Published in final edited form as:

Brain Res. 2011 March 4; 1377: 67–77. doi:10.1016/j.brainres.2011.01.003.

The integration of disparity, shading and motion parallax cues for depth perception in humans and monkeys

Peter H. Schiller, Warren M. Slocum, Brian Jao, and Veronica S. Weiner

77 Massachusetts Ave., Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

Abstract

A visual stimulus display was created that enabled us to examine how effectively the three depth cues of disparity, motion parallax and shading can be integrated in humans and monkeys. The display was designed to allow us to present these three depth cues separately and in various combinations. Depth was processed most effectively and most rapidly when all three cues were presented together indicating that these separate cues are integrated at yet unknown sites in the brain. Testing in humans and monkeys yielded similar results suggesting that monkeys are a good animal model for the study of the underlying neural mechanisms of depth perception.

Keywords

Shading; motion parallax; stereopsis

1. Introduction

One of the primary tasks of the visual system is to derive the third dimension in the visual scene from the two-dimensional images that fall on the retinal surface. Several depth cues have evolved to accomplish this. Three of these are stereopsis, motion parallax and shading. In previous studies it has been established that these three are effective individual cues for depth in humans as well as in monkeys (Rogers and Graham 1979; Rogers and Graham 1982; Poggio and Poggio 1984; Ramachandran 1988; DeAngelis et al 1991; Sekuler and Blake 1994; Harris et al 1997; DeAngelis et al 1998; Freeman 1999; Janssen et al 2000; Kontsevich and Tyler 2000; Cumming and DeAngelis 2001; Parker and Cumming 2001; Grunewald et al 2002; Howard 2002; Howard and Rogers 2002; DeAngelis and Uka 2003; Born and Bradley 2005; Schiller and Carvey 2006; Roe et al 2007; Schiller et al 2007; Nadler et al 2008; Zhang and Schiller 2008). Studies have also shown additive interactions among various depth cues (Bulthoff and Mallot 1988; Johnston et al 1993; Young et al 1993; Bradshaw et al 1998; Schiller et al 2007).

In recent studies we have compared interactions between pairs of depth cues in both humans and monkeys. We have examined stereopsis and motion parallax as well as shading and stereopsis interactions. We have established that these cues are integrated leading to more

Corresponding author: Peter H. Schiller, 77 Massachusetts Ave., 46-6041, Cambridge, MA 02139, USA, 1-617-253-5754 (tel), 1-617-324-3725 (fax), phschill@mit.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

accurate and more rapid analysis of depth information when they are presented conjointly as compared with them being presented singly (Schiller et al 2007; Zhang, et al 2007; Zhang and Schiller 2008). In the present study we have examined interactions among all three of these depth cues, namely disparity, motion parallax and shading. To accomplish this, a random-dot display system was devised with which we could present these three cues singly and in various combinations. Our results show that in both humans and monkeys these three cues are effectively integrated.

2. Results

Data were collected from two human and two monkey subjects for the seven presentation conditions. For each of these conditions three depth-levels were used for the distractors. The depth of the target stimulus, as defined by the three cues, **d** (disparity), **p** (parallax), **s** (shading) was kept constant. The three cues were presented either singly, in pairs or with all three of them together. The distractors appeared at three different depths, making thereby for a total of 21 stimulus conditions. Both percent correct and reaction time data were collected. The total number of trials collected per subject was as follows: Human subject P = 4,380; human subject V = 2,100; Monkey M = 9,840; Monkey K = 9,840.

In the figures we show data in three different ways. First, shown in in Figure 2 and 3, we plot performance on each of the three single-depth cues (**d**, **p** & **s**) and the triple cue condition (**d+p+s**), with Figure 2 showing the human data and Figure 3 showing the monkey data. Second, shown in Figures 4 and 5, we compare performance on all seven depth conditions used in the experiment under the smallest target/distractor difference condition: single depth cues (disparity, parallax and shading), paired depth cues (disparity plus parallax, parallax plus shading and disparity plus shading). Third, the conditions were aggregated by the number of cues (yielding results for single double and triple cues conditions), for the smallest target/distractor difference conditions as shown in Figure 6.

The three levels of difficulty for each condition were set individually for each subject. The most difficult condition was set to yield roughly similar levels of percent correct performance for each of the three single cues (**d**, **p** & **s**). As can be seen, this was realized quite well for the human subjects (Figure 2) as well as for the monkeys (Figure 3). The data in these two figures show that percent correct performance was significantly enhanced when the three cues were presented in combination. Also evident is the fact that the latencies were much shorter when all three cues were provided compared with the latencies obtained for single cue conditions.

Figures 4 and 5 show data for the single cue, paired cue and triple cue conditions for the smallest target/distractor differences.

Table 2 shows a two-way analysis of variance (ANOVA) of percent correct performance for two human subjects (P and V) and two monkeys (M and K) for the data displayed in Figures 2 and 3. The two factors are type of depth cue (**d**, **p** & **s**) and cue level (i.e. levels of disparity, relative velocity, color contrast). All four cases revealed significant main effects and interactions of depth cue and cue level.

