Reversed depth in anti-correlated random dot stereograms and the central-peripheral difference in visual inference

Li Zhaoping* and Joëlle Ackermann†
*University College London, UK, z.li@ucl.ac.uk
†ETH Zurich, Switzerland

Published in *Perception*, 2018, Volume: 47 issue: 5, page(s): 531-539

1 Abstract

In a random dot stereogram (RDS), the percept of object surfaces in a three-dimensional scene is generated by images presented to left and right eyes that comprise interocularly corresponding random black and white dots. The spatial disparities between the corresponding dots determine the depths of object surfaces. If the dots are anti-correlated, such that a black dot in one monocular image corresponds to a white dot in the other, disparity-tuned neurons in the primary visual cortex (V1) respond as if their preferred disparities become non-preferred and vice versa, thereby reversing the disparity signs reported to higher visual areas. Typically, when viewing anti-correlated RDSs presented in the central visual field, humans have great difficulty perceiving the reversed depth, or indeed any coherent depth at all. We report that the reversed depth is more easily perceived in the peripheral visual field, supporting a recently proposed central-peripheral dichotomy in the way that feedback from higher to lower visual cortical areas implements visual inference.

Keywords: Anti-correlated random dot stereograms, depth perception, central vision, peripheral vision, top-down feedback, primary visual cortex (V1), visual inference, analysis-by-synthesis.

2 Introduction

Julesz (1971) first reported that one cannot see depth or reversed depth in anti-correlated RDSs, although one can discriminate depth order when the density of dots in RDSs is very sparse (Cumming et al., 1998), and see veridical depth in simple stereograms in which the contrast between the eyes is reversed (Helmholtz, 1925; Cogan et al., 1995).

Fig. 1 shows three RDSs, each representing two depth planes, one of which contains a central disk (of

radius r) at disparity $d \neq 0$, and the other contains a surrounding ring of dots (within radius $R > r$) at zero disparity. The RDS in Fig. 1A is correlated: the two monocular images are identical to each other except for, (1), the dots within the central disk in the left-eye image are horizontally displaced from their corresponding dots in the right-eye image by a disparity d; and, (2), some dots near the border between the two depth planes are monocular due to surface occlusion. The disk appears in front of or behind the surrounding ring depending on whether $d > 0$ or $d < 0$. If the receptive field of a disparity-tuned V1 neuron covers only the dots for the central disk, the neuron would give high or low responses when the disparity d is its preferred or non-preferred disparity, respectively. However, if the contrasts of the dots within the disk are opposite in the two eyes, as in the anti-correlated RDS in Fig. 1C, the V1 neuron would respond as if the preferred disparity becomes non-preferred and vice versa (Cumming and Parker, 1997; Ohzawa et al., 1990; Haefner and Cumming, 2008). In this RDS, human observers usually cannot see depth, including the reversed depth reported by V1 neurons to higher visual areas, as the dot density is too high (Cumming et al., 1998).

Fig. 1B, the half-matched stereogram, is like Fig. 1A or Fig. 1C, except that, within the disk, each dot and its counterpart in the other monocular image have the same or opposite contrasts with equal probability. Such RDSs evoke veridical depth perception in human observers (Doi et al., 2011), and, perhaps due to response non-linearities, weaken, but do not reverse, V1 neurons' disparity preferences (Henriksen et al., 2016).

However, previous studies of depth perception in anti-correlated stereograms focused on the central visual field. We report that the reversed depth can be perceived in the peripheral visual field. Our experiments were modified from (Doi et al., 2011)'s design mainly by adding a peripheral viewing condition and using only smaller disparity values $|d| < 0.2^o$, so that they are within the range of V1 neurons' disparity preference particularly in the central visual field (Poggio et al., 1988).

