Vision Research 107 (2015) 15–21

Contents lists available at ScienceDirect

Vision Research

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

Are blur and disparity complementary cues to depth?

Michael S. Langer *, Ryan A. Siciliano

School of Computer Science, McGill University, Montreal H3A039, Canada

ARTICLE INFO

Article history:
Received 22 August 2014
Received in revised form 20 October 2014
Available online 4 December 2014

Keywords:
Binocular disparity
Blur
Defocus
Depth perception
Cue combinations

ABSTRACT

The image blur and binocular disparity of a 3D scene point both increase with distance in depth away from fixation. Perceived depth from disparity has been studied extensively and is known to be most precise near fixation. Perceived depth from blur is much less well understood. A recent experiment (Held, R. T, Cooper, E. A., & Banks, M. S. (2012). *Current Biology, 22*, 426–431) which used a volumetric stereo display found evidence that blur and disparity are complementary cues to depth, namely the disparity cue dominates over the blur cue near the fixation depth and blur dominates over disparity at depths that are far from fixation. Here we present a similar experiment but which used a traditional 3D display so that blur was produced by image processing rather than by the subjects' optics. Contrary to Held et al., we found that subjects did not rely more on blur to discriminate depth at distances far from fixation, even though a sufficient level of blur was available to do so. The discrepancy between the findings of the two studies can be explained in at least two ways. First, Held et al.'s subjects received trial-to-trial feedback in a training phase and may have learned how to perform the task using blur discrimination. Second, Held et al.'s volumetric stereo display may have provided other optical cues that indicated that the blur was real rather than rendered. The latter possibility would have significant implications about how depth is perceived from blur under different viewing conditions.

1. Introduction

When our eyes fixate on a point in 3D space, they both accommodate and converge to that point. Accommodation brings the point into sharp focus on each retina. Vergence brings the point to the center of each fovea where spatial resolution is highest. When accommodation and vergence are correct, the fixated point is in sharp focus and has zero binocular disparity. For scene points that are depths other than the fixation depth, their blur and disparity are proportional to the inverse distance (diopters) from the fixated point, with the disparity being roughly an order of magnitude larger than the blur width (Schechner & Kiryati, 2000).

Although blur and disparity both vary with inverse distance from fixation, there are differences in the visual system's sensitivity to these cues and how the visual system uses these cues in depth perception. Depth discrimination from disparity is very accurate near the fixation distance but it worsens rapidly with increasing distance from fixation, especially once diplopia occurs (Howard & Rogers, 2012). Depth discrimination from blur is much less well understood as we will discuss later. Blur discrimination itself is most accurate, not at the fixation depth, but rather at depths that are in front of and behind the fixation depth. JND's for blur obey a dipper function which achieves a minimum when the blur radius[1] is about 1 arcmin (Watson & Ahumada, 2011). In particular, there is a considerable range of depths around fixation over which all surfaces appear in focus, the so-called depth of field region.

This paper concerns the range of depths beyond the depth of field for which both disparity and blur cues are present. One might expect that over this range, the visual system combines the disparity and blur cues, for example, in a linear cue combination scheme (Landy et al., 1995). Mather and Smith examined disparities up to the limits of the fusion range but found little evidence for cue combination (Mather & Smith, 2000). This led them to an alternative hypothesis. Rather than estimating depth by combining blur and disparity cues, the visual system relies on disparities over the small depth range in which that cue is reliable, and it relies on blur to infer depth beyond that depth range. In this sense, disparity and blur would be complementary cues to depth.

* Corresponding author.
E-mail addresses: michael.langer@mcgill.ca (M.S. Langer), ryan.siciliano@mail.mcgill.ca (R.A. Siciliano).

[1] It is common to define the level of blur by the standard deviation of a Gaussian blur kernel, though we will also refer to a nominal blur "width". See Table 1 and surrounding text.

http://dx.doi.org/10.1016/j.visres.2014.10.036

Held, Cooper and Banks carried out a depth discrimination experiment that explored this hypothesis further (Held, Cooper, & Banks, 2012). Subjects discriminated the depths of two textured surfaces that both lay beyond fixation depth. Three cue combinations were tested: disparity-only, blur-only (monocular), and disparity-and-blur. At small disparity pedestals, the JND's were lower for the disparity-only condition than for the blur-only condition. This order reversed when the disparity pedestal became large. Most interesting is that, when both disparity and blur cues were present, the JND's followed the lower of the JND's for the disparity-only and blur-only condition. This suggests that subjects were relying more on the disparity cue for distances close to fixation and more on the blur cue for distances well beyond fixation[2] which is consistent with the hypothesis that blur and disparity are complementary cues to depth.

One concern that has been raised about Held et al. study is that JND's measure the precision of depth perception, but not the accuracy (Vishwanath, 2012). This is a well known distinction and indeed most studies of depth from disparity also addressed precision rather than accuracy (Ogle, 1953; Blakemore, 1970; McKee, Levi, & Bowne, 1990; Wilcox & Allison, 2009). The question of accuracy should not be neglected, however. For example, one of Held et al.'s stated motivations for studying depth perception for surfaces that are from the fixation depth is that these depth percepts would be needed for making eye movements and reaching movements.[3] But such movements surely require a high level of accuracy, not just precision.

