Perception, 2007, volume 36, pages 696 – 702

DOI:10.1068/p5406

Vertical disparity affects shape and size judgments across surfaces separated in depth

Lisa M O'Kane§#, Paul B Hibbard¶#
§School of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, Scotland, UK; #School of Psychology, St Mary's College, University of St Andrews, South Street, St Andrews, Fife KY16 9JP, Scotland, UK; e-mail: pbh2@st-andrews.ac.uk
Received 5 January 2005, in revised form 23 June 2006; published online 27 April 2007

Abstract. Vertical binocular disparity provides a useful source of information allowing three-dimensional (3-D) shape to be recovered from horizontal binocular disparity. In order to influence metric shape judgments, a large field of view is required, suggesting that vertical disparity may play a limited role in the perception of objects projecting small retinal images. This limitation could be overcome if vertical disparity information could be pooled over wide areas of 3-D space. This was investigated by assessing the effect of vertical disparity scaling of a large surround surface on the perceived size and 3-D shape of a small, central object. Observers adjusted the size and shape of a virtual, binocularly defined ellipsoid to match those of a real, hand-held tennis ball. The virtual ball was presented at three distances (200, 325, and 450 mm). Vertical disparities in a large surround surface were manipulated to be consistent with a distance of 160 mm or infinity. Both shape and size settings were influenced by this manipulation. This effect did not depend on presenting the surround and target objects at the same distance. These results suggest that the influence of vertical disparity on the perceived distance to a surface also affects the estimated distance of other visible surfaces. Vertical disparities are therefore important in the perception of metric depth, even for objects that in themselves subtend only small retinal images.

1 Introduction

Binocular vision is an important source of information regarding the three-dimensional (3-D) shape of objects. As the two eyes view the world from slightly different vantage points, a point in the world will in general project to different positions in the two retinal images, ie it will have a certain binocular disparity. Information about the 3-D shape of objects is carried predominantly by the horizontal component of binocular disparity. Horizontal disparity does not, however, specify shape uniquely, since it is determined partly by the 3-D scene structure and partly by the viewing geometry (the orientation of the two eyes in the head).

One way in which this ambiguity can be overcome is to make use of the vertical component of binocular disparity. It has been shown that vertical disparity can in principle provide the information that is necessary in order to recover depth unambiguously from horizontal disparity (Longuet-Higgins 1982; Mayhew and Longuet-Higgins 1982; Gillam and Lawergren 1983; Bishop 1989). Vertical disparities are known to be used in the estimation of distance (Rogers and Bradshaw 1995), size (Bradshaw et al 1996), slant (Backus et al 1999), and curvature (Rogers and Bradshaw 1995) of objects and surfaces.

One limitation in the use of vertical disparities, however, is that they may be most useful only for surfaces subtending a large field of view. For example, Rogers and Bradshaw (1993) showed clear effects of vertical disparities on perceived depth, distance, and size, and suggested that previous failures to find such effects (Cumming et al 1991; Sobel and Collett 1991) may be attributable to the use of smaller stimuli in these studies. Such a dependence on image size is to be expected, since vertical disparities increase with increasing eccentricity, reaching a maximum at around ±45°.

¶Author to whom all correspondence should be addressed.

Whereas Cumming et al (1991) used stimuli subtending 11 deg, the stimuli used by Rogers and Bradshaw to show effects of vertical disparities subtended 75 deg. In a further study, Bradshaw et al (1996) showed that vertical disparities may only be of benefit for stimuli subtending 20 deg or more.

This limitation may have important consequences whether or not we are able to use vertical disparities in the perception of many objects that we may encounter in our everyday environment. Consider, for example, an apple, with a diameter of 8 cm, that is 40 cm away (comfortably within reaching distance). The image of the apple will subtend an angle of around 11 deg. Does this mean that vertical disparities will be of no use in estimating the shape, size, and distance of the apple? Vertical disparities are often neglected as a relatively unimportant cue in the perception of depth owing to this dependence on large image sizes (eg Durgin et al 1994; Tittle et al 1995). However, while in this example it is true that vertical disparities in the images of the apple itself would not be expected to be effective, it is possible that those from other objects or surfaces in the scene (eg the table on which it is resting, or the tree on which it is growing) might be used in estimating its distance, shape, and size.