In detail, in each trial, observers viewed a dynamic sequence of 15 RDS frames (at 100 ms each) resembling one of those in Fig. 1, and then reported (without feedback) whether the disk of dots appeared in front of or behind the surrounding ring. All frames in a sequence shared r, R, d; had dots of the same size and density; and had the same correlation property (correlated, anti-correlated, or half-matched) for the central dots. To view the RDSs centrally, observers fixated on a binocular red cross of size s_f at distance e_o (with $r < e_o < R$) above the center of the disk (Fig. 1); to view the RDSs peripherally, they fixated on a second binocular red cross added at a distance $e_1 > R$ above the disk center.

We randomly interleaved central and peripheral trials within an experimental session to reveal the central-peripheral difference more clearly. In Experiment 1, central and peripheral trials had the same visual input properties in terms of r, R, d, etc, except for the retinal image locations for the RDSs (and the red cross within it) and the additional red cross in the peripheral trials. Relative to the visual acuity or the sizes of the receptive fields of the V1 cells responsive to the RDSs, this makes the RDS smaller for the peripheral trials. To examine if the stimulus feature sizes were essential, in Experiments 2 and 3, the sizes of the dots, the disk, the ring, the cross, and the cross-to-disk distance were halved for the central trials. In Experiment

2, the disparity step size $|d|$ was also halved; in Experiment 3, $|d|$ was instead unchanged or doubled (from the $|d|$ in Experiment 1) for the central or the peripheral trials, respectively. To control for possible effects of the complexity of visual inputs at the boundary between the two depth surfaces, such as the monocular dots or, as suggested by Bruce Cumming (personal communication), a mixture of correlated and anti-correlated dots that could fall within a single neural receptive field, Experiments 4 and 5 were added. They respectively resembled Experiment 1 and Experiment 2 except that a blank ring (of background luminance) in each image separated the disk of dots from the surrounding ring of dots and no monocular dot was present.

3 Methods

The arrangement of the equipment was identical to that in previous studies (Zhaoping, 2012, 2017). Equipment, including a Mitsubishi 21-inch cathode-ray tube (CRT), an eye tracker, and a mirror stereoscope, were purchased from Cambridge Research System (CRS), and were calibrated (e.g., gamma correction of the CRT's signal-to-luminance relationship) by CRS software.

Each experimental session randomly interleaved trials of the six test stimulus conditions, with two fixation positions x three types of RDS correlations (shown in Fig. 1). There were 50 trials per condition; in each trial, the central disk was equally likely to be in front of or behind the surrounding ring. Each trial comprised in sequence: a pre-fixation period, which was ended when the observer pressed a button; a fixation period; a period in which the RDSs were displayed; and finally a period in which the observer reported the depth order. The display screen had the same grey background as those of the RDSs. The binocular red cross(es) displayed with the RDSs were present in each period. The cross within the surrounding ring was at the same position on the display screen across the trials, whether it was the fixation cross (in the central trials) or not. To cue observers to the fixation location, binocular text strings "press any button" and "for the next trial" were displayed during the pre-fixation period to the left and right, respectively, of the designated fixation cross. As in a previous study(Zhaoping, 2017), throughout each session, a black binocular rectangular frame was constantly on the display screen to anchor vergence. This frame enclosed the RDS region, the red crosses, and the texts. All the binocular items (the red crosses and the instruction texts) had the same binocular zero-disparity defined by the vergence anchoring frame.

The fixation period was 0.7 second for observers who were not gaze tracked (this applied to each session of one subject, the first author, who participated in all the experiments, and to one session of another subject whose gaze was monitored by the experimenter through a video camera), otherwise, it ended at the first moment after 0.7 second into the fixation period when the eye tracker (tracking the right eye) determined that the observer was fixating properly. Observers maintained fixation better in the central trials; but all maintained fixation properly in more than 90% of the trials (judged by the first author's self-report, experimental monitoring, and eye tracking (at 160 Hz) data analysis as in (Zhaoping, 2012, 2017)). Only one observer (the first author) was not naive to the purpose of the study.