Table 3 shows a two-way analysis of variance (ANOVA) of saccadic latency for two human subjects (P and V) and two monkeys (M and K) for the data displayed in Figure 2 and 3. The two factors are type of depth cue (**d**, **p** & **s**) and cue level (i.e. levels of disparity, relative velocity, color contrast). All four cases revealed significant main effects and interactions of depth cue and cue level.

Tables 4 and 5 show t-tests carried out on the data displayed in Figures 4 and 5. In Table 4 t-test values are shown for the comparison of percent correct performance for one versus three depth cues. The depth cues were disparity alone (**d**), parallax alone (**p**), shading alone (**s**), or all three in combination (**d+p+s**). Cue levels correspond to the depth points on the x-axis of Figures 2 and 3 (i.e. a disparity of 3.4 minutes). Bonferroni adjustments for multiple comparisons were used to determine the significance levels (*p*-values). In general, the triple cue condition (**d+p+s**) yielded significantly better percent correct performance than any individual cue alone for the lower two levels of the depth cues; however, performance tended to saturate near 100% at the highest level of the depth cues. In Table 5 comparisons of saccadic latencies for one versus three depth cues are shown. For the various conditions the depth cue was either disparity alone (**d**), parallax alone (**p**), shading alone (**s**), or all three cues in combination (**d+p+s**). Cue levels correspond to the depth points on the x-axis of Figures 2 and 3 (i.e. a disparity of 3.4 minutes). Bonferroni adjustments for multiple comparisons were used to determine the significance levels (*p*-values). In general, the triple cue condition (**d+p+s**) yielded significantly lower saccadic latencies than any individual cue alone.

To compare percent correct performance and latencies for the single, double and triple cue conditions, we combined the disparity, parallax and shading cues for each of the two human and the monkey subjects. Figure 6 shows these data for the smallest target/distractor difference. In each display the left-most bars (1) show the combined single cue conditions, the middle bar the combined double cue conditions (2) and the right-most bar the triple condition. The statistical analyses of these data appear in Table 6 showing that the data were significant beyond the 0.01 level for 11 of the 12 conditions percent correct and for 10 of the 12 conditions for latency.

3. Discussion

The findings we report here establish that all three of the depth cues studied, disparity, shading and parallax, are effectively utilized by monkeys as they are by humans. The three cues are integrated in both humans and monkeys, leading to significantly more accurate and faster processing of depth information compared to conditions when these cues are presented singly. When two cues were provided performance was significantly faster and better than with single cues but not as effective as when all three cues were provided.

To prove that the shading cues are utilized for depth in monkeys, in a previous study we compared performance when the cues were presented in accordance with the principles of shading, based on illumination from above, with the same cues presented in a manner that was not in consonance with illumination from above (Zhang et al 2007). That study showed that performance of the monkeys was much poorer under the latter condition, indicating that they indeed use shading cues to process depth cues as do humans.

The central questions addressed by research examining the neural underpinnings of depth perception is where in the brain these various depth cues are processed, where they are integrated and what the specific neural mechanisms are that can accomplish these tasks. Numerous studies have established that stereoscopic depth perception arises already in area V1 (Pettigrew 1978; Poggio and Poggio 1984; Freeman 1999; Cumming and DeAngelis 2001). Several extrastriate areas also contribute to the processing of stereopsis, including areas V4 and MT (DeAngelis et al 1991; DeAngelis et al 1998; DeAngelis and Uka 2003). Three-dimensional shape coding has also been demonstrated in inferior temporal cortex (Janssen et al 2000; Yamane et al 2008). Taira et al (2000) had shown that neurons in the caudal intraparietal sulcus respond to orientation in random dot stereograms and solid figure stereograms, thereby responding to disparity cues. It does not appear, however, that any

single specific area is uniquely involved in the processing of stereopsis, as the removal of such regions as area V4 or MT produce only limited deficits in stereopsis; even paired V4 and MT lesions fail to eliminate depth processing (Schiller 1993; Chowdhury et al 2008). Such distributed processing is also the case for motion parallax; direction selective cells, especially cells that are selective for differential motion, a central requirement for processing motion parallax, have been found already in area V1 and are prevalent in areas MT and MST (Mikami et al 1986a, b; Duffy and Wurtz 1991; Roy et al 1992; DeAngelis et al 1998; Born and Bradley 2005; Nadler et al 2008). Also, Xiao et al (1997) had shown that MT/V5 cells respond to tilt of orientation of depth in motion depending on asymmetrical surround and speed tuning. Selective lesions of single extrastriate areas fail to eliminate depth processing by motion parallax suggesting that depth processing based on motion parallax is also distributed in several extrastriate regions (Schiller 1993). Imaging studies in humans and monkeys have revealed that both the dorsal and ventral regions of the cortex are involved in the extraction of depth from texture, while information based on depth from shading is carried out predominantly in the ventral stream (Georgieva et al 2008; Nelissen et al, 2009).

