The swinging doors of perception: Stereomotion without binocular matching

Kevin R. Brooks   School of Psychology, UNSW, Sydney, Australia

Barbara J. Gillam   School of Psychology, UNSW, Sydney, Australia

Until recently, it was considered necessary for features in the two eyes to be matched before the evaluation of differences in their locations (binocular disparities) could reveal depth information. Motion in depth can also be perceived binocularly from related changes in the locations of matched binocular features. However, unmatched features can arise when a binocular object occludes more distant features in one eye but not the other. The presence and extent of such features can provide quantitative depth information, although perceived depth relative to geometrical predictions may vary from one such arrangement to another. The ability of humans to perceive motion in depth from unmatched stimuli has not previously been explored. Here, we use B. Gillam, S. Blackburn, and K. Nakayama's (1999) "monocular gap" stimuli to investigate perception of motion in depth simulated by a change in the extent of a monocularly occluded feature in a binocular display. Settings of a motion in depth probe revealed that the magnitude of perceived motion in depth is generally as large as that for a stimulus containing matchable binocular features. We show that our stimuli provide disambiguating information not present in similar static stimuli. We conclude that in the computation of motion in depth, a binocular match is not required. A new cue—dynamic half-occlusion—can be used to reach an accurate percept.

Keywords: da Vinci stereopsis, ecological optics, half-occlusion, monocular camouflage, monocular gap, stereopsis

Introduction

It is well known by now that monocular features in binocular displays can provide ecologically valid information about depth relations and may produce quantitative depth perception. They simulate conditions in which a surface occludes the background for one eye and not the other (Gillam & Borsting, 1988; Nakayama & Shimojo, 1990). The present investigation extends these findings to explore the perception of motion in depth resulting from changes over time in the extent of a monocular feature in a binocular context.

Rashbass and Westheimer (1961) were the first to identify two separate yet linked cues to binocular motion in depth or "stereomotion." These are (a) the changing relative binocular disparity over time (referred to henceforth as changing disparity, or CD), and (b) the extent to which the directions and/or speeds of the object's two monocular half-images differ (the interocular velocity difference cue, or IOVD). The speed of motion in depth that each cue signals is shown in Equation 1 below, where D represents viewing distance, I is the interocular separation, $\omega_{\mathbf{R}}$ and $\omega_{\mathbf{L}}$ are the right and left image velocities, respectively, and δ represents disparity.

$$\mathbf{v_z} \approx \frac{\mathbf{d}\delta}{\mathbf{d}t}\frac{D^2}{I} = (\omega_{\mathbf{R}}-\omega_{\mathbf{L}})\frac{D^2}{I} \qquad (1)$$

In all but the most artificial of situations (e.g., dynamic random dot stereogram stimuli), these two cues are con-comitant. Although they arise from the same binocular geometry and carry identical motion in depth information, they may be processed by entirely separate mechanisms. For example, the CD cue must first combine and compare binocularly matching features to yield disparities before computing a motion signal from the variations in disparity over time. In contrast, the IOVD cue involves the computation of two independent monocular motion signals before binocular combination and comparison allow recovery of the 3D motion. Laboratory studies that isolate or selectively manipulate these cues have revealed that both play a part in stereomotion perception (Brooks, 2001; Brooks, 2002a, 2002b; Brooks & Mather, 2000; Brooks & Stone, 2004; Cumming & Parker, 1994; Fernandez & Farrell, 2005; Harris & Watamaniuk, 1995; Shioiri, Saisho, & Yaguchi, 2000).

In common with the mechanisms of conventional stereopsis, each of these established stereomotion cues relies on the comparison of features in the two monocular images and, hence, requires that matches be made. The CD cue requires the matching of several features in the binocular image to compute a rate of change of relative disparity over time (Collewijn, Erkelens, & Regan, 1985; Erkelens & Collewijn, 1985; Regan, Erkelens, & Collewijn, 1986), whereas the IOVD cue must derive two separate monocular motion signals from a matched binocular feature (Brooks & Stone, 2004; Cumming & Parker, 1994; Harris & Watamaniuk, 1995; Rashbass & Westheimer, 1961). It is worth pointing out at this stage that although the horizontal monocular velocities that serve as IOVD inputs can have any value including zero (for a stationary feature), it must

have a value. For a feature purely visible to the left eye, for example, no IOVD can exist since no feature means no value of ω_R in Equation 1. Clearly, the percept derived from a binocular stimulus that has motion in the left eye and none in the right (i.e., perceived motion in depth toward and away from the right eye) is quite different from that derived from similar motion in a purely monocular stimulus (lateral motion only).

There has been some debate over the exact nature of the primitives of binocular matching. The outputs of linear filters, such as luminance-defined edges or "zero crossings" (Marr & Poggio, 1979), or local luminance extrema (e.g., Mayhew & Frisbee, 1981) are often the primary candidates for matching in models of stereoscopic vision. However, a variety of other features, such as monocular edges defined by motion (Lee, 1970), flicker (Prazdny, 1984), color (Ramachandran, Rao, & Vidyasagar, 1973), or texture (Ramachandran et al., 1973), or, more recently, second-order features defined by contrast (e.g., Wilcox, 1999; Wilcox & Hess, 1996; Ziegler & Hess, 1999) have been shown to produce a percept of depth, although this percept is often only directional rather than quantitative in nature (Ziegler & Hess, 1999). There has even been an informal report of a percept of motion in depth from a change in disparity and associated IOVD in such images (Prazdny, 1984), although data have never been formally presented. Depth and motion in depth percepts from these second-order stereo stimuli still involve the matching of visible features in the two monocular images, despite the fact that these may not be defined by luminance variations. We refer to depth perception involving the matching of any such primitives (whether they are first or second order) as conventional stereopsis.

However, in natural 3D scenes, matching features are not always present. When objects are at different depths, it is not uncommon for features on the more distant object to be fully visible in one eye yet completely occluded in the other. It is also possible for features of proximal objects to be camouflaged in one eye alone against more distant objects of identical luminance. For such unmatched stimuli, the conventional stereomotion cues of CD and IOVD will also be unavailable.

Binocular depth perception in the presence of unmatched features has been extensively studied in recent years under static conditions. Depth has been found to be quantitative to various degrees in that it varies with the extent of the half-occluded or half-camouflaged feature (Brooks & Gillam, 2006; Cook & Gillam, 2004; Gillam, Blackburn, & Nakayama, 1999; Gillam & Nakayama, 1999; Häkkinen & Nyman, 1997; Malik, Anderson, & Charowhas, 1999; Pianta & Gillam, 2003a, 2003b; Tsai & Victor, 2000).