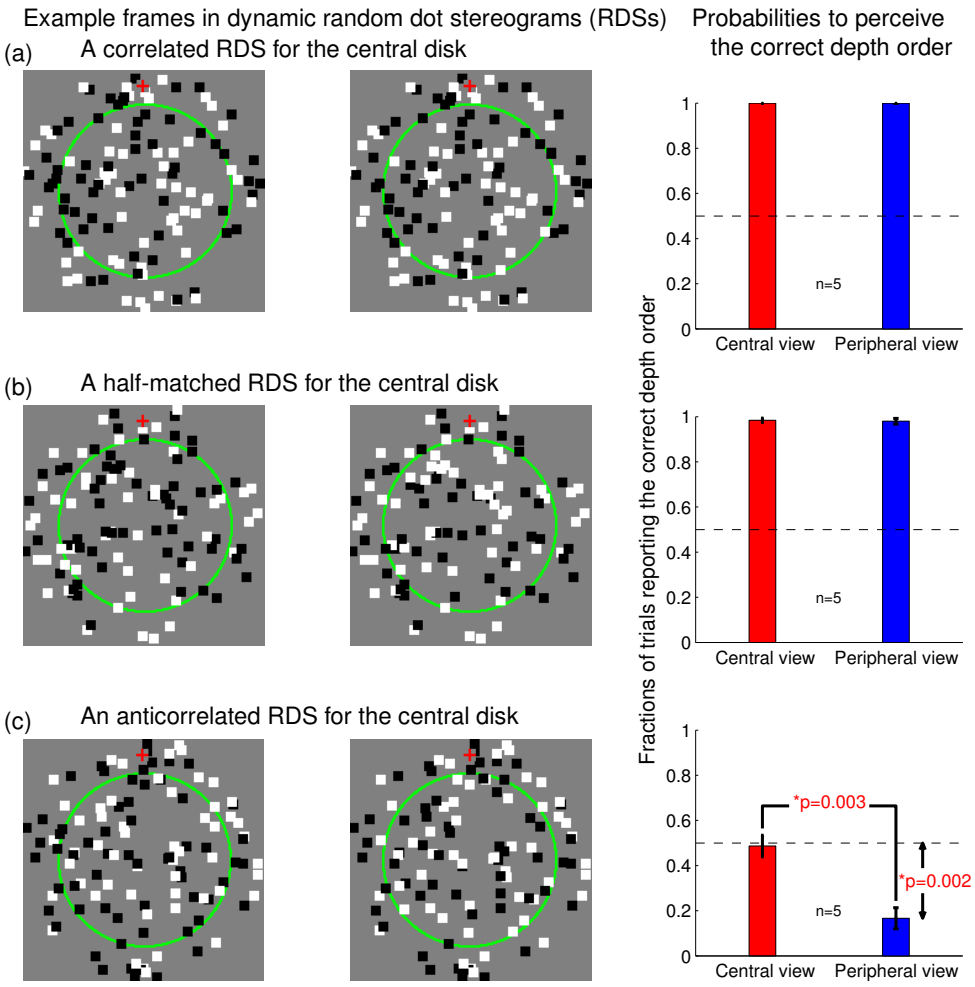

Figure 1: Examples of the RDS frames (left) for two underlying depth planes and the corresponding performance (right) of $n = 5$ observers in reporting which depth plane was closer in Experiment 1. In each RDS (the left and right monocular images), one depth plane of binocular disparity $d = \pm 0.087^o$ contains dots within a disk (radius $r = 3^o$, marked by green curves which were not shown in experiments) at the center of the images, the other depth plane at zero disparity contains the surrounding dots (within radius $R = 4.3^o$). All dots are squares with side length $s = 0.348^o$. In the correlated (A), half-matched (B), or anti-correlated (C) RDSs, the contrast (black or white) of each dot for the disk in one monocular image has a probability $p = 1, 0.5$, or 0, respectively, of matching the contrast of the corresponding dot in the other monocular image. All the dots in the ring surrounding the disk are binocularly matched in contrast, except for monocular dots near the circumference of the disk due to local surface occlusion. In the right column, the bars show the probabilities with which observers reported the correct depth order (defined by the disparity) of the two planes; the dashed line shows the chance probability (0.5). The red text linking two probability values indicates the p value associated with a t-test for whether these two values are statistically indistinguishable. Observers fixated on the binocular (zero disparity) red cross (size $s_f = 0.44^o$) at $e_o = 3.65^o$ or $e_1 = 10.1^o$ (not shown in the figure) above the center of the disk in order to view the RDSs centrally or peripherally.

Before an experimental session, each observer went through the following four groups of practice trials in order: (1), 10 central trials with fully correlated RDSs — in each trial, a static RDS frame for two seconds (rather than the dynamic RDSs for 1.5 seconds) was shown to make the task easier; (2), 10 trials as in (1) but using the dynamic (fully correlated) RDSs as in the real experimental condition; (3), 20 trials of the fully correlated experimental condition, of which 10 each, in a randomized order, were in central and peripheral views; (4), 6 example trials, one for each of the six experimental conditions. If the observer could correctly judge the depth order in no less than 90% of the trials in each condition of the practice trials in the groups (1), (2), and (3) above, he or she was then admitted to the fourth group of practice trials and the full experimental session. Roughly 50% of observers could not achieve this level of performance, particularly for the peripheral trials. Since a previous study (Doi et al., 2011) indicated that observers