Indeed there is evidence that, when disparities are large, depth perception becomes not merely imprecise but it also becomes inaccurate. For example, Richards and Kaye showed that perceived depth from disparity is not a monotonic function of physical disparity (Richards & Kaye, 1974). Rather it is a unimodal function: as the disparity increases, perceived depth at first increases but then it decreases to zero. A similar idea was discussed by Ogle (1952) who distinguished "patent stereopsis", where perceived depth increases as disparity increases, from "qualitative stereopsis" where only the sign of depth relative to fixation is perceived.[4] Ogle also noted that for sufficiently large disparities, no depth is perceived i.e. not even the sign.

In this paper, we present an experiment in which we attempted to confirm the findings of Held et al. Our experiment different from Held et al.'s in a few key ways, however. First, we used a conventional stereo display whereas they used a volumetric stereo display (Love et al., 2009). Second, our subjects had only a few minutes of training and were not given any feedback, whereas their subjects had 30 min of training with trial-to-trial feedback in all three conditions. We were concerned that the training given to Held et al.'s subjects may have led them to perform the task based on perceived blur when it was present, rather than on perceived depth (Vishwanath, 2012). Indeed Held et al. reported that one of the two naïve subjects was aware of the correlation between blur and depth and in the blur-only condition sometimes judged the blurrier stimulus as farther.

Our experiment consisted of two parts. The first part was a depth discrimination task which corresponded to the experiment of Held et al., with some differences mentioned above and others that will be described later. The second part was a blur discrimination task. The purpose was to verify that there was a sufficient amount of blur present in the stimuli for subjects to use the blur cue in the first part, where the task was to discriminate depth.

[2] Held et al. noted that they did not have sufficient statistical power to distinguish a cue switching strategy from an optimal cue combination strategy.

[3] Strictly speaking, eye movements and accommodation do not require a depth estimate. Rather they just require a disparity estimate or blur estimate, respectively. Reaching movements do require a depth estimate though.

[4] Richards and Kaye's plots are not entirely consistent with Ogle's characterization. We assume that patent stereopsis corresponds roughly to the increasing segment of the Richards and Kaye plots and qualitative stereopsis corresponds to the downward sloping segments of Richards and Kaye's plot.

2. Methods

2.1. Stimuli

The experiment was run on a Dell Precision M6700 laptop. The stimuli were generated and controlled using PsychoPy (Peirce, 2007). Stereo images were presented using 3D Vision shutter glasses by NVidia. The display screen was 1920×1080 pixels. Viewing distance was 63 cm. At this distance, each pixel subtended about $1'$ (one arcmin) of visual angle. From now on, we use units of pixels and arcmin interchangeably. The gamma of the monitor was measured to be 2.0. The images were gamma corrected so that luminance was proportional to digital gray level.

The stimuli were similar to those used by Held et al. Each image consisted of a foreground occluder and two background surfaces. The occluder was a texture composed of a grid of square tiles. Each tile was of size $64' \times 64'$. The occluder contained a fixation cross. See Fig. 1. The occluder also defined two window panels, each $512' \times 128'$ through which a background reference and test surface were shown. These background images each consisted of white squares randomly placed on a black background. Each square was $16' \times 16'$ and the density was 4 squares/deg^2. The size and density of squares was similar to the stimuli used in Held et al.

The background textures were defined offline prior to the experiment, as follows. First, a background texture of size $512' \times 512'$ was generated by placing small white squares uniformly randomly on a black background. This background texture was then blurred by a set of 2D Gaussians with varying standard deviations σ and these blurred textures were stored. On each trial of the experiment, a random cropped window from a blurred background texture was selected for the reference and for the test. Disparities were produced by selecting a cropped region for the left eye and a shifted cropped window for the right eye.

The set of reference disparities used in the experiment are listed in the first column of Table 1. These reference disparities ranged from 0 to $96'$ in steps of $24'$. For each disparity value, we define a nominal blur width ω such that the disparity to blur width ratio is 12:1, which corresponds to the ratio of the interocular distance to the pupil diameter, assuming a pillbox blur kernel (Held, Cooper, & Banks, 2012). Rather than using a pillbox kernel for blur

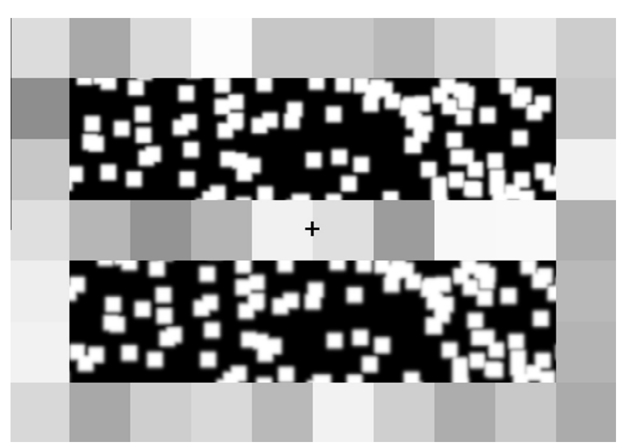

Fig. 1. Example stimulus. The top and bottom windows were rendered with a blur width ω of $6'$ and $7'$ respectively. (See Table 1.) The image should be viewed such that each foreground tile spans $64' \times 64'$, so width of just over 1 deg. See text for details.