Phantom surfaces in da Vinci stereopsis

Susan G. Wardle

School of Psychology, The University of New South Wales, Sydney, Australia

Barbara J. Gillam

School of Psychology, The University of New South Wales, Sydney, Australia

In binocular viewing of natural three-dimensional scenes, occlusion relationships between objects at different depths create regions of the background that are visible to only one eye. These monocular regions can support depth perception. There are two viewing conditions in which a monocular region can be on the nasal side of a binocular surface—(a) when a background surface is viewed through an aperture and (b) when a region is camouflaged against the background in one eye's view. We created stimuli with a monocular region using complex textures in which camouflage was not possible, and for which there was no physical aperture. For these stimuli, observers perceived a strong phantom contour in near depth at the edge of the monocular region, with the monocular texture perceived behind at the depth of the binocular surface. Depth-matching with a probe showed that the depth of the phantom occluding surface was as precise as for stimuli with regular binocular disparity. Monocular regions of texture on the opposite (temporal) side of the binocular surface were perceived behind, as predicted by occlusion geometry, and there was no phantom surface. We discuss the implications for models of da Vinci stereopsis and stereoscopic edge processing, and consider the involvement of a form of Panum's limiting case. We conclude that the visual system uses a combination of occlusion geometry and complex matching to precisely locate edges in depth that lack a luminance contour.

Introduction

Stereopsis, a powerful source of depth perception, is based on the slightly different perspective views of a three-dimensional (3D) scene presented to the left and right eyes (Wheatstone, 1838). Stereopsis requires that elements in the two eyes' views be matched so that perspective differences important for depth perception, such as a horizontal width difference between two matching elements at different depths, (horizontal disparity) can be registered. Most stereoscopic research has focused on the depth perceived from matched binocular regions. However, Leonardo da Vinci (*c.* 1508) observed that when viewing a near surface in front of a background, parts of the background hidden in one eye are visible to the other eye. It was only acknowledged relatively recently that monocular, unmatched regions in binocular scenes were not just noise in the binocular system but conveyers of useful information about spatial layout (for reviews see Gillam, 2011; Harris & Wilcox, 2009).

In several early demonstrations, a monocular region horizontally adjacent to one eye's view of a binocular surface appeared in depth relative to the surface (e.g., Kaye, 1978; Lawson & Gulick, 1967; von Szily, 1921). Additionally, Gillam and Borsting (1988) showed that the presence of the monocular regions in a Julesz random-dot stereogram reduced the latency for resolving depth. Nakayama and Shimojo (1990) conducted the first quantitative study of the perceived depth of monocular regions relative to binocular surfaces. They named depth perception from monocular regions "da Vinci stereopsis." Monocular regions have since been shown to appear in depth in a range of contexts (Brooks & Gillam, 2006; Cook & Gillam, 2004; Häkkinen & Nyman, 1996; Nakayama & Shimojo, 1990;[1] Zannoli & Mamassian, 2011), and can influence the depth of adjacent binocular regions (Gillam, Blackburn, & Nakayama, 1999; Sachtler & Gillam, 2007; Tsirlin, Wilcox, & Allison, 2010). The presence of monocular regions may also induce phantom occluding surfaces to account for the absence of part of the image in one eye (Anderson, 1994; Gillam & Nakayama, 1999; Nakayama & Shimojo, 1990). The term "da Vinci stereopsis" has been defined as any aspect of depth perception arising from monocular regions in a scene (Nakayama & Shimojo, 1990). The present paper investigates the nature of a phantom contour that arises in the conditions originally described by Leonardo da Vinci (*c.* 1508), in which a

Citation: Wardle, S. G., & Gillam, B. J. (2013). Phantom surfaces in da Vinci stereopsis. *Journal of Vision, 13*(2):16, 1–14, http://www.journalofvision.org/content/13/2/16, doi:10.1167/13.2.16.

doi: 10.1167/13.2.16 Received August 20, 2012; published February 8, 2013 ISSN 1534-7362

