QUALITY, QUANTITY AND PRECISION OF DEPTH PERCEPTION IN STEREOSCOPIC DISPLAYS

Alice E. Haines, Rebecca L. Hornsey and Paul B. Hibbard

Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ UK.

ABSTRACT

Stereoscopic 3D viewing (S3D) can create a clear and compelling improvement in the quality of the 3D experience compared with 2D displays. This improvement is distinct from any change in the amount of depth perceived, or the apparent 3D shapes of objects and the distances between them. It has been suggested instead that the enhanced feeling of realness is associated more with the precision with which we see depth. We measured the contribution of stereoscopic cues to the quality of depth perception in simple abstract images and complex natural scenes. We varied the amount of disparity present in the simple scenes in order to dissociate the magnitude and precision of perceived depth. We show that the qualitative enhancement of perceived depth in stereoscopic displays can be readily quantified, and that it is more closely related to the precision than to the magnitude of apparent depth. It is thus possible to make a distinction between scenes that contain more depth, and those that contain better depth.

Index Terms— Depth perception, stereopsis, S3D, complex natural scenes, precision, magnitude.

1. INTRODUCTION

The perception of depth in stereoscopic 3D (S3D) is enhanced through the provision of binocular cues to depth. These binocular cues consist of binocular disparities, or differences in the images seen by the two eyes, and the changes in binocular convergence as the viewer fixates on points at different depths in the scene. By providing additional information about depth, binocular cues are able to alter the apparent structure of the depicted 3D scene, and the precision with which this is perceived.

The enhancement of depth perception in S3D can be understood by considering how the visual system combines multiple cues to depth. In S3D, it is convenient to divide depth information between binocular cues, and pictorial depth cues that are available in the monocular images. These cues include, for example, linear perspective, shape from shading, and occlusion [3].

Given multiple, imperfect and possibly conflicting cues, an optimal strategy to estimate depth is to calculate a weighted average of the estimates from the individual cues, with the weights determined by the inverse of the reliability of each individual cue [15]. This weighted cue averaging predicts, for example, that a greater magnitude of depth should be seen in S3D than with binocular viewing of traditional, 2D images. In the stereoscopic case, binocular depth cues provide information about the 3D structure of the scene. In contrast, in binocular picture viewing binocular cues indicate that the scene if flat. The combination of pictorial cues to depth, and binocular cues to flatness, can thus explain why there is a relative flattening of depth in traditional pictures in comparison with stereoscopic pictures [4].

Cue combination models also predict that the precision of depth estimates will increase when more cues are available [15]. Thus, when depth is specified by both perspective and binocular cues, the reliability of our depth estimates should be greater than when either cues alone is available. This predicted increase in precision is seen both in the reduction of just noticeable differences in depth [7], and in how depth information is used to control action [8].

A further way in which our perception of depth is affected by the presence of binocular cues is a qualitative change in the depth appearance in the scene. This striking improvement in appearance is related to the concept of 'stereopsis'. Literally meaning 'solid sight', this term refers to the vivid impression of solidity and depth that is perceived in S3D, but is lacking in non-stereoscopic displays [18]. This experience of the realness and solidity of depth is a component of the assessment of quality in S3D [10,13].