Whether this possibility would occur depends on how vertical disparity is used by the visual system. Gårding et al (1995) distinguish between theories in which vertical disparities are used locally or more globally. Local theories are those in which vertical disparity is used at the same scale as the depth structures that are being recovered from horizontal disparity. According to this distinction, local theories would only predict an effect of vertical disparity if the local region were sufficiently large so as to allow them to be used. Consistent with this, Kaneko and Howard (1997) showed that manipulations of vertical disparity do not produce changes in perceived depth structure at a scale of less than around 20 deg (see also Adams et al 1996). As evidence that vertical disparity influences perceived depth over a spatial scale greater than the local depth features under consideration, Brenner et al (2001) showed that the perceived shape and size of a central object is influenced by manipulations of vertical disparity in a surround surface.

Clearly then, there exist non-local effects of vertical disparity on perceived 3-D shape. The issue addressed by the current paper is whether these effects persist for objects that are separated in depth. There are two reasons for questioning whether this would occur. First, pooling vertical disparities over too large an area, or over image regions separated by a large vertical disparity, may violate the assumptions made by some regional models, and therefore lead to unreliable results. For example, the model proposed by Mayhew and Longuet-Higgins (1982) assumes that the depth separation between viewed points is small relative to the fixation distance; this may not be true for image points on unconnected objects separated in depth (Yang and Purves 2003; Hibbard, in press).

Second, there is evidence that motion and disparity information are not always combined across distinct surfaces in order to improve perception. Retinal image motion, when combined with binocular disparity, could in principle allow the accurate recovery of 3-D shape, even though estimates of shape on the basis of either cue alone might be inaccurate (Richards 1985; Johnston et al 1994). This would then allow the possibility of accurately estimating the distance to a moving object, and also the distance to other objects in the scene. This could be done by making use of the relative disparity between the moving object and any other object, calibrated by the estimate of viewing distance provided by the combination of binocular disparity and motion information. This, in turn, should allow for the accurate perception of the shape of the second object. Brenner and Landy (1999) investigated scenes in which one object was defined by disparity and motion cues, and another, presented at a different distance, by disparity cues alone. They found that the shape of the moving object, but not the stationary object, was accurately perceived, suggesting that motion and disparity are not combined in this way.

It is certainly possible that vertical disparities across the entire image could be used in the determination of the distance of a particular object in the scene, regardless of the relative distances of objects within that scene. Longuet-Higgins (1982) showed that, for an observer fixating with a gaze angle g and a vergence angle v, the following equation holds for the locations of image points in images in the left and the right eye [(x', y'), (x'', y''), respectively]:

$$-\frac{x'}{y'}\sin(g+v)+\frac{1}{y'}\cos(g+v)=-\frac{x''}{y''}\sin(g-v)+\frac{1}{y''}\cos(g-v)\ . \tag{1}$$

It is possible to estimate the gaze and vergence angles by estimating the retinal location of each point in the image in the left and the right eye, so as to provide an equation of type (1). The set of equations produced by doing this for many points may be solved to provide estimates of v and g. In the simplest case, where one is looking at the object of interest, these may be used directly to determine object distance. It is also possible to determine the location of any other object in the scene, in addition to the one that is currently fixated, by combining estimates of gaze and vergence with the image location and disparity of objects of interest. Despite this possibility, the results of Brenner and Landy (1999) suggest that such opportunities might not be taken, and that it is only vertical disparities on surfaces at the same distance as the object of interest that might be used to determine its location, shape, and size.

We addressed this question by measuring the effect of manipulating vertical disparities in a large surround surface on the perceived shape and size of a smaller, central object. This latter task is important as it requires metric information in order to be solved. Frisby et al (1999) showed that vertical disparities may influence perceived depth when they are effective in altering the relief structure of surfaces, even for small fields of view. Only when tasks requiring a full metric reconstruction are used (Rogers and Bradshaw 1993) is image size an important consideration. Here, the central object was sufficiently small so that vertical disparities in its images would be expected to be ineffective for the size task employed, and was presented at a number of different distances relative to the surround, in order to determine whether vertical disparities demonstrate an influence over surfaces that are separated both spatially in the image and in depth. Vertical disparities in the surround were manipulated so as to be consistent with a distance that was either closer or further than the distance specified by vergence. This would be expected to lead to a correspondingly closer or further estimate of the fixation distance. Consequently, observers might then perceive the 3-D shape and size of the object as though it too was presented as distance that was closer