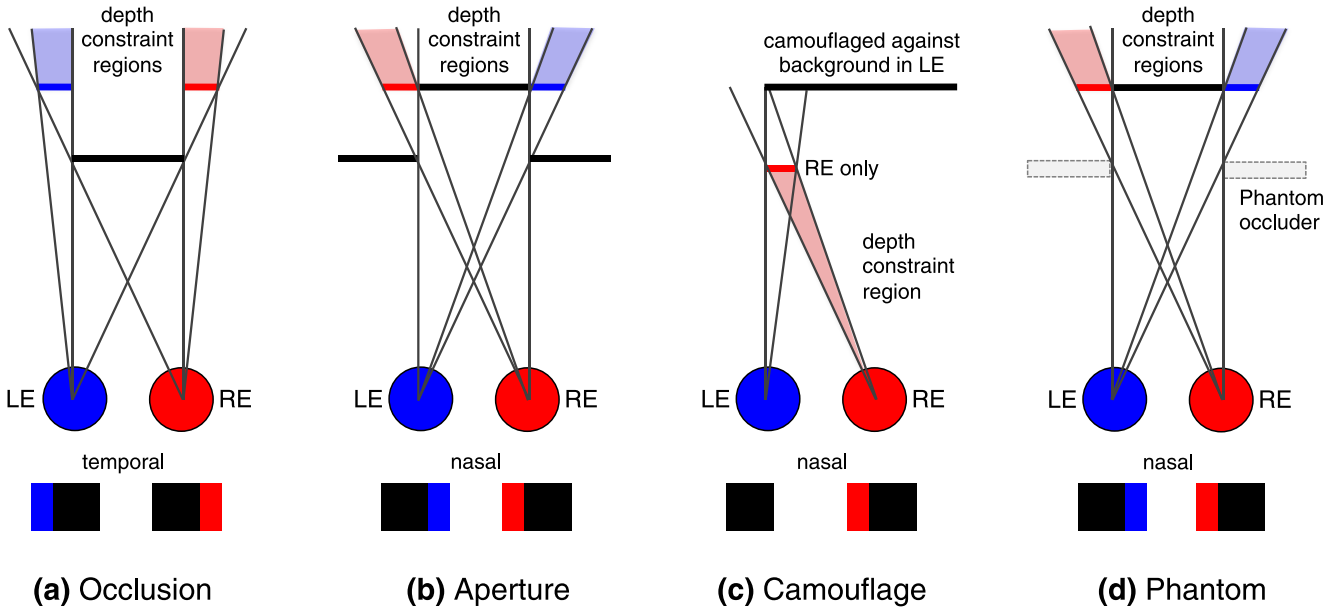

Figure 1. An aerial view of scene geometry in da Vinci stereopsis. Color is a representative marker for visibility either for the left eye (blue) or right eye (red) and does not indicate the actual color of the surface. (a) In the occlusion configuration, the monocular regions are occluded in the other eye's view by the black foreground surface. The rectangles below the figure show the monocular half-images—the monocular region is on the temporal side of the foreground surface in each eye. (b) In the aperture case, the monocular regions are occluded in the other eye by the foreground aperture. The monocular regions are on the nasal side of the background surface as shown in the half-images. (c) For camouflage, the monocular region on the nasal side (shown for right eye only) is invisible to the left eye because it is camouflaged against the background surface (thus the color and luminance of both surfaces must match). (d) The phantom is an alternative explanation for the presence of a monocular region on the nasal side; it is hidden in the other eye's view by a phantom occluding surface. See text for a description of the depth constraint regions in a–d.

monocular region horizontally adjacent to a binocular surface is seen in depth.

Conditions for da Vinci stereopsis

The location of monocular regions in binocularly viewed 3D scenes is determined by the occlusion relationships of near and far surfaces between the two eyes. In the first configuration described by Leonardo da Vinci (*c.* 1508), monocular regions arise because a near surface occludes part of a more distant surface in the other eye's view (Figure 1a). In this case the monocular regions are on the temporal side of the binocular surface (Nakayama & Shimojo, 1990). In the second case identified by da Vinci, monocular regions result from viewing a background surface through an aperture. As illustrated in Figure 1b, in this case each eye sees more of the background surface on the nasal side of the binocular region.