This qualitative experience of stereopsis in S3D, and the inferior depth experience in non-stereoscopic displays, have sometimes been linked to the conflict between binocular and pictorial cues in the latter case and also to the greater magnitude of depth perceived in the former [19]. However, it has also been proposed that stereopsis is associated with the reliability and precision of perceived depth [6,17], particularly with egocentrically scaled metric depth [18,19]. The purpose of the current study was to provide empirical data to test between these theoretical ideas.
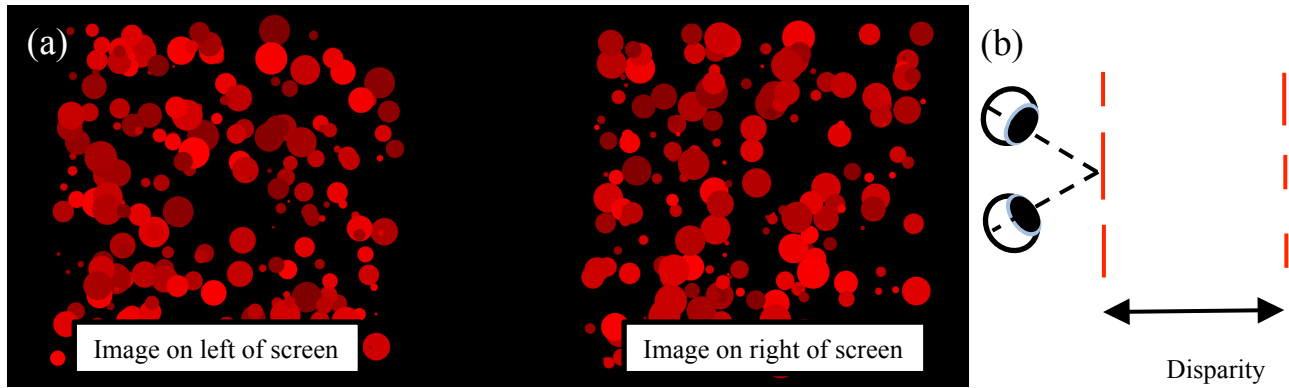
Fig. 1 (a) An example of the random circle stereogram stimuli. On each trial, two stimuli were presented, on the left and right of the screen. **(b)** The dots appeared to the participant on two separate planes in depth; the two stimuli on each trial had a different disparity.

In the current study, we conducted a direct experimental test of the association of stereopsis with the magnitude and precision of perceived depth. To do this, we measured the magnitude and quality of perceived depth as a function of the disparity in stereoscopic images. Increasing disparity is expected to increase the magnitude of perceived depth, since disparity scales approximately linearly with depth. However, the precision with which we are able to discriminate between two depth intervals decreases as disparity increases [1].

If the qualitative experience of real depth that is associated with stereopsis increases with increasing magnitude of perceived depth, then it should increase with increasing disparity. Conversely, if depth quality increases with the precision with which depth is perceived, then it is should decrease with increasing disparity. To test this, we measured the magnitude, quality and precision of perceived depth as a function of disparity in simple abstract scenes. In a second experiment, we also showed that depth quality is a directly measurable and quantifiable property of stereoscopic images of complex natural scenes.

2. METHODS

2.1 Apparatus

Stimuli were presented on a VIEWPIXX 3D monitor, viewed from a distance of 96cm. The monitor screen was 52cm wide and 29cm tall. The screen resolution was 1920x1080 pixels, with a refresh rate of 120Hz. Each pixel was 0.28mm square and subtended 1 arc min. Stimuli were presented at 8 bit resolution. Stereoscopic presentation was achieved using a 3DPixx IR emitter and

NVIDIA 3D Vision LCD shutter glasses. Participants' responses were recorded using a RESPONSEPixx response box. Stimuli were generated and presented using MATLAB and the Psychophysics Tool box extensions [2,11].

2.2 Participants

10 participants, including all the experimenters, completed the first experiment, in which the magnitude and quality of depth were assessed. 6 participants, including the experimenters, completed the other two experiments, in which we measured the precision of depth judgments, and depth quality in stereoscopic photographs of complex natural scenes. Participants did not receive any training.

2.3 Stimuli

Stimuli consisted of two-plane random-circle stereograms, and natural stereoscopic photographs.

Each random-circle stereogram consisted of 100 red circles, presented uniformly and randomly within a 10x10 degree square region. The diameter of each circle was set at a randomly chosen value between 5 and 60 arc min, and its luminance at a randomly chosen value between 50 and 100% of the maximum possible luminance. Half of the circles were presented with zero disparity, and so appeared at the depth of the screen. The other half were presented with an uncrossed disparity, and appeared on a plane behind the screen. For each stimulus, the magnitude of this disparity was 15, 30, 60, 90, 120 or 150 arc min.