In this investigation, we used dynamic versions of the "monocular gap stereogram" (Gillam et al., 1999), as shown in Figure 1, in an attempt to produce a percept of motion in depth from changes in the locations of unmatched features. Cross-fusing the images of Figure 1a produces a percept of two equal sized frontoparallel
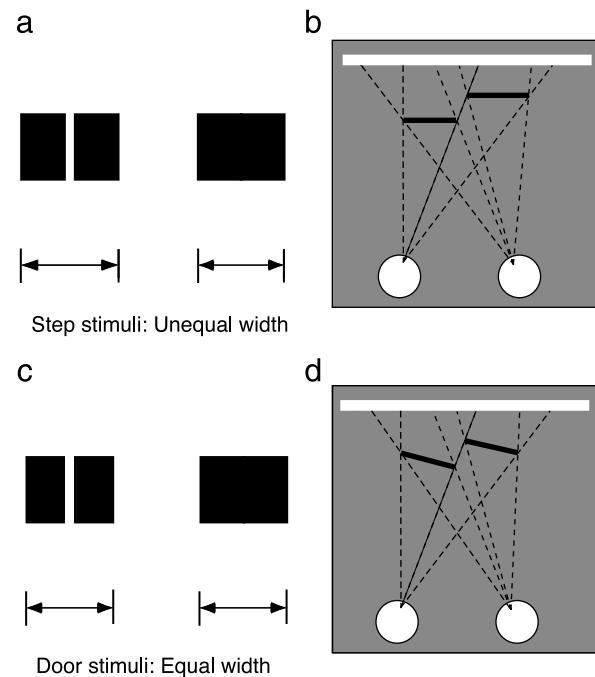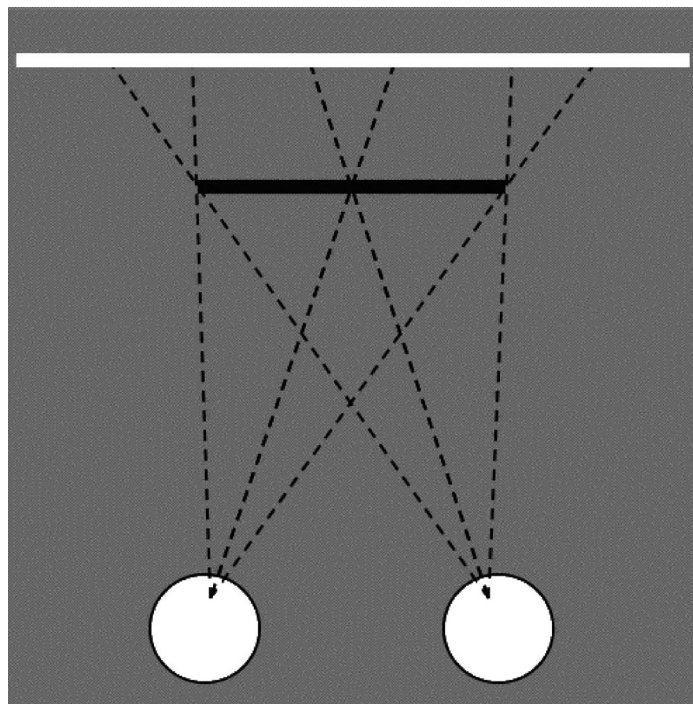

Figure 1. Snapshots of stimuli. (a) Monocular half-images of step stimuli. Crossed fusion reveals two frontoparallel planes, with the left closer than the right. Uncrossed fusion results in a reversed depth order. (b) Binocular geometry for step stimuli, showing the usual depth percept. (c) Monocular half-images of door stimuli. Crossed fusion reveals two slanted planes, with the left closer than the right. (d) Binocular geometry for door stimuli.

planes in depth (the left plane appearing closer than the right), with a central depth discontinuity between them (see Figure 1b). This we refer to as the Step stimulus. The presence of a monocular gap in one eye ecologically entails a depth discontinuity of a given sign within the figure. Although the depth of the outer edges is explicitly signaled by a conventional disparity, the inner edges of the stimulus carry no such signal. Instead, binocular geometry and the constraints of ecological validity specify a range of possible depths and surface slants (see Pianta and Gillam, 2003b, for a full discussion of the specific constraints applied). Observers reliably perceive the depth corresponding to the minimum depth constraint (marked by the bold black lines) for this stimulus (Gillam et al., 1999; Pianta & Gillam, 2003b) as well as some other varieties of unmatched stereogram (Cook & Gillam, 2004). The stereogram shown in Figure 1c is identical to Figure 1a except that the total width of the half-images is equal, such that the outer edges produce no disparity. Still, a depth discontinuity is seen at the gap, whereas the outer edges appear to be at the same depth as each other. The percept is of two approximately parallel panels slanted in depth (see Figure 1d). This we refer to as the Door stimulus. Perceived depth at the gap for this stimulus type has been shown to often be smaller than that predicted by the minimum depth constraint (Pianta & Gillam, 2003b).
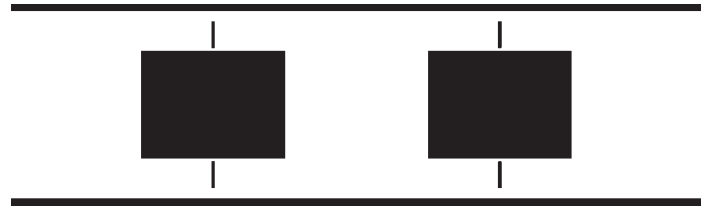
Movie 1. Representation of monocular half-images for the step stimulus. Appropriate for crossed or uncrossed fusion. Click on the image to view the movie.

By continuously changing the size of the gap in one half-image, we were able to create a percept of motion in depth for both Step and Door stereograms. In the case of the Step stimulus, the locations of the outer edges of the same half-image were manipulated in concert with gap change, preserving the size of each panel (see Movie 1) and producing a strong sense of two frontoparallel planes, one moving toward and one moving away from the observer (Movie 2). In the Door stimulus, only the gap changed (see Movie 3), giving rise to a percept of two doors swinging in opposite directions about their fixed outer edges (Movie 4). While the gap size increased and decreased in one half-image, the other image featured an unchanging black rectangle without a gap. When the gap had closed in the right eye, a gap began to open and then close in the left eye in an identical fashion, whereas the right half-image remained motionless. This cycle repeated continu-
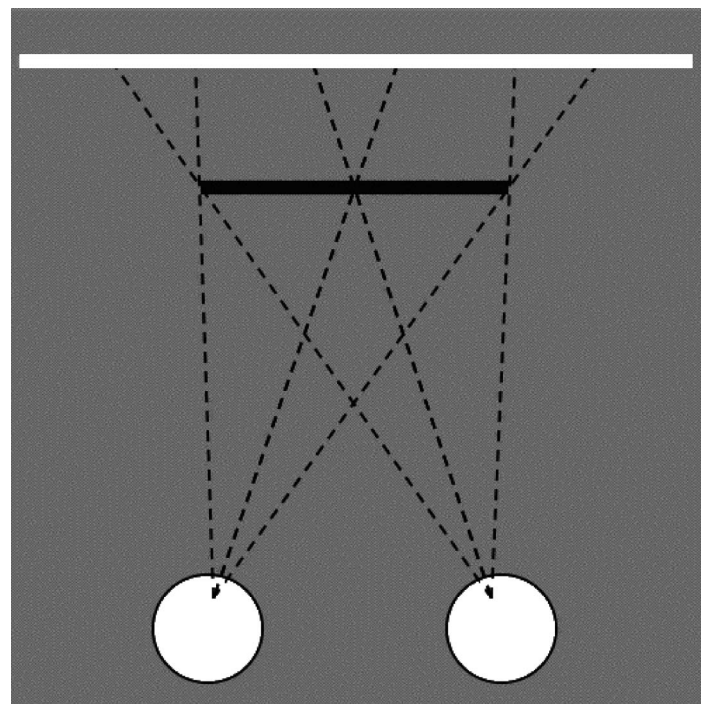

Movie 3. Representation of monocular half-images for the door stimulus. Appropriate for crossed or uncrossed fusion. Click on the image to view the movie.

ously. Since a gap was never simultaneously visible in both monocular half-images, this condition is referred to as Unmatched for both Step and Door stimuli.