There is another situation in addition to occlusion that can produce regions visible to only one eye. A surface nearer to the observer with the same color and luminance as a binocular background surface may be invisible to one eye because it is camouflaged against the background surface in that eye's view (Cook & Gillam, 2004; Gillam, 2011; Kaye, 1978; Nakayama & Shimojo, 1990; von Szily, 1921). The early stereograms of von Szily (1921) demonstrate that depth can readily be perceived under these conditions (see example stereogram in Figure 2). A black monocular region attached to the nasal side of a black binocular surface appears nearer in depth than the binocular surface, as predicted by camouflage geometry (Figure 1c).

The geometry of da Vinci stereopsis

Compared to depth from binocular disparity, which unambiguously specifies the precise location of an object in 3D space (after appropriate scaling for distance and eccentricity), the depth of monocular regions is only partially constrained. The depth constraints for four different viewing conditions are illustrated in Figure 1. In the occlusion situation (Figure 1a), the monocular region is horizontally adjacent to the temporal side of a binocular surface. The minimum depth separation between the monocular region and the binocular occluding surface is constrained because if the monocular region were nearer to the occluder, it would be visible to the contralateral eye. This "minimum depth constraint" increases with the

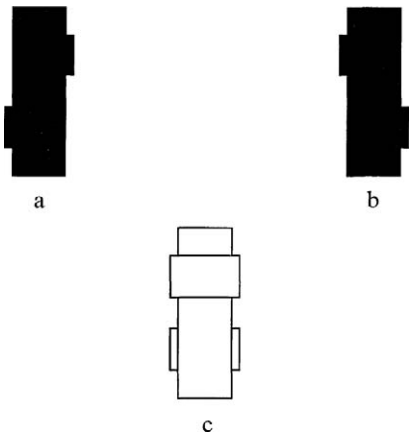

Figure 2. von Szily (1921) stimuli reprinted from Ehrenstein and Gillam (1998). Divergent fusion of images (a) and (c) will produce the percept shown in (c). The top rectangle in the fused stereogram (c) is an example of the camouflage configuration; monocular regions attached to a binocular surface on the nasal side are perceived in front of the binocular region. The bottom rectangle is an example of the occlusion configuration; temporal monocular regions are perceived behind the binocular surface. The depth order is reversed for crossed fusion. Note that the monocular tabs appear flat. Figure reproduced with permission from Pion Ltd., London (www.pion.co.uk; www.perceptionweb.com).

horizontal extent of the monocular region (Nakayama & Shimojo, 1990). However, the maximum depth of the monocular region is unconstrained because it could be located at a range of greater depths and still be invisible to the contralateral eye, indicated by the shaded "depth constraint region" in Figure 1a.

There are three possible geometric arrangements for a monocular region on the nasal side of a binocular surface (Figure 1b through d). Figure 1b shows the case of a background surface viewed through an aperture. Here the minimum depth constraint arises from the fact that if the monocular region were nearer to the aperture, it would be partially visible in the ipsilateral eye. In the second arrangement, the monocular region is camouflaged against the background surface in one eye. The minimum depth of the monocular region is constrained because the monocular region must completely overlap with the background surface along the lines of sight for the eye in which it is invisible (Figure 1c). If the monocular region were nearer to the binocular surface than the minimum depth constraint, this condition would not be met.