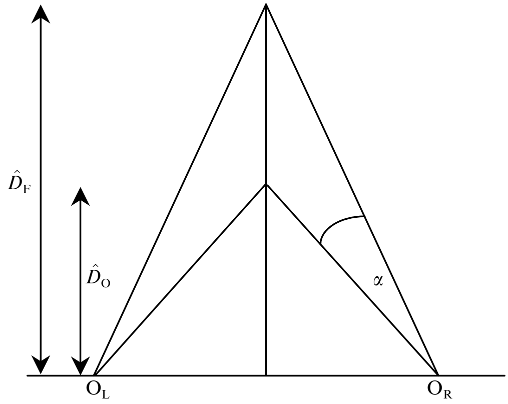

Figure 1. If we assume symmetrical fixation, an estimate of the fixation distance given by $\hat{D}_F$ and a relative disparity between the object of interest and the fixation point of 2α, then it is clear that, as $\hat{D}_F$ increases, and α remains constant, the estimate of the object distance, $\hat{D}_O$, will also increase. Thus, if the manipulation of vertical disparity increased the estimate of the fixation distance, this would have the effect of increasing the estimated distance of other objects in the scene.

or further than its actual distance, respectively. This would hold whether object distance was estimated directly on the basis of the fixation distance (in the case where the object is fixated), or by a combination of the fixation distance and the relative disparity between the fixated point and the object (in the case where it is not fixated). This is shown geometrically in figure 1, for the simple case of symmetrical fixation.

2 Method

2.1 *Observers*

Nine observers took part in the experiment, including the two experimenters.

2.2 *Apparatus*

Stimuli were presented on a 21 inch Sony Trinitron display. The resolution of the monitor was 800 × 600 pixels, and the refresh rate was 120 Hz. Separate left and right images were presented with Stereographics CrystalEyes liquid-crystal shutter glasses, so that the image in each eye was refreshed at a rate of 60 Hz. Only the red phosphor of the monitor was used, to minimise cross-talk between the views of the left and the right eye. The viewing distance was 325 mm, and was maintained by means of a chin-and-forehead rest. During the experiment, observers held a tennis ball (diameter 66 mm) in their left hand and adjusted the onscreen stimulus by means of the computer keyboard as described below.

2.3 *Stimuli*

Stimuli were random-dot stereograms, depicting a central ellipsoid framed by a frontoparallel reference surface. The size and shape (depth-to-width/height ratio) of the ellipsoid were set by the observer on each trial. Dots were positioned on the surface of the ellipsoid by the standard ray-tracing technique. The number of dots was varied, depending on the simulated distance of the surface, so that the dot density of the ellipsoid was roughly equal at all viewing distances (150 dots at the nearest viewing distance). Each dot had a Gaussian distribution of luminance, with a maximum brightness of 49.8 cd m^{-2} and a standard deviation of 1.9 min of arc. Three different distances to the centre of the ellipsoid were used: 200, 325, and 450 mm. The reference frame was a 34.2 deg × 44.2 deg frontoparallel plane. Dots were positioned at random on the plane; no dots were positioned in the central 25.0 deg × 21.6 deg region. The vertical disparities of the dots on the reference surfaces were varied, so as to be consistent with those that would be observed at one of two distances: 160 mm and infinity. The dot positions were initially calculated so they lay on a frontoparallel surface at a distance of 325 mm. The vertical disparity for a dot in this direction, at a distance of 160 mm or infinity, was then calculated, and the vertical position of the dot in the views of the left and the right eye adjusted accordingly. The two distances used were chosen so as to maximise any effect of vertical disparity. They were also chosen so as to equate the difference in the distance specified by vergence, and that specified by vertical-disparity scaling for the surround surface, in terms of the vergence angle required to fixate a surface at this distance.

2.4 *Procedure*

The observer's task was to set the shape and size of the ellipsoid so that they matched those of a real, hand-held ball (spherical, with a diameter of 66 mm). One pair of keys on the computer keyboard increased and decreased the size of the ellipsoid (isotropically in all directions), another its depth-to-width ratio. Observers were not given any specific instructions where to fixate during the experimental trials. Stimuli were presented in two experimental blocks. In each block, the vertical-disparity-scaled distance of the reference frame was held constant at one of the two values listed above.

Within a block of trials, the ellipsoid was presented at each of the three distances twelve times. The order of the distances within a block of trials was randomised. The order in which each observer performed the trial blocks was also randomised.

3 Results

3.1 *Object shape*

Figure 2a shows the mean object shape settings over the nine observers. The set depth-to-width ratio was in all cases less than one. Of primary interest in the current study is the effect of the manipulation of vertical disparities in the surround. A clear difference in the set depth-to-width ratio between the two conditions can be seen in figure 2a. Observers tended to set objects with relatively more binocular disparity when vertical disparities were consistent with a distance of 160 mm to the surround than when vertical disparities indicated a distance of infinity. This result was found to be significant in an analysis of variance ($F_{1,8} = 11.0$, $p < 0.05$). This is consistent with the object appearing closer to the observer in the former condition than in the latter, and is thus in line with our predictions.