Examining the question of where various depth cues become integrated, Liu, Vogels and Orban (2004) have shown that there is a convergence of depth from texture and disparity in the macaque inferior temporal cortex. Much less is known at this stage about where all three of the depth cues we report on in this study -- disparity, motion parallax and shading -- are integrated in the brain. Does such integration take place already in area V1, between two of these cues or among all three of them, or does this integration take place in higher areas? If the latter, one would like to determine which area it is and to what extent the integration is unique to that area. To answer these questions the response properties of single neurons need to be examined when stimulated by these cues individually and in various combinations; additionally, the effects of inactivation of selected areas needs to be carried out to determine to what extent cue integration can be disrupted.

4. Experimental Procedure

1. Subjects

The experiment was carried out on two male rhesus macaques (M and K) and two human subjects (P and V). The monkeys had a head post and a scleral search coil implanted to stabilize the head during the experimental sessions and to measure eye movements (Schiller 1993; Cao and Schiller 2002; Cao and Schiller 2003; Schiller et al 2007; Zhang et al 2007; Zhang and Schiller 2008). All surgical procedures were carried out in accordance with the NIH approved guidelines of the Department of Comparative Medicine at MIT as laid down in the publication *Guide for the Care and Use of Laboratory Animals* (National Institutes of Health publication No. 86-23, revised 1985). MIT's Committee on the Use of Humans as Experimental Subjects gave IRB approval for all research involving human subjects. The experiments were undertaken with the understanding and written consent of each human subject, and the study conformed with The Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the *British Medical Journal* (18 July 1964).

2. The display system

A display system was devised that enabled us to study the three depth cues of disparity (**d**), motion parallax (**p**) and shading (**s**) separately, and in various combinations. The system, which mimicked a rocking three-dimensional object, was presented on a color monitor and was viewed through a stereoscope. Figure 1 provides a demonstration of this arrangement. An oddity task was used in which four truncated pyramids were presented, made visible by virtue of disparity, parallax and shading, displayed singly or in various combinations. The

four protruding pyramids appeared at 3-degree eccentricities relative to the central fixation spot, with each placed at 45 degrees above and below the horizontal meridian to the left and the right. Three of the truncated pyramids, the distractors, were identical. The fourth pyramid, the target, had a different depth as defined by disparity, motion parallax and shading with these cues being presented singly and in various combinations. The depth difference between the target and the three distractors was systematically varied allowing us to obtain psychometric functions for both percent correct performance and reaction times.

Each of the dot-filled squares viewed through the stereoscope, as depicted on the top right of Figure 1, was comprised of 350 by 350 pixels and measured 9.8 by 9.8 degrees of visual angle (1 pixel = 1.68 minutes of visual angle). The four pyramids were centered at a three-degree eccentricity appearing at 45, 135, 225 and 315 degrees as shown on the left of Figure 1. The base of each pyramid extended 72 by 72 pixels and the top 24 by 24 pixels thereby making the base 2 by 2 degrees and the top 0.67 by 0.67 degrees. The random dots were 3 by 3, 4 by 4 or 5 by 5 pixels in size (5.04×5.04 , 6.72×6.72 and 8.4×8.4 minutes).

To be rewarded, the target, which appeared randomly at one of the four locations, had to be chosen by the monkey as described below. Seven presentation conditions were used which were presented in blocks of ten trials per condition: (1) disparity alone (**d**), (2) motion parallax alone (**p**), (3) shading alone (**s**), (4) **d + p**, (5) **d + s**, (6) **p + s** and (7) **d + p + s**.

The set of disparity values used was the same for all subjects for the target and the distractors. The motion parallax values for the two humans were identical. For the monkeys these values were set to be different from the humans in accordance with the calculations we had made to assure correspondence between disparity and parallax as described in previous publications (Zhang et al 2007; Zhang and Schiller 2008). The luminance values for the distractors differed among subjects over a small range as specified in Table 1. The values were adjusted to yield similar performance at the lowest contrast differences to those obtained for parallax and disparity. This was accomplished by running subjects repeatedly while making subtle adjustments in contrast values of the distractors to yield similar percent correct performance for the shading cues as those obtained for parallax and disparity. These adjustments yielded quite similar contrast values for the human subjects P, V and for Monkey M, but ended up having slightly higher values for Monkey K because this animal exhibited somewhat less sensitivity for the shading cues. For the lowest target/distractor contrasts (see #1 in Table 1 for the luminances employed) the contrast difference between the top panel background and the back panel background (see Figure 1C) was 11.2 % for subjects P, V and M; for subject K the value was 18.4 %.

The target for each of the conditions within blocks of trials was always constant, while the distractors were presented randomly at different depths, most commonly using three steps. The individual depth cues provided with disparity, motion parallax and shading for the distractors were provided separately for each subject to yield similar percent correct performance levels for each of the three lowest depth cue levels, as can be seen in Figure 2 and 3.