The third case (Figure 1d), which is geometrically similar to the aperture case, is the focus of this paper. If a monocular region is horizontally adjacent to the nasal side of a binocular surface, but there is no aperture and the ecological constraints for camouflage are not satisfied (e.g., the color or luminance of the monocular and binocular regions do not match), then what is perceived? Assee and Qian (2007) have suggested that a "featureless occluder" could account for a monocular region under these conditions. This would be the geometric equivalent to viewing through a physical aperture as in Figure 1b, but with the aperture formed by subjective rather than real contours (see Figure 1d). Although this seems unlikely, examples of such phantom surfaces in da Vinci type displays have been reported and are described below. The depth of a monocular region with a phantom aperture is under-constrained in the same way as it is for a physical aperture. However, if the border of the texture is taken to be the edge of the figure in each eye's half-image, the depth of the phantom surface itself is *precisely determined* by the lines of sight for the edge of the monocular region in one eye and the edge of the binocular region in the other eye (see Figure 1d). The geometry is consistent with only one depth for the phantom surface. It is possible for the phantom to be located at a greater depth (closer to the observer) if an arbitrary amount of white space on the temporal sides of each eye's half-image is included. If this is the case, only the *minimum* depth of the phantom surface is constrained. Although this is a valid possibility, it is more likely that the texture-defined region of the stimulus is taken as the edge of the figure. Thus, for the rest of the article, we make the explicit assumption that the edge of the figure is the texture-defined border, and consequently, that the location of the phantom surface is precisely determined.

Quantitative studies of da Vinci stereopsis

Even though the depth is under-constrained, the perceived depth of monocular regions in da Vinci stereopsis has been found to satisfy the minimum depth constraint under certain conditions. Cook and Gillam (2004) measured the depth perceived for a monocular tab intrusion into a dumbbell shaped binocular surface. Monocular *intrusions* into binocular surfaces arise from different occlusion arrangements than the more usual monocular *extrusions* shown in Figure 1 (for details see Cook & Gillam, 2004; Gillam, 2011); however, a minimum depth constraint is still specified by the intrusion width. Cook and Gillam found quantitative depth for intrusions in both occlusion and camouflage arrangements, and depth magnitude closely followed the predications of the minimum depth constraint. In a control experiment, Cook and Gillam (2004) demonstrated that observers did not perceive the curved three-dimensional shape, which would be predicted if they had fused the curved edges of the dumbbell with the straight edge of the monocular intrusion. Thus the quantitative depth found for monocular intrusions cannot be attributed to fusion. It is interesting that for

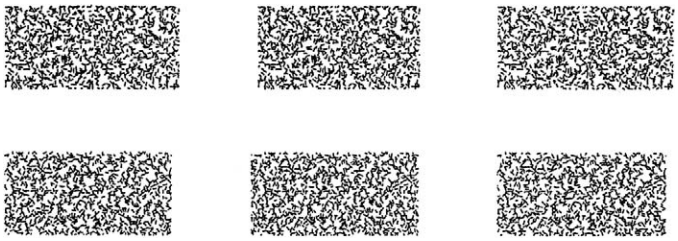

Figure 3. Slant stimuli from Gillam and Blackburn (1998) without monocular sidebands (top row) and with monocular sidebands on either side (bottom row). When the bottom row is fused with crossed fusion, the phantom surface is on the right side of the left pair and the left side of the right pair (reversed for divergent fusion). Figure reproduced with permission from Pion Ltd., London (www.pion.co.uk; www.perceptionweb.com).

other stimuli such as a short monocular bar (Cook & Gillam, 2004) or disc (Gillam, Cook, & Blackburn, 2003) not attached to the binocular region, depth was perceived but was not quantitative (see also Häkkinen & Nyman, 1996). The reasons for the differences in perception of attached and unattached monocular elements in da Vinci type configurations have yet to be determined.

Phantom surfaces and monocular regions

We define phantom surfaces as those that are only seen with binocular viewing and do not have apparent luminance edges. They have been observed in several da Vinci type displays, but in most cases the depth of the phantom surface has not been measured. Nakayama and Shimojo (1990) described seeing a subjective occluder in sparse binocular dot displays, when a column of monocular dots was added to the nasal side of one eye's view, but they did not measure its depth. Gillam and Blackburn (1998) observed a phantom occluder in their investigation of the effect of monocular regions on stereoscopic slant. They added monocular texture to the sides of the narrower binocular image in one eye's view of a slanted surface to eliminate a global image width difference in the two eyes. For the monocular region on the nasal side of the binocular slanted surface, a phantom aperture was observed (see Figure 3). As these stimuli were composed of random textures, camouflage of the monocular region against the binocular surface was not possible. Gillam and Blackburn (1998) found that the presence of the monocular regions enhanced perceived slant of the binocular surface, but they did not measure the depth of the phantom occluder.