Experiment 1

To quantify these percepts, we adopted a method of adjustment paradigm. Observers set the amplitude of motion of a probe featuring conventional stereomotion cues to equal the amplitude of motion in depth seen at the inner edges of the 3D figure. In addition to the Unmatched stimulus, matches were made to similar stimuli featuring conventional cues to stereomotion. This Matched condition featured a binocular gap with changing edge disparities

Movie 2. Changing binocular geometry for the step stimulus and the observed perceptual resolution. Click on the image to view the movie.

Movie 4. Changing binocular geometry for the door stimulus and the observed perceptual resolution. Click on the image to view the movie.

and associated IOVDs. A Synoptic condition provided identical stimuli (the Unmatched stimulus' half-image showing a gap) simultaneously to both the left and right eyes. This condition controlled for the possibility that any motion in depth effects were somehow due to lateral motion signals alone and established a baseline. Finally, two monocular control conditions were included. A Left-Eye-Only condition was used wherein the left eye resembled the Unmatched stimulus, whereas the right eye saw only an unchanging uniform background. With the Alternate stimulus, the stationary large black rectangle was always omitted, leaving a changing gap stimulus that appeared in the left eye, then in the right, and so forth. These conditions assessed the possibility that an IOVD system could produce a motion in depth percept when only one eye features a motion signal with no suggestion of half-occlusion. The motion signals presented to either eye are identical to those in the Unmatched condition without the binocular context within which the monocular feature indicates half-occlusion.

Since Step stimuli involve a changing outer edge disparity, one could argue that stereomotion signals are propagated inward from these edges. These signals may be applied at the unmatched feature without that feature directly contributing to the perception of depth or motion in depth. If the perceived motion in depth in Step stimuli is entirely due to outer edge signals propagating inward and the monocular gap does not impose a motion in depth signal, the changing depth signals at the outer edges alone must be large enough to predict the observed results. This possibility was addressed by asking observers to match the outer edges of the Unmatched Step stimulus and of an equivalent stimulus that lacked a gap (i.e., a dynamic slant stimulus) but contained identical stereomotion signals at its outer edges to the Unmatched Step stimulus.

Methods

Stereoscopic stimuli were presented on two Samsung SynchMaster 957DF CRT monitors driven by an ATI Radeon 8500 dual-head video board and synchronized at a rate of 60 Hz. The gamma non-linearity of each monitor was corrected using the lookup table. Images were superimposed using a modified Wheatstone stereoscope with convergence distance adjusted to match the optical distance of 86 cm, while maintaining perpendicular lines of sight to the screens. At this distance, each screen subtended 24.3 × 19 deg. Subpixel resolution was achieved by antialiasing edge positions to 1/60th of a 0.62-min-wide pixel. Stimuli were black (0.5 cd/m^2), presented on a white (95 cd/m^2) background. Black vertical lines (1.9 × 14.9 min) were placed above and below the center of each monocular stimulus as a fusion lock and stereoscopic reference. Probe stimuli consisted of one 81.24 × 60 min monocular rectangle in each half-image, one featuring a central vertical gap of 1.24 min width and the other featuring a gap whose width changed linearly, frame by frame,

with an amplitude controlled by the observer. Probe gap amplitude could vary up to 13.64 min to give a maximum gap difference (and hence disparity) of 12.4 min. This produced matchable binocular features, and hence, conventional stereomotion cues were present. The observer was able to manipulate the amplitude of probe motion and the relative phase (0 or 180 deg) of probe to test.

Five conditions were included for each stimulus type (Steps and Doors). These were the Unmatched, Matched, Synoptic, Alternate, and Left-Eye-Only conditions, as described above.

The motion cycle repeated at a rate of 0.25 Hz for all stimuli and conditions. Test stimuli (80 × 60 min) were presented simultaneously 5 deg above the probe and were randomized in their horizontal position by up to ±12.4 min. Each had a maximum gap size of 9.3 min and a minimum of 0 min, except for the Matched stimuli whose maximum gap size was 10.54 min with a minimum of 1.24 min. All stimuli featured a one-frame deletion, timed to coincide with the end of one half-image's deletion and the other's creation, to mask the transient signal in the Alternate condition as the figure disappeared in one eye and reappeared in the other. For the Synoptic, Left-Eye-Only, and Alternate stimuli, the monocular gap reopened as soon as it had closed. This ensured that, as in the Unmatched and Matched stimuli, the observer always perceived motion in the combined, cyclopean display, rather than viewing a stimulus that remained motionless for half of its period (as in the Unmatched stimulus' monocular images).

Four observers were asked to match the amplitude of motion in depth seen at either the inner or outer edges of the probe stimulus to that seen in the test stimulus. This amplitude was controlled using the keyboard cursor keys. If observers made a further reduction of amplitude when the gap width was at a minimum, monocular gaps then began to increase in anti-phase with the test stimulus. Initial amplitude of the probe was randomized with a range of 9.3 min and could be randomly in-phase or antiphase. After practice sessions, all observers contributed nine settings for each task and condition. The order of stimulus conditions was randomized within each block of testing. Step and Door data were collected separately, as were inner edge and outer edge data. No feedback was given. Although two subjects had knowledge of the details of binocular depth perception, none were aware of the conditions and hypotheses being tested.

One observer's data contained one outlier that was more than 8 standard deviations from the mean of responses (observer N.S., Door stimuli, Unmatched condition). This value, believed to represent a finger error, was removed. Statistical analyses were performed by means of one-way analyses of variance (ANOVAs) for each observer. Thirty-two specific planned comparisons between unmatched data and all other conditions were performed using linear contrasts (four for each observer for each stimulus). Critical values were adjusted to maintain a per-experiment α level of .05. For significant results, we report only the smallest F value across observers.