The three disparity and parallax differences between the target and the distractors are specified on the abscissa in Figures 2 and 3. Shown are the disparity differences between the target and distractors in minutes and the differential rocking velocities in degrees per second. The corresponding shading values used appear in Table 1 for the target and the distractors.

To create a display which mimics real three dimensional objects consisting of four rocking pyramids, the level of disparity and parallax changed in value from the tip of the pyramid, where they were maximal, to the base, where they were minimal, reaching the same value as

the vertical surface of the display. The values in Table 1 specify the maximal disparity and parallax values that were used at the tip of the truncated pyramid.

The values for the shading cues are more complex. To provide appropriate shading cues, different luminance values were set for the top, side and bottom of each face in the pyramid. Furthermore, these values, as required by the rules of reflectance, were set differently for the background (bkgnd) and the dots for each surface. These values are shown in Table 1 separately for the target and the distractors. Again, the values for the target remained constant for blocks of trials while the values for the three distractors were varied randomly allowing us thereby to generate psychometric functions.

Data were collected for both percent correct performance and for latencies. Monkeys were trained to make saccadic eye movements to the target and humans to press buttons as described below.

3. Testing procedures

a. Monkeys—During the experimental sessions each animal was seated in a monkey chair and faced a color monitor placed at a distance of 57.3 cm from the eye whereby a 1 cm extent on the screen equals 1 degree of visual angle. After the animal had been placed into the testing apparatus with the head fixed, the stereoscope, through which the visual display was viewed, was placed in position and a tube for the delivery of drops of apple juice was put into the animal's mouth. Each trial began with the appearance of a fixation spot in the center of a frame. The background was homogeneous, with the luminance set to 7.9 cd/m². After the monkey fixated the fixation spot for 180 to 220 ms, the stimulus display appeared as depicted in Figure 1. Target selection by the monkey was accomplished by making a saccadic eye movement to the target; the stimuli were extinguished right after a selection had been made. A reward, consisting of a drop of apple juice, was dispensed only when the animal made a direct saccadic eye movement to the target. The data collected were reaction times and percent correct performance.

b. Humans—The human subjects viewed a similar display as did the monkey using a stereoscope designed for human use with the monitor screen also placed at a distance of 57.3 cm. A chin rest was used to facilitate comfortable viewing. For humans, the selection process involved the pressing of one of four buttons on a panel which was placed horizontally on the table below the stereoscope. Subjects used the left-hand and right-hand thumbs and forefingers for the pressing. They were instructed to rest their fingers on the buttons and to press the appropriate button with dispatch, and to make their choice while maintaining central fixation. Feedback for correct choices was provided with a brief beep sound.

Throughout, four truncated pyramids were presented. As with the monkeys, seven presentation conditions were used presented in blocks: (1) disparity alone (**d**), (2) motion parallax alone (**p**), (3) shading alone (**s**), (4) **d+p**, (5) **d+s**, (6) **p+s** and (7) **d+p+s**. Each set consisted of 10 blocks making for 120 trials when three conditions were used and 160 trials when four levels were used. The intertrial interval was set at 1200 ms. In each block of trials the target values of disparity, and/or motion parallax, and/or shading was kept constant while three different values were used for the three identical distractors allowing us thereby to generate performance functions for percent correct and latencies as a function of the difficulty level of the task.

4. Statistical analyses

Statistical analyses were performed to assess the effect of the display parameters on saccadic latency and percent correct performance. Mean saccadic latencies and percent correct were calculated and presented with 95% confidence intervals. Two-way analysis of variance (ANOVA) test were used for each experiment. When an ANOVA test revealed significant effects or interactions of the experimental conditions, planned comparisons were performed among conditions of interest using Student *t*-tests with a Bonferroni adjustment for multiple comparisons.

All analyses were performed using MATLAB software (The MathWorks, Inc., Natick, Massachusetts, USA). Data from each human subject or monkey were analyzed independently. These statistical analyses follow methods employed by Schiller et al (2007).

Acknowledgments

The research reported here was supported by grant EY014884 from the National Eye Institute. The authors thank Christina E. Carvey and Michelle C. Kwak for their assistance in putting together this manuscript.

Abbreviations

V1	striate cortex
V4	extrastriate visual cortical area V4
MT	middle temporal cortex
MST	medial superior temporal cortex