could adequately perceive the depth order in half-matched RDSs in central vision, if an observer incorrectly reported the depth order in more than 25% of the experimental trials with half-matched RDSs in either fixation position, the data of this session were excluded from further analysis (one such session was therefore removed). Significant results in data analysis were obtained by t-tests (matched-sample when comparing conditions within observers).

In Experiments 1–3, each frame of the dynamic RDSs was made independently of any other frame as follows. The process started with two concentric disks (radius r and $R > r$) of random dots. For each disk, each dot was placed at any location within the disk with equal probability and was set to be black or white with equal probability; a fraction $f = 25\%$ of the disk area would be covered by the dots if they did not occlude each other. The two monocular images of the RDS were made from these two disks of dots as follows. In the unit of pixel size, the disparity of the central disk was an integer d. If d was even, then the left and right monocular images contained the dots from the smaller disk (radius r) after these dots were shifted horizontally by $d/2$ and $-d/2$, respectively. If d was odd, then the magnitude of the shift in one image was $(|d| + 1)/2$ and in the other image was $(|d| - 1)/2$; the choice of which image had which magnitude was randomized. For each monocular image, the dots for the surrounding ring were those from the larger disk, excluding any dot that was either within the image area of the shifted smaller disk or would overlap in the monocular image with any dots from the smaller disk. The RDSs for Experiment 4/5 were made similarly, except that any dot in the larger disk was excluded if it was within radius R_1 ($r < R_1 < R$) from the center of the disk or if it was monocular.

In Experiments 1 to 5, respectively, the dot size in central/peripheral trials was $s = 0.348^o/0.348^o$, $0.174^o/0.348^o$, $0.174^o/0.348^o$, $0.348^o/0.348^o$, $0.174^o/0.348^o$; the radius of the central disk was $r = 3^o/3^o$, $1.5^o/3^o$, $1.5^o/3^o$, $3^o/3^o$, $1.5^o/3^o$; the surrounding ring had an outer radius $R = 4.3^o/4.3^o$, $2.15^o/4.3^o$, $2.15^o/4.3^o$, $4.7^o/4.7^o$, $2.35^o/4.7^o$, the disparity step was $|d| = 0.087^o/0.087^o$, $0.044^o/0.087^o$, $0.087^o/0.174^o$, $0.087^o/0.087^o$, $0.044^o/0.087^o$, the distance from the red cross in the surrounding ring to the center of the central disk was $e_o = 3.65^o/3.65^o$, $1.83^o/3.65^o$, $1.83^o/3.65^o$, $4.1^o/4.1^o$, $2.05^o/4.1^o$; the size of this red cross was $s_f = 0.44^o/0.44^o$, $0.22^o/0.44^o$, $0.22^o/0.44^o$, $0.44^o/0.44^o$, $0.22^o/0.44^o$. The distance from the peripheral

