tag{1}$$

where d is the object distance, v the image distance and F the focal length. Note that $d>0$ for objects in front of the lens and $v>0$ for images behind the lens. In these conditions, d is al so called ‘depth’. The authors of Nagata et al. (2012) study the case of equal image distances v and different perceived depths d, d', caused by the use of different focal lengths, say F_g, F_r with $F_g<F_r$ (‘g’ for green, ‘r’ for red). Taking the two equalities and deleting one from the other it is immediate to arrive at

$$d'=\frac{d}{1+df}=\frac{1}{f}\left(1-\frac{1}{1+df}\right), \quad f\equiv\frac{1}{F_g}-\frac{1}{F_r}, \tag{2}$$

which is Eq. (1) in Nagata et al. (2012). The meaning of the present magnitudes is d' = estimated distance, d = true distance.

For this type of lens, the general law relating the absolute values of object and image sizes R_o, R, focal length F and object distance d reads

$$R=R_o\,\frac{F}{d-F} \tag{3}$$

In the studied situations, $d>F$.

2.1. Aberration method

Different focal values F give rise to different image sizes R. Let R_g, R_r indicate the image sizes R of (3) for $F=F_g$ and for $F=F_r$ respectively. The difference R_r-R_g amounts to the lateral or transverse chromatic aberration. We call $\mathcal{C}$ the ratio between this chromatic aberration and one of the two sizes, i.e.

$$\mathcal{C}=\frac{|R_r-R_g|}{R_r}=\frac{d}{d-F_g}F_g f. \tag{4}$$

Thus, $\mathcal{C}>0$ for $d>F_g$. From this equation it is straightforward to find the true distance d as a function of $\mathcal{C}$

$$d=\frac{F_g}{1-\frac{\mathcal{C}_\infty}{\mathcal{C}}}\equiv d(\mathcal{C}), \quad \mathcal{C}_\infty\equiv F_g f. \tag{5}$$

Next, we can consider what happens when ‘g’ and ‘r’ are interchanged in Eq. (4), i.e., the reference is ‘r’ instead of ‘g’, while the f value is kept unchanged. Thus, instead of (4), we are left with

$$\mathcal{C}'=\frac{|R_r-R_g|}{R_g}=\frac{d}{d-F_r}F_r f, \tag{6}$$

where we have made use again of (3). At this point, evaluating the $d(\mathcal{C})$ function of Eq. (5) with argument $\mathcal{C}'$, we obtain

$$d'=d(\mathcal{C}')=\frac{F_g}{1-\frac{\mathcal{C}_\infty}{\mathcal{C}'}}. \tag{7}$$

One may care to check that Eqs. (4)–(7) indeed lead to the relation between d, d' already established by Eq. (2).

For this mechanism to work it is necessary to find the relative aberration $\mathcal{C}$ from two image planes by examining the regions where green is distinct and where red is distinct. When the figure is a circle centred at the origin it is possible to obtain $\mathcal{C}=1-R_g/R_r=1-\sqrt{A_g/A_r}$, where A_g, A_r indicate the figure areas in image space for the green-distinct and red-distinct parts (this is an ideal set-up, rather infrequent in the world of real images; the general problem of extracting distance cues is a question which we are here sidestepping). Although simple, the ‘lateral chromatic aberration’ method requires simultaneous inputs from two planes.

2.2. Blur method

In each of the two considered image planes one colour shows a distinct outline and the other a blurred one. Following e.g. Chaudhuri and Rajagopalan (1999) or Pentland (1987), or studying the similar triangles in Fig. 3A of (Nagata et al., 2012), it is easy to reason that the blur radius b amounts to

$$b=\rho\,\frac{|\Delta v|}{v}, \tag{8}$$

where ρ is the lens aperture radius, v denotes the distance from the image plane (focus plane) to the lens, and Δv indicates the separation between that plane and the employed sensor plane. From Eqs. (1) and (2), setting equal d and different F, v values for the two cases of F_r, v_r and of F_g, v_g, we arrive at $\Delta v=v_r-v_g=v_r v_g f$. Taking into account the colours of the distinct and blurred parts in every case, the radii for the red and green blurs are

$$b_{r,g}=\rho\,v_{r,g}f \tag{9}$$

By combining these relations and (1) again, we obtain two possible forms for the d distance in terms of the blur radii, which are

$$d=\frac{F_r}{1-\frac{F_r f\rho}{b_r}} \tag{10}$$

$$=\frac{F_g}{1-\frac{F_g f\rho}{b_g}}=\frac{F_g}{1-\frac{b_\infty}{b_g}}, \quad b_\infty\equiv F_g f\rho. \tag{11}$$