2009

Stereoscopic depth magnitudes at greater distances in an old steam railway tunnel

Barbara Gillam
University of New South Wales

Stephen Palmisano
University of Wollongong, stephenp@uow.edu.au

Donovan Govan
University of New South Wales, dgovan@uow.edu.au

Robert Allison
York University

Julie Harris
University of St. Andrews

Stereoscopic depth magnitudes at greater distances in an old steam railway tunnel

Abstract

Abstract presented at the 32nd European Conference on Visual Perception, 24-28 August 2009, Regensburg, Germany

Keywords

stereoscopic, railway, depth, tunnel, magnitudes, greater, distances, old, steam

Disciplines

Education | Social and Behavioral Sciences

Publication Details

Gillam, B., Palmisano, S., Govan, D., Allison, R. & Harris, J. M. (2009). Stereoscopic depth magnitudes at greater distances in an old steam railway tunnel. In 32nd European Conference on Visual Perception, 24-28 August, Regensburgh, Germany. Perception, 38 (Supplement), 59-60.

disc presented at 10 deg eccentricity in a random quadrant. Task difficulty was altered by decreasing the interocular correlation, making stereo correspondence harder due to the increasing number of dots with no match in the other eye. We find that adding on-screen vertical disparities which simulate infinite viewing distance can actually increase task performance. Thus, human stereo correspondence may concentrate its search near the epipolar lines appropriate to infinity, even when this is inconsistent with the oculomotor system. However, an inconsistent vertical disparity field, where the periphery (> 20 deg) indicates the correct 30 cm viewing distance while the region < 20 deg indicates infinite viewing distance, has an even more detrimental effect. Thus, the brain may not simply search a fixed zone on the retina, but may adjust its search according to the epipolar geometry indicated by the global vertical disparity field.

◆ **Alternation frequency thresholds for stereopsis reveal different types of stereoscopic difficulties**
S I Rychkova, J Ninio¶ (Eye Microsurgery Clinic, Irkutsk, Russia; ¶ Laboratoire de Physique Statistique, Ecole Normale Supérieure, Paris, France; e-mail: rych.sv@mail.ru)

When stereoscopic images are presented alternately to the two eyes, stereopsis occurs at at least 3 Hz full cycle frequencies (eg, Ludwig et al, 2007 *Perception & Pychophysics* **69** 92 – 102). We studied the transition to stereopsis with twelve subjects, using 21 different stereograms presented on a synoptophore. For all stimuli, the subject-averaged minimal frequency F for stereopsis with stable fusion was 0.7 Hz higher than the minimal frequency for depth with apparent motion in the images. The lowest F values (4.2 to 4.6 Hz) were observed with stimuli involving 2 to 4 simple disjoint elements (circles, arcs, rectangles). Higher F values were needed for stimuli containing slanted elements (1 to 2 Hz increment), curved surfaces (2 to 4 Hz increment), overlapping elements at two different depths (1 to 1.5 Hz increment) or camouflaged overlapping surfaces (> 4 Hz). A textured cylindrical surface thus appeared easier to interpret (7.5 Hz) than a pair of slanted segments separated in depth but forming a cross in projection (8.5 Hz). Training effects were minimal, and F usually increased as disparities were reduced. The hierarchy of difficulties thus revealed may shed light on various problems that the brain needs to solve during stereoscopic interpretation.

◆ **Depth information does not speed up a texture segmentation task unless the required reaction time is longer than about 1 s: implications for the relative roles of striate and extra-striate cortices in input driven attentional guidance**
L Zhaoping, N Guyader¶, A Lewis§ (Department of Computer Science, University College London, UK; ¶ Image and Signal, Grenoble Image Parole Signal Automatique Lab, France; § Goldman Sachs International; e-mail: z.li@ucl.ac.uk)

We consider segmentation of two neighboring textures of uniformly, but differently, oriented oblique bars. The reaction time (RT) for this task becomes much longer when an irrelevant texture, comprising spatially alternating horizontal and vertical bars, is superposed on the same grid. This is a predicted signature (Zhaoping and May, 2007 *PLoS Computational Biology* **3** e62) of the bottom – up saliency mechanisms in primary visual cortex (V1). Putting the task-irrelevant texture behind the relevant texture in depth reduces interference substantially, suggesting depth cues might also contribute to attentional guidance. Since depth separation is caused by a horizontal shift between the two textures in one or both monocular images in a dichoptic display, an important question is whether a flat 2-D stimulus lacking depth cues, made by presenting the modified monocular images identically to both eyes, reduces interference or RT as effectively. We found that interference is reduced as effectively, unless the 2-D stimulus requires an RT (of subjects' report for the task) longer than about 1 s. Since physiological data suggest that extra-striate cortical areas, but not V1, are significantly involved in depth and surface perception, our observations imply that V1 dominates in input-driven attentional guidance initially after stimulus onset, and extra-striate cortices contribute only later.

