



Feedback from higher to lower visual areas for visual recognition may be weaker in the periphery: Glimpses from the perception of brief dichoptic stimuli



Li Zhaoping

University College London, United Kingdom

ARTICLE INFO

Article history:

Received 2 November 2016

Received in revised form 13 April 2017

Accepted 1 May 2017

Available online 7 June 2017

Number of Reviews = 2

Keywords:

Central vision

Peripheral vision

Dichoptic stimuli

Visual decoding

Top-down feedback

Primary visual cortex (V1)

Analysis-by-synthesis

ABSTRACT

Eye movements bring attended visual inputs to the center of vision for further processing. Thus, central and peripheral vision should have different functional roles. Here, we use observations of visual perception under dichoptic stimuli to infer that there is a difference in the top-down feedback from higher brain centers to primary visual cortex. Visual stimuli to the two eyes were designed such that the sum and difference of the binocular input from the two eyes have the form of two different gratings. These gratings differed in their motion direction, tilt direction, or color, and duly evoked ambiguous percepts for the corresponding feature. Observers were more likely to perceive the feature in the binocular summation rather than the difference channel. However, this perceptual bias towards the binocular summation signal was weaker or absent in peripheral vision, even when central and peripheral vision showed no difference in contrast sensitivity to the binocular summation signal relative to that to the binocular difference signal. We propose that this bias can arise from top-down feedback as part of an analysis-by-synthesis computation. The feedback is of the input predicted using prior information by the upper level perceptual hypothesis about the visual scene; the hypothesis is verified by comparing the feedback with the actual visual input. We illustrate this process using a conceptual circuit model. In this framework, a bias towards binocular summation can arise from the prior knowledge that inputs are usually correlated between the two eyes. Accordingly, a weaker bias in the periphery implies that the top-down feedback is weaker there. Testable experimental predictions are presented and discussed.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The most obvious difference between central and peripheral vision in humans is the spatial resolution of visual perception (Anstis, 1974). This difference is in line with the differential brain resources devoted to central and peripheral vision. Each unit solid angle in the central visual field has devoted to it more photoreceptors and retinal ganglion cells, and a larger surface area in primary visual cortex and most of the extrastriate visual cortical areas (Osterberg, 1935; van & Anderson, 1995) than the same angle in the periphery. One might therefore wonder whether central and peripheral vision would be equal if the different spatial resolutions were compensated for by scaling the visual inputs spatially (Koenderink, Bouman, de Mesquita, & Slappendel, 1978).

However, a more important difference between central and peripheral vision is the extra difficulty in recognizing visual shape or form in the latter even when visual inputs in the periphery are

enlarged to compensate for the lower spatial resolution, see (Strasburger, Rentschler, & Jüttner, 2011) for a review. This peripheral problem is twofold, one is the difficulty or impossibility of recognizing a single item or pattern, such as a numeral (Strasburger, Rentschler, & Harvey, 1994; Strasburger & Rentschler, 1996) or a spatial pattern for hyperacuity tasks (Fendick & Westheimer, 1983; Westheimer, 1982), even after input enlargement; the other is the additional difficulty when the input pattern to be recognized is surrounded by neighboring inputs (Anstis, 1974; Pelli, Palomares, & Majaj, 2004; Strasburger, Harvey, & Rentschler, 1991; Strasburger, 2014), for review see (Levi, 2008; Strasburger et al., 2011). The latter is referred to as crowding, as if the contextual inputs inhibit recognition by making the visual scene more cluttered. The underlying neural causes for these peripheral difficulties are poorly understood; there are only some observations suggesting that visual cortical areas like V1, V2, and V4 might be involved (Chen et al., 2014a; Flom, Heath, & Takahashi, 1963; Levi, 2008; Strasburger et al., 1994; Strasburger & Rentschler, 1996; Tyler & Likova, 2007).

E-mail address: z.li@ucl.ac.uk

A different perspective on the differences between central and peripheral vision comes from considering visual attention. The brain has limited resources; thus only a small fraction of visual inputs can be selected to be brought into the attentional spotlight for further processing. New destinations for selection typically start in the periphery; thus attending them happens in two stages. First, the attentional spotlight is shifted very briefly to the peripheral target. Second, a saccade brings that destination to the center of the visual field (Hoffman, 1998). We refer to this process of selecting the small fraction of inputs and bringing them into the attentional spotlight as the process of looking.

After looking is complete, another process, which we refer to as seeing, focuses on the selected inputs within the attentional spotlight. It recognizes or infers properties of visual scenes. While a single word “attention” is often used to refer to both looking and seeing, it is important to note that looking involves visual selection (attentional selection of visual inputs) while seeing involves visual decoding (of inputs within the attentional spotlight) (Zhaoping, 2014). In natural visual behavior, shifting one’s gaze to attended inputs is mandatory (Deubel & Schneider, 1996; Hoffman, 1998). Thus, peripheral and central vision should differentially emphasize looking and seeing respectively (Zhaoping, 2014), with peripheral vision’s role in looking compromising its ability to see. This inferiority is manifest in its lower spatial resolution (Aubert & Foerster, 1857; Koenderink et al., 1978; Weymouth, 1958), reduced performance in recognition tasks such as hyperacuity and character recognition (Fendick & Westheimer, 1983; Strasburger et al., 1991; Strasburger & Rentschler, 1996), and its affliction to visual crowding (Levi, 2008; Pelli & Tillman, 2008; Strasburger et al., 1991; Strasburger et al., 1994).