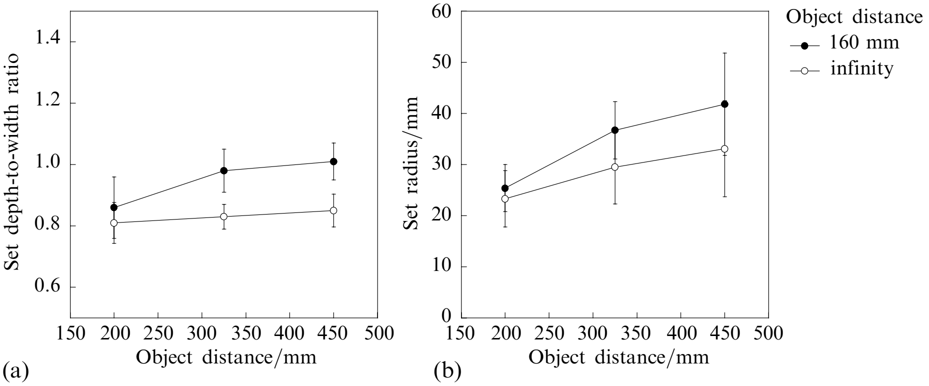

Figure 2. (a) Set object shape, and (b) set object size as a function of object distance, and vertical disparity scaling of the surround surface. Observers ($n = 9$) set objects that were both larger, and more extended in depth, when vertical disparity indicated a closer distance. Error bars show ±1 SD.

No significant effect of object distance ($F_{2,16} = 1.1$, ns) or interaction ($F_{2,16} = 0.786$, ns) were observed. Although an effect of object distance might have been expected, and some trend in this direction is evident in figure 2, the lack of a significant result may be attributed to the relatively narrow range of distances used here.

3.2 *Object size*

Figure 2b shows the mean set width and height of the object across the nine observers. The actual radius of the reference tennis ball was 33 mm. Again, the effects of vertical disparity and distance on these settings were tested with an analysis of variance.

A significant effect of the vertical disparity of the surround was observed ($F_{1,8} = 12.1$, $p < 0.01$). Observers set objects to be smaller when the surround vertical disparities were consistent with a 160 mm viewing distance, compared to when they were consistent with a distance of infinity. Again, this is consistent with the use of a closer distance to scale for the size of the tennis ball in the former case.

A significant effect of distance on size settings was also observed ($F_{2,16} = 42.3$, $p < 0.001$). The set size increased with increasing object distance, consistent with previous findings (Brenner and van Damme 1999). A significant interaction was also observed ($F_{2,16} = 3.9$, $p < 0.05$). Pairwise comparisons between the two vertical disparity conditions at each distance showed a significant effect of vertical disparity at the middle and far distance.

4 Discussion

The results reported here show a clear influence of the vertical disparity of a large, surrounding surface on the perceived size and shape of a small, central target object. These results are consistent with those reported by Brenner et al (2001), in showing that vertical disparity scaling in one surface can affect a task relating to the metric structure of another, smaller surface. The current result extends this to demonstrate that this effect occurs when the two surfaces are separated in depth.

As stated in the introduction, models of vertical disparity processing may usefully be classified as local or regional. The current results show a clear regional effect, spanning objects presented at different depths. These results are consistent with a global employment of vertical disparity in estimating the viewing parameters. Alternatively, vertical disparity could be used more locally in determining the distance to the surround surface, which might then affect the perception of the central object indirectly via its relative disparity. The current results are thus consistent with any model in which vertical disparity could influence estimation of the distance to the surround surface.

Gårding et al (1995) proposed that the effects of vertical disparity can be broken down into disparity correction, to recover surface shape up to a relief transform, and disparity normalisation, to recover full metric shape. Because we found an effect on a metric-size-estimation task, these results cannot be explained in terms of disparity correction, but must relate to a stage of disparity normalisation. These results do not speak to the question of whether the first stage is in fact implemented. Indeed, there is compelling evidence that the effects of vertical disparity are much stronger at this earlier stage of processing (Gårding et al 1995). However, the current results show pooling across objects separated in depth at the disparity-normalisation stage. Effects at the disparity-normalisation stage would be particularly important given the increased need for pooling over a wide area at this stage. Moreover, the need for this pooling to incorporate image features across a range of distances is important, since the variability in the distances to objects in an image region will increase as the size of that region increases (Yang and Purves 2003; Hibbard, in press).

The results found here therefore differ from those relating to the influence of motion information in one object on the perceived shape of another object, where it is found that the perceived shape of the latter object remains unaffected by the motion in the former. This might be explained if motion information is used to provide veridical shape information directly, rather than by acting to improve estimates of object distance used for scaling other sources of information such as binocular disparity. This may be contrasted with the case of vertical disparities, which appear to play a greater role in the estimation of the distances to surfaces.