On each trial, two stereograms were presented. The stimuli were vertically centered on the screen, at eye-

Fig. 2. An example of the stereoscopic images used, arrange for uncrossed (left and centre) or crossed (centre and right) viewing.

height, with one presented 7.7 degrees to the left of centre, and the other at 7.7 degrees to the right. Example stimuli are shown in Fig 1.

Stereoscopic photographs were of complex natural scenes, containing 'still-life' arrangements of vegetables and plants. An example image pair is provided in Fig. 2. Images were taken with an inter-camera distance of 6.5cm; full details of the camera setup and calibration are described elsewhere [5].

2.4 Procedure

2.4.1 Depth magnitude in random-circle stereograms

On each trial, two stereograms were presented, and the participant's task was to determine which had the greater depth between the front and back plane. The presentation time was unlimited, and the stimuli were presented until the participant indicated their response by pressing the left or right key on the response box. The two stereograms had a different disparity between the two planes. All 15 possible pairs of different disparities were presented 10 times, in random order, within a block of trials. Two blocks of trials were completed, resulting in 20 repetitions of each disparity pair. The participant's responses were used to generate a scale of perceived depth magnitude from these two-alternative forced choice responses, as explained in detail in section 3.1.

2.4.2 Depth quality in random-circle stereograms

The stimuli were exactly the same as for the depth magnitude task, but in these blocks of trials the participants were asked to judge the quality of apparent depth. They were asked to judge which of the two stimuli looked more convincing as a 3D scene, and more tangible, solid and real. These forced choice responses were used to generate a scale of depth quality, in the was as for depth magnitude.

2.4.3 Just noticeable differences in depth in random-circle stereograms
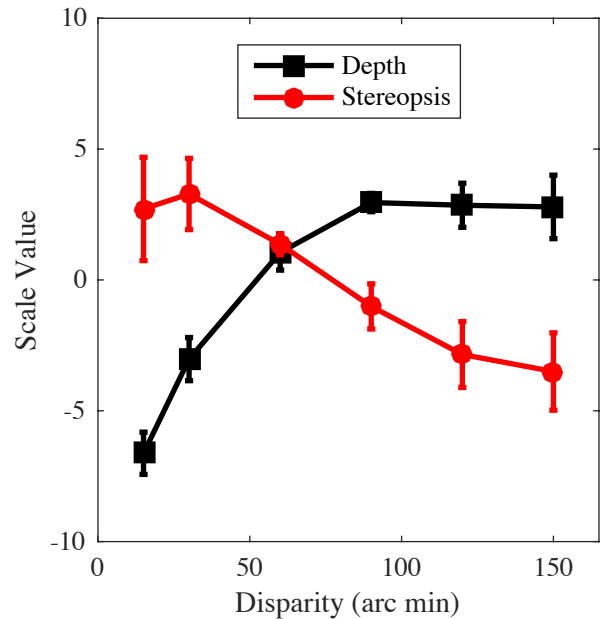

Fig. 3. Apparent depth magnitude (black squares) and depth quality (red circles) as a function of the stimulus disparity for random-circle stereograms. Error bars show ±1 standard error.

The stimuli and task were the same as for the depth magnitude task, except that disparity values were chosen that allowed us to estimate the just noticeable difference in depth between the two planes. Within a block of trials, one stimulus was always presented with a reference disparity of 15, 30, 60, 90, 120 or 150 arc min. The other was presented with one of 9 values, chosen on the basis of pilot experiments so that the extreme values were reliably seen as a smaller or larger depth interval than the reference stimulus. The position of the reference stimulus was randomly selected to be on the left or right of the screen on each trial. Within each block of trials, each of the 9 stimulus pairs was presented 40 times.