Here, we probe differences in the ways that central and peripheral vision perform seeing or visual recognition tasks. These differences go beyond mere spatial resolution or visual input sensitivity; and so we examine them using computationally motivated visual psychophysical methods. In particular, we consider different influences over ambiguous percepts coming from prior knowledge about the statistics of natural visual inputs.

We use specially designed dichoptic inputs that have ambiguous perceptual interpretations; the relevant prior knowledge in this case is that the inputs to the two eyes are normally correlated (Li & Atick, 1994). This prior expectation can influence the perceptual outcome in a case of ambiguous perception by favoring the interpretation associated with similar or identical inputs to the two eyes over that associated with dissimilar or opposite inputs to the two eyes. To anticipate, we report that this prior influence was indeed observed and was significantly stronger in central than in peripheral vision.

Our particular prior influence must use information about the eye-of-origin of visual inputs to tell whether inputs from the two eyes are similar or dissimilar. This information is abundant in the primary visual cortex (V1) but is absent or very scarce beyond V1 along the visual pathway (Burkhalter & Van Essen, 1986; Hubel & Wiesel, 1968). Therefore, we argue that this prior influence cannot act in higher visual areas beyond V1 in a purely feedforward manner, but has to involve feedback processes to V1. In particular, we argue that it acts through the following processes which compute analysis-by-synthesis: (1) feedforward visual inputs from V1 suggest initial perceptual hypotheses about properties (e.g., the motion direction of a drifting grating) of the visual scene; (2) the visual inputs expected under each hypothesis are generated or synthesized according to prior knowledge of the visual world, (3) these synthesized visual inputs are fed back from higher brain areas to be compared with the actual visual inputs in the primary visual cortex, allowing the match between hypothetical and actual input to be assessed (Carpenter & Grossberg, 2011); and (4) a perceptual hypothesis is strengthened or weakened for

close or far matches, respectively. In these processes, the prior knowledge shapes the hypothesized inputs and, consequently, biases the final perceptual outcome. Accordingly, a weaker prior influence in the periphery should suggest that the feedback pathway in the periphery is weaker.

Note that although we suggested that peripheral and central vision differ according to their differential foci on looking and seeing, we investigated the difference by considering feedback influences in the seeing computation of analysis-by-synthesis. This paper makes no commitment to other computational or mechanistic differences. In particular, a weaker or absent feedback in peripheral vision for analysis-by-synthesis does not necessarily mean that visual attention does not operate in peripheral vision. After all, we argued above that peripheral vision has to play an important or leading role in attentional selection (looking). Some such selections are exogenous or stimulus driven, governed mainly by a bottom-up saliency map generated by feedforward and intracortical mechanisms (Li, 2002; Zhaoping, 2014) (if we exclude the mechanism of reading out this saliency map to guide gaze shifts through the midbrain). Other such selections are endogenous or driven by top-down factors such as in feature-based attention, by which visual inputs having a particular feature across the visual field are preferentially selected for further scrutiny (e.g., red-colored inputs across visual field are preferentially selected when one is searching for a red cup). Questions such as how feature-based attentional selection operates, e.g., whether it operates through feedback fibers from higher to lower brain areas or by weighting the feedforward fibers according to the goals of ongoing visual tasks (Müller, Reimann, & Krümmenacher, 2003; Wolfe, Cave, & Franzel, 1989), lie beyond the scope of this study.

Parts of the data in this study and their theoretical implications have been previously presented in abstract form (Zhaoping, 2013a, 2013b, 2015) and discussed in Zhaoping (2014).

2. Ambiguous perception using dichoptic stimuli

Shadlen and Carney (1986) presented two brief flashing horizontal gratings, one each to the two eyes; the two gratings differed from each other in both spatial and temporal phase by 90°, and the authors reported that observers perceived a drifting grating which was the sum of the two flashing gratings. More specifically, let S_L and S_R be the input stimuli to the left and right eye, respectively (for simplicity, the background luminance used to ensure that $S_{L,R}$ is non-negative is omitted in the following expressions),

$$\begin{aligned} S_L &= c \cdot \cos(k \cdot y + \phi_y) \cos(2\pi\omega \cdot t + \phi_t), \\ S_R &= c \cdot \sin(k \cdot y + \phi_y) \sin(2\pi\omega \cdot t + \phi_t), \end{aligned} \quad (1)$$

where y denotes vertical spatial location and t denotes time, ϕ_y and ϕ_t are two arbitrary phase values; $c \leq 1$ is the contrast amplitude of the gratings, and k and ω are their spatial and temporal frequencies, respectively. The binocular sum

$$\begin{aligned} S_+ &\equiv S_L + S_R \\ &= c \cdot \cos \left[k \cdot (y - 2\pi \frac{\omega}{k} t) + \phi_y - \phi_t \right] \end{aligned} \quad (2)$$

is the perceived drifting grating, drifting at speed $v = 2\pi\omega/k$.

However, it is known that, because natural inputs to the two eyes are correlated, the primary visual cortex (V1) represents these inputs more efficiently (Barlow, 1961; Li & Atick, 1994) by transforming them into two decorrelated channels: one for the sum of, and the other for the difference between, the inputs in the two eyes (see section 6.4 for physiological details and discussions). The sensitivities to the binocular summation and difference signals can independently adapt to the visual inputs statistics (May & Zhaoping, 2016; May, Zhaoping, & Hibbard, 2012). We thus predicted, and duly found, that the difference S_- between the two

Download English Version:

<https://daneshyari.com/en/article/5705858>

Download Persian Version:

<https://daneshyari.com/article/5705858>

[Daneshyari.com](https://daneshyari.com)