Gillam and Grove (2004) measured the depth of a phantom occluding surface in an unusual type of da Vinci display in which a constant additional width was added to one eye's view of a set of horizontal lines aligned on one side but of different widths. The different horizontal disparities that this produced for each line (which would predict different slants) were overridden by the simpler solution of seeing a single vertical (phantom) occluder differentially hiding the lines for the two eyes. This solution was preferred even though it required perceiving a phantom. The depth of the phantom occluder was found to be quantitative, and close to the depth predicted from treating the additional constant width as a monocular region. If the two eyes' views were reversed so that the extra width added to the lines violated occlusion geometry, there was no perceived occluder and the lines did indeed appear at locally different slants according to their individual disparities.

Anderson (1994) showed that vertical image differences result in the perception of oblique phantom occluding contours at the image boundaries. This reflects natural viewing, in which an oblique edge occluding a vertical edge at a greater depth does give rise to vertical image differences in the two eyes. The depth of the phantom occluder in this case has not been measured. Gillam and Nakayama (1999) found that a phantom rectangle is perceived for a stereogram consisting of a pair of thin vertical bars with a central gap in the left bar of the right eye's view and vice versa. The phantom rectangle "accounts for" the monocular gaps. The perceived depth of the phantom increases with increasing width of the bars, although it is seen at a considerably greater depth than that predicted by the minimum depth constraint (Gillam & Nakayama, 1999; Grove, Gillam, & Ono, 2002; Kuroki & Nakamizo, 2006; Mitsudo, Nakamizo, & Ono, 2005).

The aim of the present experiments is to investigate the perceived depth of phantom occluders of the da Vinci type in which a monocular region is horizontally adjacent to the nasal side of one eye's view of a binocular surface (Figure 1d). Under these conditions the monocular region cannot be camouflaged against the surface and there is no physical aperture, and it is known that a phantom occluder will be perceived (Gillam & Blackburn, 1998; Nakayama & Shimojo, 1990). However, it is not known whether the depth of a phantom occluding surface is quantitative or precise under these conditions. This is a particularly interesting question considering the viewing geometry shown in Figure 1d, in which the depth of the phantom is precisely determined by the size of the monocular region if the monocular region is densely textured and it is assumed that its edge is the edge of the surface.

We created stimuli from complex, nonrepeating patterns, in which camouflage of the monocular region against the binocular surface is impossible. In these stimuli, a phantom occluding surface is clearly observed for a monocular region on the nasal side of a binocular surface, and it is only observed with

binocular viewing. As the depth of the phantom occluder in this case (given the edge assumption discussed above) is precisely determined by the size of the monocular region (see Figure 1d) we expected that this would be reflected by accurate measurements of its perceived depth for different sizes of monocular region. In Experiment 1 we compared the depth of the phantom occluding surface to that of temporal monocular regions (which were perceived behind the binocular surface) and also with the minimum depth constraint. In Experiment 2 we added monocular regions to both eyes' images to equate the image width between the two eyes, and compared the depth of the nasal monocular regions with the depth of the phantom surface occluding them.

Experiment 1

Methods

Observers

Four observers with normal stereo vision and normal or corrected-to-normal visual acuity participated. Stereovision was assessed using the Stereo Titmus test (Stereo Optical, Chicago, IL, USA), and a binocular training task (see Procedure). All observers were right-eye dominant (hole-in-the-card test). One observer was an author (SW); the other three observers were naive to the aims of the experiment and were reimbursed for participation. The experimental procedure followed the Declaration of Helsinki.

Stimuli

Stereoscopic stimuli were generated on a MacPro computer with MATLAB (The MathWorks) and functions from the Psychtoolbox (Brainard, 1997; Pelli, 1997). A mirror stereoscope was used to present separate images to each eye, which were displayed on two identical LCD monitors (Dell U2412M) at a screen resolution of 1920 × 1200 pixels.