2.4.4 Depth quality in natural stereoscopic photographs

We also used a pairwise comparison task to measure depth quality for stereoscopic photographs of natural scenes, and to compare quality between monocular and stereoscopic viewing. Stimuli consisted of four stereoscopic photographs

Individual stimuli were viewed monocularly or stereoscopically. In the monocular case, the stimulus was presented only to the participant's dominant eye. This gave a total of eight possible stimuli: each of the four photographs under monocular or stereoscopic viewing. On each trial, participants were presented with two stimuli, and judged which had the better depth quality. A
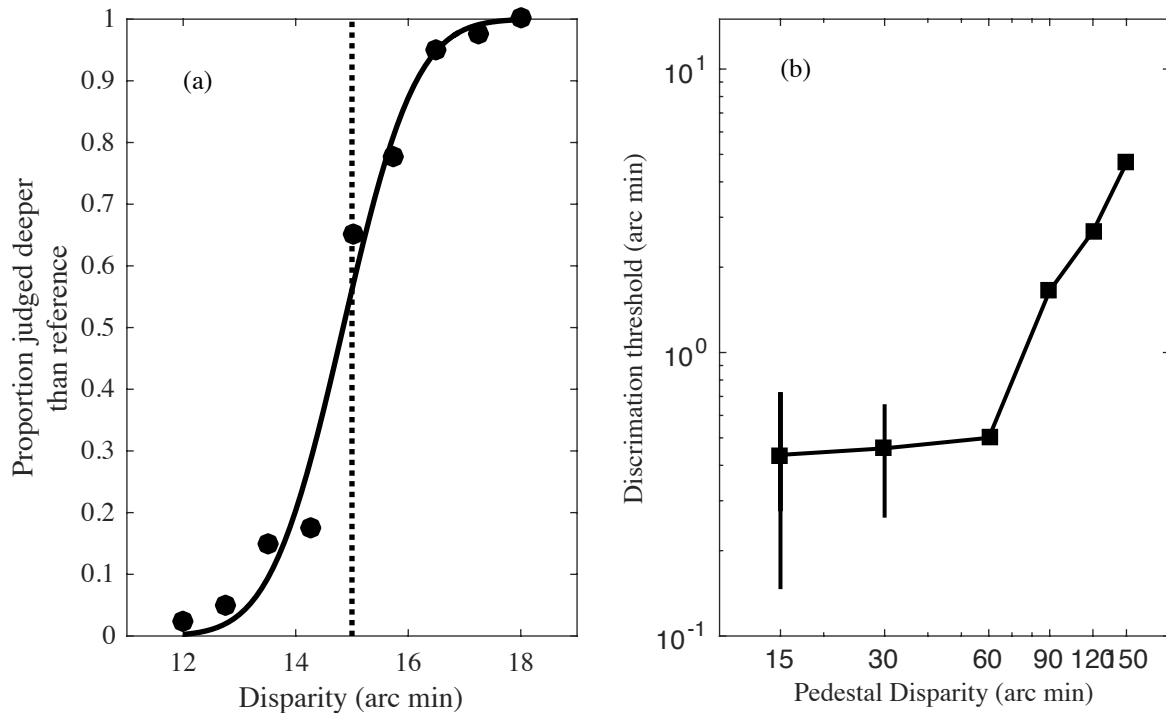
Fig 4. (a) An example of a psychometric function, showing the proportion of trials on which the test stimulus was chosen as being deeper than the reference stimulus, as a function of its disparity. The disparity of the reference stimulus is indicated by the vertical dotted line. The black circles show the participant's responses, the black curve the best fitting cumulative Gaussian. (b) Disparity discrimination thresholds plotted as a function of the disparity of the reference stimulus. Error bars show ±1 standard error

.

two-interval technique was used. One stimulus was presented in the centre of the screen for 1.5s. The screen then went blank for 1s, after which the second stimulus was presented for 1.5s. The screen again went blank, and the participant indicated whether the depth quality was greater in the first or second interval. Each of the 28 possible pairings of different stimuli was presented 10 times, across 2 blocks of trials.

3. RESULTS

3.1 Depth magnitude and quality in random-circle stereograms

Each pair of disparity values was compared 20 times, and we recorded the number of times that each disparity was judged as the larger of the two. These were then used to create a scale of apparent depth using a Thurstone scale [14]. For each comparison between two stimuli, A and B, we first calculated the proportion of times that stimulus A was judged as having more depth. The difference between the apparent depth of A and B was then calculated as the inverse of the normal cumulative error function for this proportion. If the proportions calculated were 0 or 1, they were replaced with values of 0.05 or 0.95, to allow depth scale values to be calculated [16]. The final depth magnitude for each stimulus was then calculated as the mean of the depths calculated for comparisons with all other stimuli. The same technique was used to calculate a scale of the quality of perceived depth for each stimulus.