Results

Results for perceived motion in depth of the inner edges of the Step stimuli are shown in Figure 2 (first column). All observers set large amplitudes of motion in depth in the probe to equal the perceived motion in depth in the Unmatched conditions. Settings for Matched displays were equal to those for Unmatched for three subjects, at near veridical levels, whereas one subject made a slightly higher setting for Matched than Unmatched stimuli. No consistent motion in depth was seen for control conditions. A significant effect of condition emerged for all observers, $F(4,32) \geq 9.71$, $p < .0001$. Planned contrasts showed significant differences for comparisons of Unmatched data with Synoptic, Alternate, and Left-Eye-Only conditions for each observer, $F(1,40) \geq 33.973$, $p < .0001$. Only observer B.S. showed a significant difference between Matched and Unmatched data, $F(1,40) = 57.562$, $p < .0001$.

Observers saw little or no change in depth for the No-Gap outer edges, making settings near zero, whereas motion in depth for the Unmatched condition was perceived clearly at the outer edge as at the inner edge (see Figure 2, second column). A significant difference between the two conditions was shown for all observers, $F(1,16) \geq 114.17$, $p < .0001$. Our results for the Unmatched Step stereograms cannot be explained in terms of conventional stereomotion cues at the outer edges.
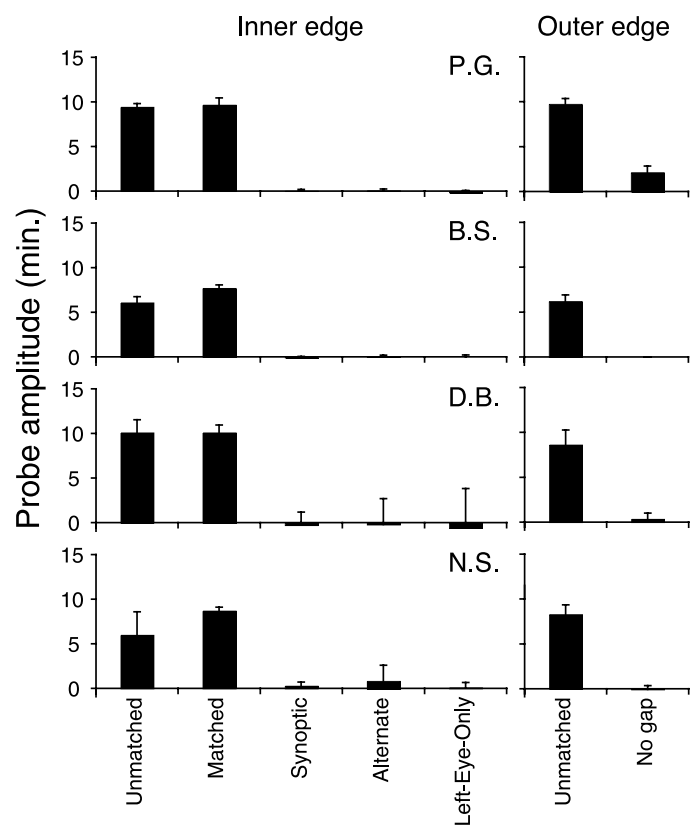
Motion in depth matches for the inner edges of the Door stimuli are shown in Figure 3. Here, a pattern similar to that for the Step stimuli emerges, where observers make equal magnitude settings for Unmatched and Matched conditions, whereas control displays lacked any coherent percept of motion in depth. A significant effect of condition was present for all subjects, $F(4,32) \geq 10.47$, $p < .0001$. Planned contrasts confirmed the statistical significance of differences between Unmatched data and Synoptic, Alternate, or Left-Eye-Only data for all observers, $F(1,40) \geq 11.683$, $p < .05$. No comparisons of Unmatched with Matched data were significant. Although means in the two monocular conditions (Alternate, Left-Eye-Only) for Door stimuli all fall near zero, the larger error bars shown for two observers (D.B. and N.S.) are noteworthy. These observers' data showed a bimodal pattern, such that on many trials, a non-zero setting was made, reflecting a percept of motion in depth. However, these settings were as likely to be out of phase with the test stimulus as they were in-phase. We consider that these stimuli are

Figure 2. Results for step stimuli. (a) Inner edges. (b) Outer edges. Error bars represent 95% confidence intervals.
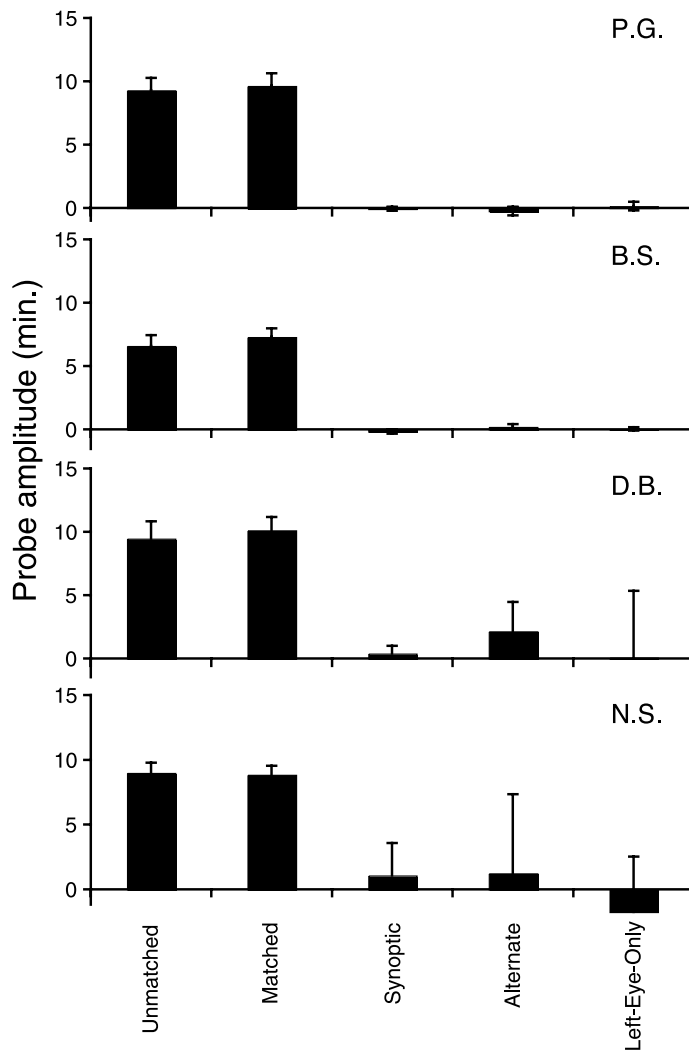

Figure 3. Results for door stimuli. Error bars represent 95% confidence intervals.

capable of producing a percept of motion in depth simply due to progressive foreshortening over time for some observers. This percept is necessarily ambiguous and quite different to the percept created by either Matched or Unmatched stimuli, whose phase was never confused by any observer.

Discussion

The data show that a change in the degree of half-occlusion provides a clear and unambiguous percept of motion in depth. This cue was effective in the absence of binocular matching (and hence CD or IOVD information) and was effectively set in conflict with looming information since the constant height of the stimuli was not consistent with the isotropic expansion expected during motion in depth.