Literature References

- Born RT, Bradley DC. Structure and function of visual area MT. *Annu Rev Neurosci* 2005;28:157–189. [PubMed: 16022593]
- Bradshaw MF, Parton AD, Eagle RA. The interaction of binocular disparity and motion parallax in determining perceived depth and perceived size. *Perception* 1998;27:1317–1331. [PubMed: 10505177]
- Bulthoff HH, Mallot HA. Integration of depth modules: stereo and shading. *J Opt Soc Am A* 1988;5:1749–1758. [PubMed: 3204438]
- Cao A, Schiller PH. Behavioral assessment of motion parallax and stereopsis as depth cues in rhesus monkeys. *Vision Res* 2002;42:1953–1961. [PubMed: 12160568]
- Cao A, Schiller PH. Neural responses to relative speed in the primary visual cortex of rhesus monkey. *Vis Neurosci* 2003;20:77–84. [PubMed: 12699085]
- Chowdhury SA, DeAngelis GC. Fine discrimination training alters the causal contribution of macaque area MT to depth perception. *Neuron* 2008;60:367–377. [PubMed: 18957227]
- Cumming BG, DeAngelis GC. The physiology of stereopsis. *Annu Rev Neurosci* 2001;24:203–238. [PubMed: 11283310]
- DeAngelis GC, Cumming BG, Newsome WT. Cortical area MT and the perception of stereoscopic depth. *Nature* 1998;394:677–680. [PubMed: 9716130]
- DeAngelis GC, Ohzawa I, Freeman RD. Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature* 1991;352:156–159. [PubMed: 2067576]
- DeAngelis GC, Uka T. Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 2003;89:1094–1111. [PubMed: 12574483]
- Duffy CJ, Wurtz RH. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J Neurophysiol* 1991;65:1329–1345. [PubMed: 1875243]
- Freeman RD. Stereoscopic vision: Which parts of the brain are involved? *Curr Biol* 1999;9:R610–613. [PubMed: 10469583]

- Georgieva SS, Todd JT, Peeters R, Orban GA. The extraction of 3D shape from texture and shading in the human brain. *Cereb Cortex* 2008;18:2416–2438. [PubMed: 18281304]
- Grunewald A, Bradley DC, Andersen RA. Neural correlates of structure-from-motion perception in macaque V1 and MT. *J Neurosci* 2002;22:6195–6207. [PubMed: 12122078]
- Harris JM, McKee SP, Smallman HS. Fine-scale processing in human binocular stereopsis. *J Opt Soc Am A Opt Image Sci Vis* 1997;14:1673–1683. [PubMed: 9248059]
- Howard, IP. Seeing in depth. Vol. I. Porteous; Toronto: 2002.
- Howard, IP.; Rogers, BJ. Seeing in depth. Vol. II. Porteous; Toronto: 2002.
- Janssen P, Vogels R, Orban GA. Three-dimensional shape coding in inferior temporal cortex. *Neuron* 2000;27:385–397. [PubMed: 10985357]
- Johnston EB, Cumming BG, Parker AJ. Integration of depth modules: stereopsis and texture. *Vision Res* 1993;33:813–826. [PubMed: 8351852]
- Kontsevich LL, Tyler CW. Relative contributions of sustained and transient pathways to human stereoprocessing. *Vision Res* 2000;40:3245–3255. [PubMed: 11008141]
- Liu Y, Vogels R, Orban GA. Convergence of depth from texture and depth from disparity in macaque inferior temporal cortex. *J Neurosci* 2004;24:3795–3800. [PubMed: 15084660]
- Mikami A, Newsome WT, Wurtz RH. Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *J Neurophysiol* 1986;55:1308–1327. [PubMed: 3016210]
- Mikami A, Newsome WT, Wurtz RH. Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *J Neurophysiol* 1986;55:1328–1339. [PubMed: 3734858]
- Nadler JW, Angelaki DE, DeAngelis GC. A neural representation of depth from motion parallax in macaque visual cortex. *Nature* 2008;452:642–645. [PubMed: 18344979]
- Nelissen K, Joly O, Durand JB, Todd JT, Vanduffel W, Orban GA. The extraction of depth structure from shading and texture in the macaque brain. *PLoS One* 2009;4:e8306. [PubMed: 20011540]
- Parker AJ, Cumming BG. Cortical mechanisms of binocular stereoscopic vision. *Prog Brain Res* 2001;134:205–216. [PubMed: 11702545]
- Pettigrew J. Stereoscopic visual processing. *Nature* 1978;273:9–11. [PubMed: 692676]
- Poggio GF, Poggio T. The analysis of stereopsis. *Annu Rev Neurosci* 1984;7:379–412. [PubMed: 6370081]
- Ramachandran VS. Perceiving shape from shading. *Sci Am* 1988;259:76–83. [PubMed: 3064296]
- Roe AW, Parker AJ, Born RT, DeAngelis GC. Disparity channels in early vision. *J Neurosci* 2007;27:11820–11831. [PubMed: 17978018]
- Rogers B, Graham M. Motion parallax as an independent cue for depth perception. *Perception* 1979;8:125–134. [PubMed: 471676]
- Rogers B, Graham M. Similarities between motion parallax and stereopsis in human depth perception. *Vision Res* 1982;22:261–270. [PubMed: 7101762]
- Roy JP, Komatsu H, Wurtz RH. Disparity sensitivity of neurons in monkey extrastriate area MST. *J Neurosci* 1992;12:2478–2492. [PubMed: 1613542]
- Schiller PH. The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis Neurosci* 1993;10:717–746. [PubMed: 8338809]
- Schiller PH, Carvey CE. Demonstrations of spatiotemporal integration and what they tell us about the visual system. *Perception* 2006;35:1521–1555. [PubMed: 17286122]
- Schiller PH, Slocum WM, Weiner VS. How the parallel channels of the retina contribute to depth processing. *Eur J Neurosci* 2007;26:1307–1321. [PubMed: 17767508]
- Sekuler, R.; Blake, R. Perception. McGraw-Hill; New York: 1994.
- Taira M, Tsutsui KI, Jiang M, Yara K, Sakata H. Parietal neurons represent surface orientation from the gradient of binocular disparity. *J Neurophysiol* 2000;83:3140–3146. [PubMed: 10805708]
- Xiao DK, Marcar VL, Raiguel SE, Orban GA. Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. *Eur J Neurosci* 1997;9:956–964. [PubMed: 9182948]