Depth magnitude values are plotted as the black squares in Fig. 3. As expected, perceived depth increased with increasing disparity. This levelled off for disparities greater than 90 arc min [12]. The quality of depth is plotted as the red circles in Fig. 3. In contrast to depth magnitude, depth quality decreased with increasing disparity.

Depth magnitude and quality data were analysed separately using a mixed linear regression model for each task, with random intercepts and disparity as a fixed covariate. Perceived depth magnitude increased significantly with increasing disparity (slope=0.0658; t(58)=7.798; p<0.001), while depth quality decreased significantly with increasing disparity (slope=-0.0538; t(58)=-4.907; p<0.001).

3.2 Precision of depth judgments in random-circle stereograms

Each reference disparity was compared with 9 test magnitudes of disparity, and in each case the number of times that the test disparity was chosen as the larger depth interval was recorded. These responses were used to create psychometric functions; an example function in show in Fig. 4a. A cumulative Gaussian function was fit to each function, with Palamedes [9], and used to calculate a depth discrimination threshold, or just noticeable difference, in each case. This was calculated as half the difference between the disparity values at the 25% and 75% points on the psychometric function. The JND thus represents the magnitude of disparity required
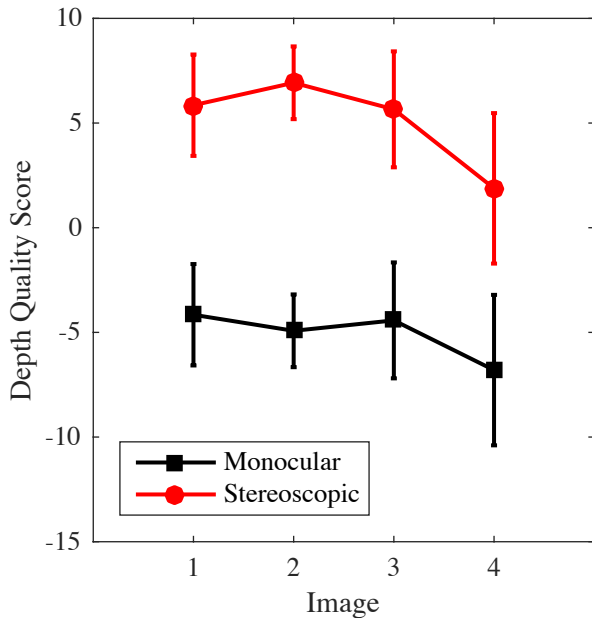

Fig 5. Depth quality values for the four photographs, under monocular (black squares) and stereoscopic (red circles) viewing. Error bars show ±1 standard error.

to discriminate the sign of disparity reliably on 75% of trials. JNDs increased as the reference disparity increased, as shown in Fig. 4b. These results were analysed with a mixed-effect linear regression, with random intercepts and disparity as a covariate. Thresholds increased significantly with increasing disparity (slope=0.0742; t(29)=3.166; p=0.004).

3.3 Quality of depth in natural stereoscopic photographs

The Thurstone scaling technique outlined in section 3.1 was used to assign a depth quality value to each of the four photographs, under both monocular and stereoscopic viewing. These data are shown in Fig 5. These data were analysed using a 4 (photograph) x 2 (viewing condition) repeated measures ANOVA. This showed a significant effect of viewing condition (F(1,4)=29.515; p=0.006), no significant effect of the photograph viewed (F(3,12)=3.330; p=0.056) and no significant interaction (F(3,12)=0.757; p=0.539). This experiment shows that the quality of depth is greater with stereoscopic viewing, as expected, and that depth quality in natural stereoscopic photographs can be reliably quantified using pairwise comparison scaling techniques.