It is clear that a continuous change of the central monocular gap in a binocular object acts as a motion in depth signal in its own right. If the changing monocular gap simply signaled a discontinuity to which the CD or IOVD signal at the outer edges of the Step stimulus could be propagated, the stereomotion signal at the outer edges would have to be at least as large as that seen at the inner edges. The No-Gap stimulus presents stereomotion signals at the outer edges identical to those in the Unmatched stereogram with no depth signal in the middle. When asked to match the motion in depth of the outer edges, observers saw little or no change in depth for the No-Gap stimuli, making near-zero settings. Our results for the Unmatched stereograms thus cannot be explained in terms of the stereomotion cues at the outer edges. In addition, Door stimuli contain explicit signals that their outer edges are stationary, yet show a robust percept of motion in depth with equivalent accuracy in Matched and Unmatched conditions. This also shows that the motion in depth seen is not dependent upon the conventional stereomotion cues at the stimulus' outer edges.

Meanwhile, motion in depth for the Unmatched condition was perceived as clearly at the outer edge as at the inner edge (see Figure 2, Column 2). The signal given by the monocular gap that a depth step is present greatly increases the stereoscopic response to the disparity at the outer edges. In this case, rather than the disparity at the outer edges propagating inward to the discontinuity, information that there is a discontinuity propagates outward to the edge disparity. This is a novel finding that has not been shown in a static context.

The possibility of artifactual matching

Although the inner edges of the two visible rectangles in one monocular image of our stimuli have no obvious corresponding feature in the other eye, we should at least entertain the possibility that these edges could be matched with other visible features in the other eye. Such possibil-

ities and the predictions they make for the Door stimuli are explored in Figure 4. (Equivalent arguments can be applied to the Step stimuli.)

The most obvious of these possibilities is that the inner edges (E_RH_R and F_RG_R) of the two rectangles in one eye are Matched over a long range to the single rectangle's outer edges in the other image (B_LC_L and A_LD_L, respectively), with edges of the same contrast polarity being Matched (see Figure 4a). This effectively represents the maximum depth constraint for this stimulus and resembles Panum's limiting case (Panum, 1858), where two objects in one eye must be Matched with a single object in the other. Such matches would predict a very large relative depth between the two inner edges (89.3 min) and would be expected to yield very large amplitude of motion in depth (see Figure 4b). Furthermore, when the gap disappeared in one eye and appeared in the other, this would correspond to a momentary discontinuity in the CD or IOVD signal, where each edge instantaneously goes from a large positive to a large negative depth or vice versa. Neither this pattern of results nor these percepts of large depths, slants, and motion discontinuities were observed.

It has also been suggested that the matching of horizontal lines with differing positions or of differing lengths can give rise to a stereoscopic percept of depth (Gillam, 1995; Gillam & Nakayama, 1999; Grove, Brooks, Anderson, & Gillam, 2006). For example, length A_LB_L might be double matched with extents A_RE_R and F_RB_R, with similar matches between the rectangles' bottom edges. Any depth signal or change of depth signal that might emerge from the matching of such features would produce the same predictions as the long-range, same-polarity matches mentioned above and, hence, are unable to explain our effect.

Stereoscopic matches involving stimuli with opposite contrast polarity have been demonstrated, although their effectiveness is somewhat reduced compared with same contrast polarity matches (Cogan, Kontsevich, Lomakin, Halpern, & Blake, 1995; Cogan, Lomakin, & Rossi, 1993; Cumming & Parker, 1997; Pope, Edwards, & Schor, 1999). If such matches were occurring here, between E_RH_R and A_LD_L and between F_RG_R and B_LC_L from Figure 4a, CD and IOVD cues might be able to operate. In this scheme, it is not clear to which of the panels (the left or the right) the inner edge reversed-polarity disparity signals would be applied (i.e., does the left panel stretch from A_LA_R to A_LE_R or to B_LF_R?), but the possible solutions produce a situation where the gap is either as wide as the entire stimulus or nonexistent. Considering the motion in depth signals alone, such matches make explicit predictions as to the pattern of results and the percept derived, as shown in Figure 4c. The relative depths of the inner edges would again be very large (70.7 min). Although there would be a degree of motion in depth between them, the sign of the depth discontinuity between the two panels would predict that the observers would set the phase of motion in depth opposite to that shown in their data. This motion sequence would also result in a "jump" in the middle when the gap switches from one
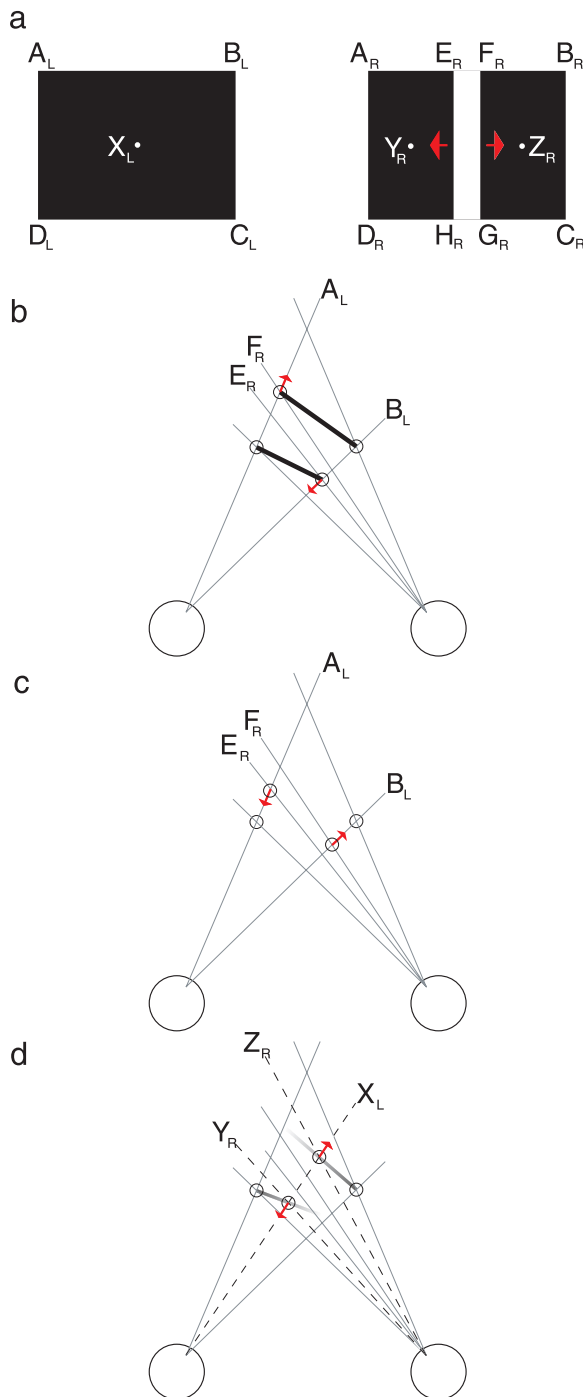
Figure 4. Possible matching solutions. (a) A schematic of the door stimulus during expansion phase (red arrows). Edges in the left and right eye's half-images are labeled with subscripts L and R, respectively. The centroids of contiguous areas are marked as X_L, Y_R, and Z_R. It is assumed that edges $A_R D_R$ and $B_R C_R$ will match $A_L D_L$ and $B_L C_L$. (b) Panum's limiting case: the maximum depth constraint. Plan view of matches of $E_R H_R$ with $B_L C_L$ and of $F_R G_R$ with $A_L D_L$. Matches are encircled in black. (c) Opposite contrast polarity match. Matches of edges $E_R H_R$ with $A_L D_L$ and of $F_R G_R$ with $B_L C_L$, repectively. Planes are not drawn, as this model does not specify their details. (d) Centroid match. Matches of right eye centroids Y_R and Z_R with X_L. This model does not specify depth signals beyond the centroids.

monocular image to the next. Again, this does not correspond to the percept reported or to the results shown by our observers.