red cross (size 0.44^o) to the center of the central disk was $e_1 = 10.1^o$ in all experiments. The inner radius R_1 bordering the surrounding ring of dots was $3.65^o/3.65^o$ in Experiment 4 and $1.825^o/3.65^o$ in Experiment 5.

4 Results

When the disparity step $|d| = 0.087^o$ between the two depth planes in Experiment 1 (Fig. 1), observers could adequately judge the depth order when the dots for the central disk were correlated or half-matched. This was so for both the central and peripheral fixation, at, respectively, 0.65^o and 7.1^o above the closest depth discontinuity, in line with previous findings (Westheimer, 1982; Doi et al., 2011). However, with anti-correlated RDSs for the disk, observers' performance was at chance in central vision, as in previous studies (Cumming et al., 1998; Julesz, 1971) (also see Doi et al. (2011) for $|d|$ values within V1 neurons' disparity sampling range (Poggio et al., 1988)). However, in peripheral vision, performance was significantly lower than chance, i.e., was better than chance at perceiving the reversed depth order reported by V1 neurons (Qian, 1997), even though visual acuity is worse in peripheral vision.

Fig. 2 shows the results from anti-correlated RDSs in the control Experiments 2–5, implying that the central-peripheral difference continued to hold. Halving the sizes of relevant visual stimulus features for the central trials in Experiments 2 and 3 did not impact the depth perception substantially, whether $|d|$ was also halved centrally (Experiment 2) or instead doubled peripherally (Experiment 3). Hence, the sizes of the visual feature (Read and Eagle, 2000) are not critical for our results. Experiments 4 and 5, which were like Experiments 1 and 2 except for the introduction of a dot-free ring of thickness 0.65^o or 0.325^o between the central disk and the surrounding ring of dots and the removal of all the monocular dots, showed that the central-peripheral difference was not due to the complexity of visual inputs at the boundary between the two depth surfaces.

5 Discussion

Our study was designed to test a recent proposal (Zhaoping, 2017) that hierarchical feedback involved in perceptual inference is weaker or absent in peripheral vision compared to central vision. According to this proposal, perceptual inference in central vision is a form of analysis-by-synthesis: V1 responses are feedforward signals (e.g., reporting an input binocular disparity) that suggest to higher visual areas an initial hypothesis for a perceptual decision (e.g., as to the depth of an underlying object surface). If perception is unclear or ambiguous in central vision, the initial hypothesis is re-evaluated by the brain in three steps. First, according to brain's internal model or prior knowledge of the visual world, the higher brain centers synthesize a would-be visual input that should resemble the actual visual input if the hypothesis is correct. Second, the synthesized input is fed back to lower visual areas such as V1 for comparison with the actual visual input. Third, the strength of the initial perceptual hypothesis is reweighted according to the degree