4. DISCUSSION

We used a pairwise comparison technique to measure depth magnitude and quality in simple abstract stereograms and natural photographs. We showed that this technique can be used to quantify depth quality reliably for simple and complex stimuli. As expected, depth quality in complex scenes was better with stereoscopic viewing.

The magnitude of perceived depth increased with disparity, while the precision with which depth differences could be judged decreased. The quality of depth experience decreased with increasing disparity over the range used here. These results show that increasing apparent depth does not necessarily increase the quality of the 3D experience. Rather, as predicted by theoretical considerations [6,17,18,19], depth quality, or stereopsis, was more closely related to the precision than to the magnitude of perceived depth. Our results are consistent with the idea that the qualitative experience of stereopsis is associated with the precise representation of depth information [6,18,19].

5. REFERENCES

[1] R. Badcock, and C.Schor. Depth-increment detection function for individual spatial channels, *Journal of the Optical Society of America A*, 2:1211-1216, 1985.

[2] D.H. Brainard. The psychophysics toolbox. *Spatial Vision*, 10:433–436, 1997

[3] J.E. Cutting, and P.M. Vishton. Perceiving layout: The integration, relative dominance, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition: Vol. 5. Perception of space and motion,* pages 69–117. New York, NY: Academic Press, 1995

[4] P.C.A. Doorschot, A.M.L. Kappers, and J.J. Koenderink. The combined influence of binocular disparity and shading on pictorial shape, *Perception and Psychophysics*, 63:1038-1047, 2001.

[5] P.B. Hibbard. Binocular energy responses to natural images.*Vision Research*, 48:1427-1439, 2008.

[6] P.B. Hibbard. Can appearance be so deceptive? Representationalism and binocular vision. *Spatial Vision* 21:549-559, 2008.

[7] J.M. Hillis, S.J. Watt, M.S. Landy, and M.S. Banks. Slant from texture and disparity cues: optimal cue combination, *Journal of Vision*, 4:967-992, 2004.

[8] B.D. Keefe, P.B. Hibbard, and S.J. Watt. Depth-cue integration in grasp programming: no evidence for a binocular specialism, *Neuropsychologia*, 49:1246-1257, 2011.

[9] F.A.A. Kingdom, and A. Prins. *Psychophysics: a practical introduction*, London, Academic Press, 2009.

[10] T. Kim, J. Kang, S. Lee, and A.C. Bovik. Multimodal interactive continuous scoring of subjective 3D video quality of experience, IEEE Transactions on Multimedia, 16:387-402. 2014.

[11] M. Kleiner, D. Brainard, and D. Pelli, D. What's new in psychtoolbox-3? *Perception*, 36(ECVP Supplement):s14, 2007.

[12] K.N. Ogle. On the limits of stereoscopic vision, Journal of Experimental Psychology,44:253-259, 1952.

[13] J.C.A. Read and I. Bohr. User experience while viewing stereoscopic 3D television, *Ergonomics*, doi: 10.1080/00140139.2014.914581, 2014.

[14] L.L. Thurstone. The method of paired comparisons for social values, Journal of Abnormal and Social Psychology, 21:384-400, 1927.

[15] J. Trommershauser, K. Kording, and M.S. Landy. Sensory Cue Integration, Oxford, Oxford University Press

[16] K. Tsukida, and M.R. Gupta. How to analyze paired comparison data, *Technical Report UWEETR-2011-0004*, Seattle, WA: University of Washington, 2011.

[17] M. Tye. Visual qualia and visual content revisited, in: D.J. Chalmers (ed), *Philosophy of Mind, Classical and Contemporary Readings*, Oxford, Oxford University Press, 2002.

[18] D. Vishwanath D. Towards a new theory of stereopsis, *Psychological Review*, 121:151-178, 2014.

[19] D. Vishwanath, and P.B. Hibbard. Seeing in 3D with just one eye: Stereopsis in the absence of binocular disparities. *Psychological Science,* 24:1673–1685, 2013.

6. ACKNOWLEDGMENTS