In addition to the matching of second-order edges and of contrast maxima mentioned earlier, Mitchell (1969, 1970) has reported that qualitative stereopsis is possible for images presented briefly at a large disparity even if they are completely dissimilar in shape (e.g., a horizontal line and a vertical line; an "o" and an "x"). Such stimuli are also capable of initiating a vergence response. It is unlikely that the matches made in determining the depth of such stimuli would involve the images' edges since they are so dissimilar and cannot be fused. Rather, this coarse stereo mechanism would be more likely to match the centroids of the images in question, despite the fact that they are featureless. If a coarse stereo mechanism were at work here, there is the possibility that the center of each of the two panels in one eye (Y_R and Z_R in Figure 4a) might be matched with the center of the large rectangle (X_L) in the other eye, as shown in Figure 4d. This would again predict a percept of motion in depth amplitude far higher (44.65 min) than that observed in our data, as well as a discontinuity as the gap disappeared in one eye and reappeared in the other. Although the coarse stereopsis system identified by Mitchell can elicit some impression of depth order, this does not seem sufficient to explain a motion in depth percept such as ours. Furthermore, although a vergence response might be initiated, it has been shown that vergence alone is not sufficient to create a motion in depth percept (Collewijn et al., 1985; Erkelens & Collewijn, 1985; Regan et al., 1986).

For any matching solution to make predictions corresponding to our observed pattern of results, it would have to involve a match between the inner edges of the two panels with an inferred edge in the center of the single panel in the other eye. This would effectively be a solution corresponding to the minimum depth constraint. This is a solution that is reached for some static stimuli, although in some circumstances, a larger depth is seen when the two panels appear to overlap rather than abut (see Pianta & Gillam, 2003b). However, as there is no edge (either first order, second order, or even subjective or illusory) at this central location with which our inner edges could be matched, this cannot be considered within the bounds of any known stereoscopic matching process.

Since our simple stimuli entail only textureless black figures on a white background, there do not seem to be any other features available as primitives for a matching process. As such, no current stereo models would appear able to generate predictions consistent with the motion in depth perceived here. Instead, considerations of the geometry of binocular vision, occlusion, and camouflage may be informative.

The constraints of ecological optics

It is striking that even the Door stimuli, which entirely lack any continuous change in binocular disparity or IOVD, produce a motion in depth setting very close to that

predicted by the minimum depth constraint. Analogous static stimuli have been found to produce significantly attenuated depth relative to this constraint (Pianta & Gillam, 2003b), perhaps reflecting a particular difficulty in resolving ambiguous surface slants when no frontal plane solution is possible. Although the depth in static Door stimuli is constrained, it is not uniquely specified by binocular geometry in the way that a binocularly matchable stimulus is, leaving a degree of ambiguity. The present findings suggest that our motion sequence removes any such ambiguity by providing crucial additional stimulus information, leaving the changing depth fully constrained in geometric terms.

The ranges of possible slants seen in static images are bounded by minimum and maximum depth constraints (see Figure 5). These are constructed using the intersection of the rectangles' visible inner edges in one image and either the rectangle's outer edges (maximum) or the assumed position of abutting edges (minimum) in the other eye. While complete overlap corresponds to the maximum, a lack of overlap (abutting edges) corresponds to the minimum depth. Although in static images the degree of overlap is unknown, in the motion in depth versions used here the surfaces swing past each other, and so cannot overlap if they are to remain rigid and solid. Instead, they must abut, which means that only the minimum depth constraint is relevant and that depths (and associated slants) cannot increase up to the maximum depth constraint that previously applied to static images. In addition, the static image lacks any information on the relative sizes of the two component surfaces, leading to many possible minimum depth constraints (see Figures 5a and b). However, the motion in depth version features a symmetrically expanding gap that opens centrally in each eye sequentially. This gives explicit information that the surfaces are equal in width (Figure 5c) and that the abutting location of the mutually occluding/camouflaging edges must be central. These two pieces of information lead us to a unique geometrical solution (given the assumptions of rigidity and solidity), removing any ambiguity and, with it, the depth attenuation found with static Unmatched Door stimuli.

Dynamic half-occlusion: A special case of IOVD?

We have argued that although a velocity of zero is valid as a second monocular input to the IOVD system, allowing a motion in depth trajectory and speed to be uniquely computed (given the viewing distance and interocular separation), a difference between two velocities being computed when only one feature is present to carry such a velocity signal is difficult to imagine. If, in general, this combination was sufficient as an input to such a system, unambiguous motion in depth should also have been perceived in the monocular control conditions. However, it has been suggested that the visual system might deal with largely binocular stimuli with an unmatched feature in a
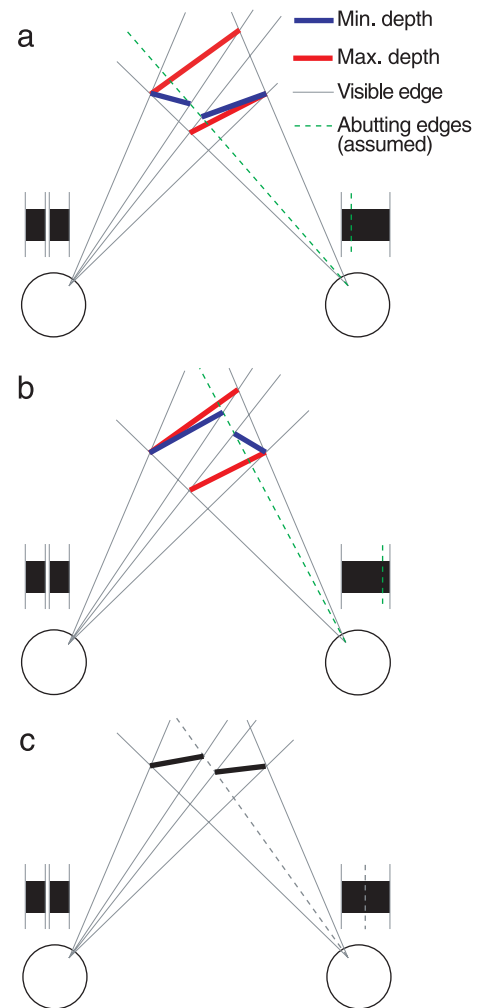