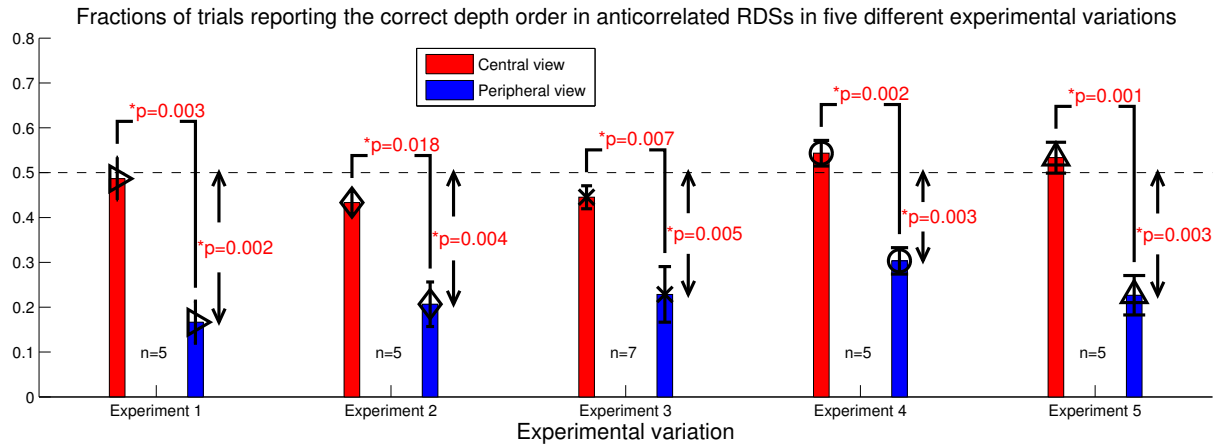

Figure 2: The tendency to perceive the reversed depth was stronger in peripheral than central vision regardless of the sizes of the visual stimuli or the presence of a dot-free gap between the image regions for the two depth surfaces. The performances of observers for anti-correlated RDSs are shown for five different experiments: Experiments 1–5 (involving $n = 5$ or 7 observers each). The results for Experiment 1 were reported in Fig. 1. Experiment 2 differed from Experiment 1 by halving the sizes of the dots, the disk, the ring, the fixation cross, the cross-to-disk distance, and the disparity step for central trials. Experiment 3 differed from Experiment 2 by doubling the disparity step in central/peripheral trials, making $|d| = 0.087^o/0.174^o$. Experiment 4 differed from Experiment 1 by the insertion of a dot-free background ring of thickness 0.65^o in the image space between the disk of dots and the surrounding ring of dots; Experiment 5 differed from Experiment 4 by halving all the relevant stimulus feature sizes (r, R_1, R, e_o, s, s_f, and $|d|$) for the central trials.

of the match between the synthesized and the actual inputs, such that a good or poor match strengthens or weakens, respectively, the initial hypothesis for the ultimate perceptual outcome. This inference process is called the feedforward-feedback-verify-weight (FFVW) process (Zhaoping, 2017).

Consider the application of this process to the perception of depth of the RDSs that we considered. When the input contains an anti-correlated RDS of disparity d, V1 neurons respond as if the input disparity is $-d$ (Cumming and Parker, 1997) and feed an initial hypothesis of this opposite disparity $-d$ forward. Higher brain centers generate the synthesized inputs of disparity $-d$; however, according to the (conventional) internal model of the visual world, these synthesized inputs to the two eyes are correlated rather than anti-correlated. The synthesized inputs are fed back to V1, and fail to match the actual, anti-correlated, inputs (or indeed the actual disparity d) as represented by the population V1 responses. Consequently, the initial hypothesis of a disparity $-d$ (the reversed depth) fails the verification, is weakened or vetoed by the re-weighting. This makes it hard to be perceived as the ultimate perceptual outcome.

According to the proposal, feedback is weaker in the peripheral visual field; this expectation is motivated by recent observations of tilt, motion, and color perception (Zhaoping, 2017). Accordingly, the initial hypothesis of a reversed depth in the peripheral visual field is not easily verified or vetoed, and is thus more

likely to become the perceptual outcome.