- Yamane Y, Carlson ET, Bowman KC, Wang Z, Connor CE. A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat Neurosci* 2008;11:1352–1360. [PubMed: 18836443]
- Young MJ, Landy MS, Maloney LT. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res* 1993;33:2685–2696. [PubMed: 8296465]
- Zhang Y, Schiller PH. The effect of overall stimulus velocity on motion parallax. *Vis Neurosci* 2008;25:3–15. [PubMed: 18282306]
- Zhang Y, Weiner VS, Slocum WM, Schiller PH. Depth from shading and disparity in humans and monkeys. *Vis Neurosci* 2007;24:207–215. [PubMed: 17640412]

Figure 1. Truncated pyramid display providing disparity, parallax and shading cues separately or in combination

The basic layout of the display system. A: A three-dimensional rendition of the display, shown with shading cues; four truncated pyramids are visible, one of which, the target, appears to protrude more than the other three, the distractors. B: Head-on view of the left and right eye random images viewed through a stereoscope (D). The two displays are fused and appear as a single unit which rocks back and forth along a central vertical axis. C: The head-on appearance of one of the truncated pyramids made visible by virtue of shading cues. D: The stereoscope used in these experiments.

Figure 2. Separate and combined disparity, parallax and shading depth cues, humans
 Saccadic latencies (panels A & C) and percent correct performance (panels B & D) for two human subjects P (panels A & B) and V (panels C & D). Data are shown for three separate depth cues – disparity (**d**), parallax (**p**), and shading (**s**) – as well as for a combination of the three cues (**d+p+s**). Units for the abscissae are degrees of visual angle (disparity) and degrees per second of differential velocity (parallax). The luminance values for shading vary by subject and are shown in Table 1. The number of trials (N) in each data set is indicated.

Figure 3. Separate and combined disparity, parallax and shading depth cues, monkeys Saccadic latencies (panels A & C) and percent correct performance (panels B & D) for two monkeys M (A & B) and K (C & D). Data are shown for three separate depth cues – disparity (**d**), parallax (**p**), and shading (**s**) – as well as for a combination of the three cues (**d+p+s**). Units for the abscissae are degrees of visual angle (disparity) and degrees per second of differential velocity (parallax). The luminance values used for each subject for shading are shown in Table 1. The number of trials (N) in each data set is indicated.

Figure 4. Disparity, parallax and shading cues presented singly and in various combinations under the smallest target/distractor difference, humans

Mean saccadic latencies (panels A & C) and percent correct performance (panels B & D) for two human subjects P (panels A & B) and V (panels C & D) under the smallest target-distractor difference employed (i.e. 3.4 minutes). Error bars indicate the 95% confidence interval for the estimate of the mean. Results are shown for three separate depth cues – disparity (d), parallax (p), and shading (s) – as well as for each possible pair of cues (d+p, p+s, d+s) and for the combination of all three cues (d+p+s). The number of trials (N) in each data set is indicated.

Figure 5. Disparity, parallax and shading cues presented singly and in various combinations under the smallest target/distractor difference, monkeys

Mean saccadic latencies (A & C) and percent correct performance (B & D) for two monkeys M (A & C) and K (C & D) under the smallest target-distractor difference employed (i.e. 3.4 minutes). Error bars indicate the 95% confidence interval for the estimate of the mean.

Results are shown for three separate depth cues – disparity (d), parallax (p), and shading (s) – as well as for each possible pair of cues (d+p, p+s, d+s) and for the combination of all three cues (d+p+s). The number of trials (N) in each data set is indicated.

Figure 6. Mean latencies and percent correct for single cues, paired cues, and all three cues
Mean saccadic latencies (top row) and percent correct performance (bottom row) for two human subjects (P & V) and two monkeys (M & K). Data have been aggregated by number of cues: one cue (**d**, **p**, **s**), two cues (**d+p**, **p+s**, **d+s**), and three cues (**d+p+s**). Error bars indicate the 95% confidence intervals for the estimate of the mean. Data are shown for the smallest target-distractor difference employed (i.e. 3.4 minutes).