Figure 5. Perceived depth and lines of constraint. (a) Binocular geometry produces constraints of minimum (bold blue lines) and maximum depth (bold red lines). While the maximum depth constraint corresponds to total overlap of the surfaces in the eye seeing no gap, the minimum depth corresponds to no overlap, where the two surfaces' inner edges abut from this viewpoint. Three possible resolutions of the same static monocular stimuli are shown, where the assumed location of the abutting edges lies (a) to the left or (b) to the right of center. (c) When the motion of the stimulus specifies no overlap and equal surface width, stimulus slant is fully specified by binocular geometry.

very different way compared to purely monocular stimuli. Since we have already considered and rejected the possibility of unorthodox matches, alternative schemes must be considered. One might suggest the possibility that, in this instance, an IOVD system might use a default value of zero as a second input in lieu of an explicit signal. However, when objects are visible only in one eye, such as in our monocular control conditions, this may fail. Such a process would explain our data without any need to appeal to ecological constraints. In Experiment 2, we assessed the perception of motion in depth for ecologically valid and invalid stimuli.

Experiment 2

In this experiment, we used the Matched and Unmatched Door stimuli from Experiment 1. Half-images constituted ecologically valid pairs for binocular fusion, and thus, a stable percept of motion in depth was expected, as in Experiment 1. We compared this with the motion in depth seen at similar gaps in half-images that did not constitute ecologically valid pairs. These half-images had the same widths (equal to those of the valid stimuli) but very different heights (one being three times the other). If a gap was present in both images (the Invalid Matched condition), then an IOVD signal and a CD signal would be available, although in a somewhat rivalrous context. We predicted that depth would be seen in this case. However, if a gap was present in one image only (the Invalid Unmatched condition), then no IOVD signal would be available unless a default value of zero was assigned to the rectangle in the other eye. Unlike the Valid Unmatched condition, the binocular context was not consistent with attribution of the gap to monocular occlusion in this case, and we predicted that motion in depth would resemble that obtained in the control conditions of Experiment 1 in which only one image was present at any one time.

Methods

The details for Experiment 2 differed from those in Experiment 1 only in the following respects. We used a 2×2 design where the two factors were occlusion validity (two levels: Valid and Invalid) and matchability (two levels: Matched and Unmatched). While vertical dimensions of the Valid stimuli were the same as in Experiment 1, Invalid stimuli measured 180 min in height. Gap motion was confined to one half-image throughout the motion sequence of each stereo pair. This resulted in a "swinging doors" percept similar to that seen before, although instead of traversing the plane of the screen, each door seemed to "rebound" and continue its motion while remaining on the same side. This continued at 0.5 Hz, resulting in the same speeds as in Experiment 1. For invalid stimuli, the changing gap always appeared on the smaller of the two half-images. However, which eye received which half-image was determined at random, varying the depth order (left near or right near). Again, observers were required to match the amplitude and direction of motion in depth for each stimulus to those in a probe stimulus identical to the Valid Matched stimulus. The probe had a randomly determined initial depth order and amplitude. These parameters could be changed independently by the observer, using the cursor keys. Depth settings consistent with the depth order of the stimulus (as defined by which eye saw a changing gap) were scored as positive, whereas those not corresponding to the correct depth order were scored negatively. Two observers contributed 20 probe settings for each condition. These were author B.G. and one naïve observer from Experiment 1: D.B. Statistical significance was assessed by means of 2×2 repeated measures ANOVAs with a planned contrast to examine the difference between Unmatched Valid and Unmatched Invalid conditions for each subject.

Results

Results for both subjects are shown in Figure 6. Here, large, near-veridical settings are made for Valid stimuli, as predicted from the results of Experiment 1. For the Invalid stimuli, Matched and Unmatched stimuli show very different patterns of responding. While the Matched condition consistently produced large settings of the correct depth order, the Invalid case resulted mostly in near-zero settings or larger settings that could be either positive or negative in sign. In line with these results, both observers often reported a multistable percept of motion in depth, where each panel could appear to swing in either direction independently. As in Experiment 1, we believe that this ambiguous motion in depth percept is due to monocular foreshortening of the image and is quite distinct from the impression created by all other stimuli, where the direction of depth or motion in depth was never unclear.

The results of two-way ANOVAs revealed a significant interaction effect for each subject, $F(1,19) \geq 5.967$, $p \leq .0245$. In addition, linear contrasts confirmed the specific difference in settings for Matched Invalid and Unmatched Invalid conditions, $F(1,19) \geq 7.808$, $p \leq .012$.
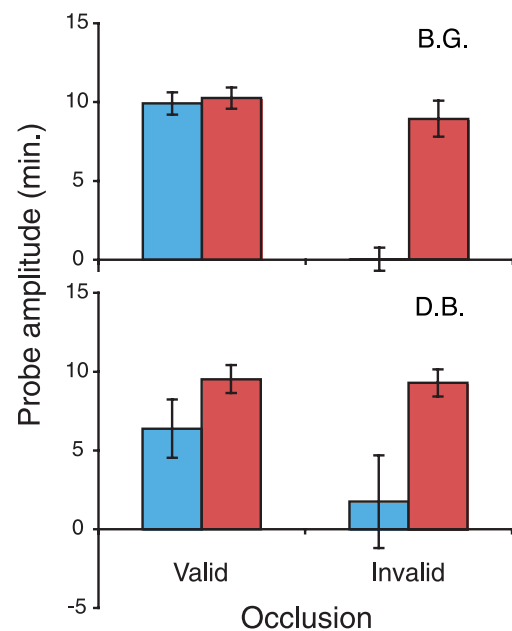

Figure 6. Results for Experiment 2. Light blue bars represent unmatched stimuli, whereas red bars represent matched stimuli. Error bars represent 95% confidence intervals.

Discussion

In this experiment, matchable vertical edges could yield stable and consistent motion in depth percepts through the CD and IOVD cues even when the horizontal edges of the two half-images did not correspond. However, when the vertical inner edges were unmatched, no such percept was evident unless the monocular figures were of the same height. Changes in the width of the monocular gap are effective as a motion in depth cue only when presented in a binocular context where an occlusion solution is valid. This allows us to discount the possibility that our phenomenon is a special case of IOVD, where an explicit motion signal in one eye is combined with a default zero motion signal in the other. If this process were responsible, motion in depth should have been seen equally in all conditions. Instead, we must regard the effect of dynamic half-occlusion as an example of the reconstruction of the 3D details of a changing external environment by the imposition of constraints of ecologically validity without relying on binocular matching.