The FFVW process is based on the long-standing idea (MacKay, 1956; Carpenter and Grossberg, 1987; Li, 1990; Kawato et al., 1993; Dayan et al., 1995; Yuille and Kersten, 2006) that the brain uses both bottom-up and top-down processes (and knowledge) for sensory inference. The contribution of the recent work (Zhaoping, 2017) is to propose that the feedback in the FFVW process is weaker or absent in the peripheral visual field, using ambiguous perception of motion direction, tilt, and color of gratings to infer and reveal the central-peripheral dichotomy. The new contribution here is a test of the proposal using reversed depth perception. It was for this reason that we used small disparity values for the RDSs. According to the FFVW process, making the disparities lie within the range of preferred disparities of the V1 neurons (Poggio et al., 1988) allowed them to evoke feedforward depth signals. It is beyond our scope to explain stereopsis in general, or, in particular, the interesting observation (Doi et al., 2011, 2013) that humans can perceive reversed depth in anti-correlated RDSs in central visual field when disparity values (e.g., 0.24^o and 0.48^o) are larger than preferred by V1 neurons (Poggio et al., 1988).

Reversed depth perception in anti-correlated RDSs has an analogue in the domain of motion, namely reversed Phi motion. In this, when a picture is followed (via a dissolve) by its slightly displaced photo-negative, the perceived direction of motion is opposite to that of the physical displacement (Anstis, 1970; Anstis and Rogers, 1975). However, when the picture and its (displaced) photo-negative are presented in central vision successively without a temporal dissolve, the perceived direction of motion is indeterminate (Anstis and Rogers, 1975). This is analogous to the difficulties our observers had in seeing the reversed depth in central vision. It has also been noted that reversed Phi motion is more easily seen in the peripheral visual field (Anstis, 1970; Anstis and Rogers, 1975), while the physical motion direction is more easily seen in central visual field (Lu and Sperling, 1999), reflecting the same central-peripheral dichotomy that we report. However, perhaps because the horizontal (rather than, e.g., vertical) displacement between the two images is special in stereopsis but not so much in motion perception, the analogy between reversed depth and reversed motion does not extend to perception when information is integrated at different input orientations (Read and Eagle, 2000).

Reversed depth (in central vision) is more easily perceived in dynamic RDSs flickering with higher frequencies (Doi et al., 2013). This makes sense because faster flickering makes the appearance of each input frame too brief to be easily verified by feedback in the FFVW process. Indeed, in central vision, motion direction of drifting gratings having ocularly opponent contrasts is more likely perceived when the drifting frequency is high; similarly, the tilt of ocularly opponent gratings is more likely perceived through a shorter viewing duration (Zhaoping, 2017). Although anti-correlated RDSs rarely evoke reversed depth perception in central vision, they can oppose non-reversed depth perception when mixed with correlated RDSs in inputs (Neri et al., 1999), or evoke depth aftereffects by adapting V1 neurons (Hayashi et al., 2003), suggesting their contribution to the feedforward part of the FFVW process. If anti-correlated dots constitute the zero-disparity reference plane, they should activate, rather than inhibit, V1's tuned-(zero)-inhibitory cells (Poggio

et al., 1988) while activating the tuned-near and tuned-far cells equally. Consequently, this reference plane generates indefinite feed-forward signals and the resulting perception of its depth (and the sign of depth), regardless of the feedback process, makes it harder to discriminate the depth order between the testing and referencing planes, especially if depth perception of the testing plane itself is already weakened by reversed depths (Read and Eagle, 2000; Tanabe et al., 2008). Noting that reversed depth perception in some recent studies (Tanabe et al., 2008; Aoki et al., 2017) had observers fixate at more than 2^o from the depth edge, one could investigate whether the reversed perception in these studies would be weaker or stronger if fixation were less or more eccentric. Future studies should provide further falsifiable theoretical predictions to test the central-peripheral dichotomy against alternative hypotheses, and reveal the underlying neural mechanisms.

Acknowledgement

This study is supported in part by the Gatsby Charitable Foundation.

References

Anstis, S. (1970). Phi movement as a subtraction process. *Vision research*, 10(12):1411–IN5.

Anstis, S. M. and Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision research*, 15(8–9):957–961.