◆ **Stereoscopic depth magnitudes at greater distances in an old steam railway tunnel**
B Gillam, S Palmisano¶, D Govan, R Allison§, J M Harris# (Department of Psychology, University of New South Wales, Australia; ¶ Department of Psychology, University of Wollongong, Australia; § Department of Computer science, York University, UK; # Department of Psychology, University of St Andrews, UK; e-mail: b.gillam@unsw.edu.au)

Unlike disparity thresholds, depth magnitudes (which require scaling for distance), have been little measured beyond a few metres. The present studies used remote bluetooth-controlled LED targets presented in a disused steam railway tunnel at Helensburgh, NSW, with distances of 20 m and 40 m to the nearest target and disparities up to 4.5 min of arc between targets. Conditions were either completely dark or lighted to provide perspective cues up to the nearest LED. Observers estimated depth in metres. Even in the dark, depth was scaled by disparity and

superior to monocular depth. With perspective scaling available, monocular depth did not improve but stereo depth increased strongly for most observers and was considerably greater at 40 m than at 20 m. Even the greatest depths reported were strongly attenuated from geometric prediction to about half predicted value at 20 m (similar to Allison et al, 2009) at 9 m). At 40 m the greatest depths reported were about 4 m for predicted depths of about 30 m. Clearly disparity influences perceived depth magnitude and perspective acts as a scaling cue at distances greater than 20 m. Depth compression cannot be attributed to perceived distance in either the light (reasonably accurate) or the dark.

◆ **Binocular disparities, motion parallax and geometric perspective in Patrick Hughes's 'Reverspectives'**

B J Rogers, A Gyani (Department of Experimental Psychology, University of Oxford, UK; e-mail: bjr@psy.ox.ac.uk)

Patrick Hughes's 'Reverspective' artwork, in which the perceived depth of the truncated pyramids is opposite in direction to their physical structure, have been of significant interest for vision scientists. Not only is the depth reversed in 'Reverspective' pictures but also the depth-reversed 3-D structure appears to rotate with the observer during side-to-side head movements. Papathomas [2007 *Spatial Vison* **21**(1/2) 79 – 95] argued that 'top – down' processes play an important role in the effect and he attributed the illusory rotationary motion to a "discrepancy between the expected and actual retinal optic flows". The aim of the present experiments was to determine the role of disparities, parallax and the different components of perspective information—linear-, size-, and edge-perspective—in creating the effect, using a series of reverspective models in which both the slant of the pyramid sides and the texture covering the surfaces were manipulated. The strength of the different cues under monocular and binocular viewing was derived from the distance at which the illusory effect broke down. Linear-perspective and edge-perspective effects were found to be more important than size-perspective. We also show that there is no need to invoke 'top – down' processes to explain the effect if the nature and limitations of motion parallax information are understood correctly.

◆ **Inferred motion perception of light sources in three-dimensional scenes is color-blind**

H E Gerhard, L T Maloney ¶ (Department of Psychology [¶ Center for Neural Science], New York University, USA; e-mail: holly.gerhard@nyu.edu)

An impressive function of the visual system is to estimate surface properties from retinal images that confound object properties with the light field's spatially varying chromaticity and luminance. Estimating the light field and object geometry constrains surface property estimates, and previous work indicates that light field parameters are estimated (eg, Gerhard and Maloney, 2008 *Vision Sciences*). Can the visual system also track temporal changes in the light field? If so, what information is used? For example, standard motion perception is driven by luminance while color is uninformative. We evaluate whether the same is true of inferred light source motion. Observers viewed three-dimensional surfaces illuminated by an out-of-view moving punctate source and a diffuse source. On half the trials, the punctate and diffuse sources had different chromaticities, thereby providing more information about motion direction. Of seven observers, all reported motion direction above chance (31% – 77% correct). Only the observer with the poorest performance benefited significantly from the added color information. We explored the impact of other dynamic and static image properties using a generalized linear model. We conclude that observers rely on several scene statistics to monitor temporal changes in the light field's intensity distribution, but color does not play a major role.

◆ **The image-based acuity model: A general model for image recognition**

A B Watson, A J Ahumada Jr (Human Systems Integration Division, NASA Ames Research Center, USA; e-mail: andrew.b.watson@nasa.gov)

Watson and Ahumada (2008, http://journalofvision.org/8/4/17/) proposed an image-based model for predicting acuity from optical aberrations. In this model a 'neural' image is computed from the optics and a neural contrast sensitivity function. In the ideal observer version of the model, the neural images for all letters of a given size serve as templates. When a target is presented, white 'internal' noise is added to its neural image, and the response of the observer is the nearest template. The performance of this model is only a function of the noise level and the matrix of cross-correlations among the neural images. The noise level we estimated from the observers in Cheng et al (2004, http://journalofvision.org/4/11/1/) ranged from 0.5×10^4 to 10^5 deg^{-2}, consistent with the noise level of the Modelfest observers (2005, http://journalofvision.org/5/9/6/). Watson and Ahumada (2008) presented simple acuity metrics that assumed no correlation among templates. Here we derived and tested metrics that include the cross-correlations among templates. Predictions were computed for blurred and unblurred tumbling E images and Sloan letters. The new metrics