Table 1**Shading values**

Shading values (cd/m^2) are shown for the dots and background used to define the base, target, and distractor regions of the visual display. Shading levels are defined separately for the top (t), bottom (b), and sides (s) of the pyramidal images. See Figure 1 for details. Data are shown for two human subjects (P and V) and for two monkeys (M and K). Distractor shading values were adjusted for each experimental subject.

Shading values in cd/m^2									
Base luminance					Target luminances				
		dots	bkgnd				dots	bkgnd	
		31.4	7.9		t		51.1	27.5	
					b		19.6	2.4	
					s		39.3	15.7	
Distractor luminances									
		1		2		3			
		dots	bkgnd	dots	bkgnd	dots	bkgnd		
P	t	47.1	19.6	49.1	23.6	49.9	25.1		
	b	23.6	3.5	21.6	2.8	20.4	2.4		
	s	35.4	11.8	37.3	13.7	37.7	14.5		
V	t	47.1	19.6	49.1	23.6	49.9	25.9		
	b	23.6	3.5	21.6	2.8	20.0	2.4		
	s	35.4	11.8	37.3	13.7	38.1	15.3		
M	t	47.1	19.6	49.1	23.6	49.9	25.1		
	b	23.6	3.5	21.6	2.8	20.4	2.4		
	s	35.4	11.8	37.3	13.7	37.7	14.5		
K	t	45.2	15.7	47.1	19.6	49.1	23.6		
	b	27.5	4.7	23.6	3.5	21.6	2.8		
	s	33.4	9.8	35.4	11.8	37.3	13.7		

Table 2

Analysis of percent correct – ANOVA

Two-way analysis of variance (ANOVA) of percent correct performance for two human subjects (P and V) and two monkeys (M and K). The two factors are type of depth cue (d, p, s) and cue level (i.e. levels of disparity, relative velocity, color contrast). All four cases revealed significant main effects and interactions of depth cue and cue level.

Analysis of percent correct performance, ANOVA test									
Variable	Subject P				Subject V				
	F-statistic	P-value	d.f.	N trials	F-statistic	P-value	d.f.	N trials	
Depth cue	148.4	p << 0.01	6	4380	63.8	p << 0.01	6	1920	
Cue level	1413	p << 0.01	2	4380	987.5	p << 0.01	2	1920	
Interaction	37.7	p << 0.01	12	4380	26.0	p << 0.01	12	1920	

Analysis of percent correct performance, ANOVA test									
Variable	Monkey M				Monkey K				
	F-statistic	P-value	d.f.	N trials	F-statistic	P-value	d.f.	N trials	
Depth cue	542.7	p << 0.01	6	9840	375.2	p << 0.01	6	9360	
Cue level	4451	p << 0.01	2	9840	3749	p << 0.01	2	9360	
Interaction	180.3	p << 0.01	12	9840	22.8	p << 0.01	12	9360	

Table 3

Analysis of saccadic latency – ANOVA

Two-way analysis of variance (ANOVA) of saccadic latency for two human subjects (P and V) and two monkeys (M and K). The two factors are type of depth cue (d, p, s) and cue level (i.e. levels of disparity, relative velocity, color contrast). All four cases revealed significant main effects and interactions of depth cue and cue level.

Analysis of saccadic latency, ANOVA test									
Variable	Subject P				Subject V				
	F-statistic	P-value	d.f.	N trials	F-statistic	P-value	d.f.	N trials	
Depth cue	125.5	p << 0.01	6	3676	40.2	p << 0.01	6	1590	
Cue level	529.2	p << 0.01	2	3676	190.4	p << 0.01	2	1590	
Interaction	15.5	p << 0.01	12	3676	5.2	p << 0.01	12	1590	

Analysis of saccadic latency, ANOVA test									
Variable	Monkey M				Monkey K				
	F-statistic	P-value	d.f.	N trials	F-statistic	P-value	d.f.	N trials	
Depth cue	744.0	p << 0.01	6	8214	496.6	p << 0.01	6	6078	
Cue level	2497.6	p << 0.01	2	8214	192.2	p << 0.01	2	6078	
Interaction	109.1	p << 0.01	12	8214	16.0	p << 0.01	12	6078	

Table 4
Analysis of percent correct performance – *posthoc* comparisons

Comparison of percent correct performance for one versus three depth cues. Depth cue was either disparity alone (d), parallax alone (p), shading alone (s), or all three in combination (d+p+s). Cue levels correspond to the lowest depth point on the x-axis of Figures 2 and 3 (i.e. a disparity of 3.4 minutes). Bonferroni adjustments for multiple comparisons were used to determine the significance levels (*p*-values). In general, the triple cue condition (d+p+s) yielded significantly better percent correct performance than any individual cue alone for the lower two levels of the depth cues; however, performance tended to saturate near 100% at the highest level of the depth cues.