The current results extend reports (Adams et al 1996; Kaneko and Howard 1997; Brenner et al 2001) that vertical disparity may be pooled over relatively large image regions, to show that this may be done so as to affect the perceived 3-D shape and size of objects that are clearly separated both spatially and in depth. This is an encouraging result, since in situations in which accurate metric shape information would be most useful (objects of graspable size within a reachable distance), the angle subtended by individual objects may be insufficient to allow for the use of vertical disparity to influence estimates of their metric properties. The non-local influences demonstrated here would thus allow for a role for vertical disparities in situations in which they would be particularly useful.

References

Adams W, Frisby J P, Buckley D, Gårding J, Hippisley-Cox S D, Porrill J, 1996 "Pooling of vertical disparities by the human visual system" *Perception* **25** 165 – 176

Backus B T, Banks M S, Ee R van, Crowell J A, 1999 "Horizontal and vertical disparity, eye position, and stereoscopic slant perception" *Vision Research* **39** 1143 – 1170

Bishop P O, 1989 "Vertical disparity, egocentric distance and stereoscopic depth constancy: a new interpretation" *Proceedings of the Royal Society of London, Series B* **204** 415 – 444

Bradshaw M F, Glennerster A, Rogers B J, 1996 "The effect of display size on disparity scaling from differential perspective and vergence cues" *Vision Research* **36** 1255 – 1264

Brenner E, Damme W J M van, 1999 "Perceived distance, shape and size" *Vision Research* **39** 975 – 986

Brenner E, Landy M S, 1999 "Interaction between the perceived shape of two objects" *Vision Research* **39** 3834 – 3848

Brenner E, Smeets J B J, Landy M S, 2001 "How vertical disparities assist judgements of distance" *Vision Research* **41** 3455 – 3465

Cumming B G, Johnston E B, Parker A J, 1991 "Vertical disparities and the perception of three-dimensional shape" *Nature* **349** 411 – 413

Durgin F H, Proffit D R, Olson T J, Reinke K S, 1994 "Comparing depth from binocular disparity to depth from motion" *Journal of Experimental Psychology: Human Perception and Performance* **21** 679 – 699

Frisby J P, Buckley D, Grant H, Gårding J, Horsman J M, Hippisley-Cox S D, Porrill J, 1999 "An orientation anisotropy in the effects of scaling vertical disparities" *Vision Research* **39** 481 – 492

Gårding J, Porrill J, Mayhew J E W, Frisby J P, 1995 "Stereopsis, vertical disparity and relief transformations" *Vision Research* **35** 703 – 722

Gillam B, Lawergren B, 1983 "The induced effect, vertical disparity, and stereoscopic theory" *Perception & Psychophysics* **34** 121 – 130

Hibbard (in press) "A statistical model of binocular disparity" *Visual Cognition*

Johnston E B, Cumming B G, Landy M S, 1994 "Integration of stereopsis and motion shape cues" *Vision Research* **34** 2259 – 2275

Kaneko H, Howard I P, 1997 "Spatial limitation of vertical-size disparity processing" *Vision Research* **37** 2871 – 2878

Longuet-Higgins H C, 1982 "The role of the vertical dimension in stereoscopic vision" *Perception* **11** 377 – 386

Mayhew J E W, Longuet-Higgins H C, 1982 "A computational model of binocular depth perception" *Nature* **297** 376 – 379

Richards W, 1985 "Structure from stereo and motion" *Journal of the Optical Society of America A* **2** 343 – 349

Rogers B J, Bradshaw M F, 1993 "Vertical disparities, differential perspectives and binocular stereopsis" *Nature* **361** 253 – 255

Rogers B J, Bradshaw M F, 1995 "Disparity scaling and the perception of frontoparallel surfaces" *Perception* **24** 155 – 179

Sobel E C, Collett T S, 1991 "Does vertical disparity scale the perception of stereoscopic depth?" *Proceedings of the Royal Society of London, Series B* **244** 87 – 90

Tittle J S, Todd J T, Perotti V J, Norman J F, 1995 "Systematic distortion of perceived 3-dimensional structure-from-motion and binocular stereopsis" *Journal of Experimental Psychology: Human Perception and Performance* **21** 663 – 678

Yang Z, Purves D, 2003 "A statistical explanation of visual space" *Nature Neuroscience* **6** 632 – 640

© 2007 a Pion publication

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 36 2007

www.perceptionweb.com

Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.