Aoki, S. C., Shiozaki, H. M., and Fujita, I. (2017). A relative frame of reference underlies reversed depth perception in anticorrelated random-dot stereograms. *Journal of Vision*, 17(12), article 17

Carpenter, G. and Grossberg, S. (1987). Art 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, 26(23):4919–4930.

Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., and Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception*, 24(1):33–47.

Cumming, B. G. and Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, 389(6648):280–283.

Cumming, B. G., Shapiro, S. E., and Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, 27(11):1367–1377.

Dayan, P., Hinton, G., Neal, R., and Zemel, R. (1995). The Helmholtz machine. *Neural Computation*, 7(5):889–904.

Doi, T., Takano, M., and Fujita, I. (2013). Temporal channels and disparity representations in stereoscopic depth perception. *Journal of Vision*, 13(13), article 26

Doi, T., Tanabe, S., and Fujita, I. (2011). Matching and correlation computations in stereoscopic depth perception. *Journal of Vision*, 11(3), article 1.

Haefner, R. M. and Cumming, B. G. (2008). Adaptation to natural binocular disparities in primate V1 explained by a generalized energy model. *Neuron*, 57(1):147–158.

Hayashi, R., Miyawaki, Y., Maeda, T., and Tachi, S. (2003). Unconscious adaptation: a new illusion of depth induced by stimulus features without depth. *Vision research*, 43(26):2773–2782.

Helmholtz, H. v. (1925). *Physiological Optics (translated by J P C Southall)*. New York: Optical Society of America, Dover Press.

Henriksen, S., Read, J. C., and Cumming, B. G. (2016). Neurons in striate cortex signal disparity in half-matched random-dot stereograms. *Journal of Neuroscience*, 36(34):8967–8976.

Julesz, B. (1971). *Foundations of cyclopean perception.* University of Chicago Press.

Kawato, M., Hayakawa, H., and Inui, T. (1993). A forward-inverse optics model of reciprocal connections between visual cortical areas. *Network: Computation in Neural Systems*, 4(4):415–422.

Li, Z. (1990). A model of olfactory adaptation and sensitivity enhancement in the olfactory bulb. *Biological Cybernetics*, 62(4):349–361.

Lu, Z.-L. and Sperling, G. (1999). Second-order reversed phi. *Attention, Perception, & Psychophysics*, 61(6):1075–1088.

MacKay, D. (1956). Towards an information flow model of human behavior. *British Journal of Psychology*, 47(1):30–43.

Neri, P., Parker, A. J., and Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401(6754):695–698.

Ohzawa, I., DeAngelis, G. C., Freeman, R. D., et al. (1990). Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science*, 249(4972):1037–1041.

Poggio, G. F., Gonzalez, F., and Krause, F. (1988). Stereocopic mechanisms in monkey visual cortex: binocular correlation and disiparity selectivity. *Journal of Neuroscience*, 8(12):4531–4550.

Qian, N. (1997). Binocular disparity and the perception of depth. *Neuron*, 18(3):359–368.

Read, J. C. and Eagle, R. A. (2000). Reversed stereo depth and motion direction with anti-correlated stimuli. *Vision research*, 40(24):3345–3358.

Tanabe, S., Yasuoka, S., and Fujita, I. (2008). Disparity-energy signals in perceived stereoscopic depth. *Journal of Vision*, 8(3), article 22.

Westheimer, G. (1982). The spatial grain of the perifoveal visual field. *Vision Research*, 22(1):157–162.

Yuille, A. and Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, 10(7):301–308.

Zhaoping, L. (2012). Gaze capture by eye-of-origin singletons: Interdependence with awareness. *Journal of Vision*, 12(2):article 17.

Zhaoping, L. (2017). Feedback from higher to lower visual areas for visual recognition may be weaker in the periphery: Glimpses from the perception of brief dichoptic stimuli. *Vision Research*, 136:32–49.