Comparison of percent correct for different depth conditions						
Cue level (disparity in minutes of visual angle, and corresponding values for parallax and shading)	Subject P			Subject V		
	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s
3.4	p < 0.01, t = 6.37	p < 0.01, t = 5.60	p < 0.01, t = 7.85	p < 0.01, t = 8.51	p < 0.01, t = 8.11	p < 0.01, t = 6.95
6.7	p < 0.01, t = 3.18	p < 0.01, t = 5.63	p < 0.01, t = 9.18	p < 0.01, t = 3.78	p < 0.01, t = 3.60	p < 0.01, t = 3.40
13.4	not significant t = 1.02	not significant t = 0.00	p < 0.01, t = 3.57	not significant t = 0.99	not significant t = 1.43	not significant t = 2.05

Comparison of percent correct for different depth conditions						
Cue level (disparity in minutes of visual angle, and corresponding values for parallax and shading)	Monkey M			Monkey K		
	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s
3.4	p < 0.01, t = 16.2	p < 0.01, t = 18.3	p < 0.01, t = 15.7	p < 0.01, t = 8.30	p < 0.01, t = 4.93	p < 0.01, t = 3.32
6.7	p < 0.01, t = 8.30	p < 0.01, t = 15.7	p < 0.01, t = 4.26	p < 0.01, t = 12.0	p < 0.01, t = 10.1	p < 0.01, t = 10.9
13.4	not significant t = 0.22	p < 0.01, t = 4.49	not significant t = 0.57	p < 0.01, t = 13.0	p < 0.01, t = 10.4	p < 0.01, t = 8.59

Table 5
Analysis of saccadic latency performance – *posthoc* comparisons

Comparison of saccadic latency for one versus three depth cues. Depth cue was either disparity alone (d), parallax alone (p), shading alone (s), or all three cues in combination (d+p+s). Cue levels correspond to the lowest depth point on the x-axis of Figures 2 and 3 (i.e. a disparity of 3.4 minutes). Bonferroni adjustments for multiple comparisons were used to determine the significance levels (*p*-values). In general, the triple cue condition (d+p+s) yielded significantly lower saccadic latencies than any individual cue alone.

Comparison of saccadic latency for different depth conditions						
Cue level (disparity in minutes of visual angle, and corresponding values for parallax and shading)	Subject P			Subject V		
	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s	d+p+s d vs.	d+p+s p vs.	d+p+s s vs.
3.4	p < 0.01, t = 4.43	p < 0.01, t = 6.01	p < 0.05, t = 2.41	not significant t = 0.49	p < 0.05, t = 2.66	p < 0.05, t = 2.63
6.7	p < 0.01, t = 19.9	p < 0.01, t = 19.2	p < 0.01, t = 18.7	p < 0.01, t = 5.13	p < 0.01, t = 10.3	p < 0.01, t = 9.28
13.4	p < 0.01, t = 18.0	p < 0.01, t = 28.4	p < 0.01, t = 21.4	p < 0.01, t = 9.42	p < 0.01, t = 19.0	p < 0.01, t = 10.1

Comparison of saccadic latency for different depth conditions						
Cue level (disparity in minutes of visual angle, and corresponding values for parallax and shading)	Monkey M			Monkey K		
	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s	d+p+s vs. d	d+p+s vs. p	d+p+s vs. s
3.4	p < 0.01, t = 8.54	p < 0.01, t = 13.1	p < 0.01, t = 12.7	p < 0.01, t = 5.87	p < 0.01, t = 13.1	p < 0.01, t = 4.43
6.7	p < 0.01, t = 19.2	p < 0.01, t = 35.0	p < 0.01, t = 26.3	p < 0.01, t = 19.9	p < 0.01, t = 15.8	p < 0.01, t = 14.0
13.4	p < 0.01, t = 22.4	p < 0.01, t = 48.2	p < 0.01, t = 37.3	p < 0.01, t = 33.7	p < 0.01, t = 34.0	p < 0.01, t = 22.2

Table 6
Comparison of saccadic latencies and percent correct for varying number of cues

Comparison of saccadic latency (top panel) and percent correct performance (bottom panel) for varying numbers of depth cues presented. Data are shown for two human subjects (P & V) and two monkeys (M & K). Data have been aggregated by number of cues: one cue (**d, p, s**), two cues (**d+p, p+s, d+s**), and three cues (**d+p+s**). Data are shown for the smallest target-distractor difference employed (i.e. 3.4 minutes). The table presents p-values indicating the significance level of a t-test for the difference of means.

Saccadic latency			
Subjects	Two cues vs. one cue	Three cues vs. one cue	Three cues vs. two cues
Subjects P	p = 0.06	p << 0.01	p << 0.01
Subjects V	p = 0.07	p << 0.01	p = 0.017
Monkeys M	p << 0.01	p << 0.01	p << 0.01
Monkeys K	p << 0.01	p << 0.01	p << 0.01

Percent correct			
Subjects	Two cues vs. one cue	Three cues vs. one cue	Three cues vs. two cues
Subjects P	p << 0.01	p << 0.01	p = 0.11
Subjects V	p << 0.01	p << 0.01	p << 0.01
Monkeys M	p << 0.01	p << 0.01	p << 0.01
Monkeys K	p << 0.01	p << 0.01	p << 0.01