In Experiment 1, the constraints of ecological optics give information that the two panels abut in one eye's view, as discussed previously. The constraints are slightly weaker in Experiment 2, where the planes never actually pass each other and, hence, need not abut or be of equal size; this may explain why subject D.B.'s settings for the Unmatched Valid condition were slightly lower than her settings for the Unmatched condition in Experiment 1. Although the constraints are weakened, the symmetrical motion sequence may support the interpretation that the panels are similar in size and that the edges of both are near the center. However, this information does not mean that we can regard the stimulus for motion in depth as a CD or IOVD cue between edges in one eye and an inferred contour in the other eye. A stereoscopic process in which one of the matching contours is not physically present but is inferred or implicit from contextual information is very different from conventional stereopsis and cannot be accounted for by stereo models based on conventional disparity. This has been recognized in the static literature. In the motion in depth case, it represents a clearly different stimulus from IOVD or CD and deserves investigation and recognition as a separate cue. This may involve investigating how this new source of motion in depth information is related to the two conventional stereomotion cues or to monocular sources of motion in depth information.

Conclusions

The perception of motion in depth from a change in the extent of binocularly unmatched features has not previously been reported. We can now include dynamic half-occlusion as a third stereomotion cue, along with the established cues of CD and IOVD. This new cue's importance is clear since it is available when and only when the other two are absent, as both CD and IOVD rely on explicit matchable features.

References

Brooks, K. (2001). Stereomotion speed perception is contrast dependent. *Perception, 30,* 725–731. [PubMed]

Brooks, K. R. (2002a). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision, 2*(3), 218–231, http://journalofvision.org/2/3/2/, doi:10.1167/2.3.2. [PubMed] [Article]

Brooks, K. R. (2002b). Monocular motion adaptation affects the perceived trajectory of stereomotion. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1470–1482. [PubMed]

Brooks, K., & Mather, G. (2000). Perceived speed of motion in depth is reduced in the periphery. *Vision Research, 40,* 3507–3516. [PubMed]

Brooks, K. R., & Gillam, B. J. (2006). Quantitative perceived depth from sequential monocular decamouflage, *Vision Research, 46,* 605–613. [PubMed]

Brooks, K. R., & Stone, L. S. (2004). Stereomotion speed perception: Contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision, 4*(12), 1061–1079, http://journalofvision.org/4/12/6/, doi:10.1167/4.12.6. [PubMed] [Article]

Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., & Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception, 24,* 33–47. [PubMed]

Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anticorrelated stereograms: Effects of spatial density and interocular delay. *Vision Research, 33,* 1959–1975. [PubMed]

Collewijn, H., Erkelens, C. J., & Regan, D. (1985). Neither ocular vergence nor absolute disparity changes produce motion-in-depth sensation in man. *Journal of Physiology, 366,* 16P.

Cook, M., & Gillam, B. (2004). Depth of monocular elements in a binocular scene: The conditions for da Vinci stereopsis. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 92–103. [PubMed]

Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion in depth. *Vision Research, 34,* 483–495. [PubMed]

Cumming, B. G., & Parker, A. J. (1997). Response of primary visual cortical neurons to binocular disparity without depth perception. *Nature, 389,* 280–283. [PubMed]

Erkelens, C. J., & Collewijn, H. (1985). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research, 25,* 583–588. [PubMed]

Fernandez, J. M., & Farrell, B. (2005). Seeing motion in depth using inter-ocular velocity differences. *Vision Research, 45,* 2786–2798. [PubMed]

Gillam, B. (1995). Matching needed for stereopsis. *Nature, 373,* 202–203. [PubMed]

Gillam, B., Blackburn, S., & Nakayama, K. (1999). Stereopsis based on monocular gaps: Metrical encoding of depth and slant without matching contours. *Vision Research, 39,* 493–502. [PubMed]

Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception, 17,* 603–608. [PubMed]

Gillam, B., & Nakayama, K. (1999). Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. *Vision Research, 39,* 109–112. [PubMed]

Grove, P. G., Brooks, K. R., Anderson, B. L., & Gillam, B. J. (2006). Monocular transparency and unpaired stereopsis. *Vision Research, 46,* 1695–1705. [PubMed]

Häkkinen, J., & Nyman, G. (1997). Occlusion constraints and stereoscopic slant. *Perception, 26,* 29–38. [PubMed]

Harris, J. M., & Watamaniuk, S. N. (1995). Speed discrimination of motion-in-depth using binocular cues. *Vision Research, 35,* 885–896. [PubMed]

Lee, D. N. (1970). Binocular stereopsis without spatial disparity. *Perception and Psychophysics, 9,* 216–218.

Malik, J., Anderson, B. L., & Charowhas, C. E. (1999). Stereoscopic occlusion junctions. *Nature Neuroscience, 2,* 840–843. [PubMed] [Article]

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London Series B, 204,* 301–328. [PubMed]

Mayhew, J. E. W., & Frisbee, I. P. (1981). Psychophysical and computational studies towards a theory of human stereopsis. *Artificial Intelligence, 17,* 349–385.

Mitchell, D. E. (1969). Qualitative depth localization with diplopic images of dissimilar shape. *Vision Research, 9,* 991–994. [PubMed]

Mitchell, D. E. (1970). Properties of stimuli eliciting vergence eye movements and stereopsis. *Vision Research, 10,* 145–162. [PubMed]

Nakayama, K., & Shimojo, S. (1990). da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research, 30,* 1811–1825. [PubMed]

Panum, P. L. (1858). *Physiologische Untersuchungen über das Sehen mit zwei Augen.* Kiel: Schwerssche Buchhandlungen.

Pianta, M. J., & Gillam, B. J. (2003a). Monocular gap stereopsis: Manipulation of the outer edge disparity and the shape of the gap. *Vision Research, 43,* 1937–1950. [PubMed]

Pianta, M. J., & Gillam, B. J. (2003b). Paired and unpaired features can be equally effective in human depth perception. *Vision Research, 43,* 1–6. [PubMed]

Pope, D. R., Edwards, M., & Schor, C. S. (1999). Extraction of depth from opposite-contrast stimuli: Transient system can, sustained system can't. *Vision Research, 39,* 4010–4017. [PubMed]

Prazdny, K. (1984). Stereopsis from kinetic and flicker edges. *Perception & Psychophysics, 36,* 490–492. [PubMed]

Ramachandran, V. S., Rao, V. M., & Vidyasagar, T. R. (1973). The role of contours in stereopsis. *Nature, 242,* 412–414. [PubMed]

Rashbass, C., & Westheimer, G. (1961). Dysjunctive eye movements. *Journal of Physiology, 159,* 339–360. [PubMed] [Article]

Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology & Visual Science, 27,* 584–597. [PubMed]

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research, 40,* 2565–2572. [PubMed]

Tsai, J. J., & Victor, J. D. (2000). Neither occlusion constraint nor binocular disparity accounts for the perceived depth in the 'sieve effect'. *Vision Research, 40,* 2265–2276. [PubMed]

Wilcox, L. M. (1999). First and second-order contributions to surface interpolation. *Vision Research, 39,* 2335–2347. [PubMed]

Wilcox, L. M., & Hess, R. F. (1996). When stereopsis does not improve with increasing contrast. *Vision Research, 38,* 3671–3679. [PubMed]

Ziegler, L. R., & Hess, R. F. (1999). Stereoscopic depth but not shape perception from second-order stimuli. *Vision Research, 39,* 1491–1507. [PubMed]