The background luminance of the two monitors was equal within 2 cd/m^2 (52.73 cd/m^2, 54.55 cd/m^2).

The stimuli were rectangles containing a texture constructed from ellipses of varying size, color, and orientation (see Figure 4 for examples of the stimuli that can be free-fused). This texture was adapted from a "dead leaves" model (e.g.,: Lee, Mumford, & Huang, 2001; Ruderman, 1997; Wardle, Bex, Cass, & Alais, 2012). The important feature of the texture was that its nonrepeating pattern formed a complex visual surface, against which camouflage of a monocular region is not possible. The other features of the texture are not critical to the experimental design (a different random texture is used in Experiment 2). The texture was generated by filling a 3 × 3 degree central region of the screen with 1,000 ellipses with random screen location, size [3.8–28.6 arcmin], and orientation [0–360 degrees] drawn from independent uniform distributions. A unique random texture was generated on every trial. The amount of the texture visible to each eye was varied by overlaying a white image "mask" over the entire screen, which had a transparent window in the center. Consequently, only the portion of the pattern that overlapped with the transparent window was visible in each eye, and the size of the window could be adjusted. The binocular region of the stimuli was 0.9° (width) × 3° (height). Monocular regions were added to one eye's image by extending the horizontal width of

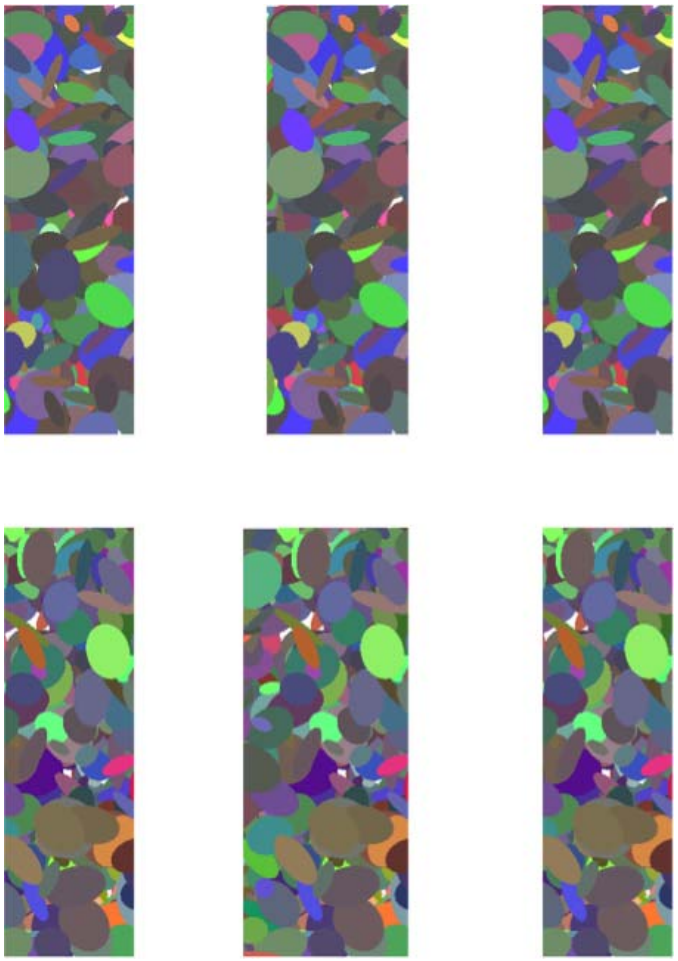

Figure 4. Example of the stimuli for Experiment 1. (top row) Crossed fusion of the middle and left figures demonstrates a temporal monocular region—the monocular strip is perceived behind the binocular surface. Crossed fusion of the middle and right Figures demonstrates a nasal monocular region—a phantom occluder is perceived next to the binocular surface. (Labels are reversed for divergent fusion.) (bottom row) Stimuli with a wider monocular region produce greater depth in both conditions.

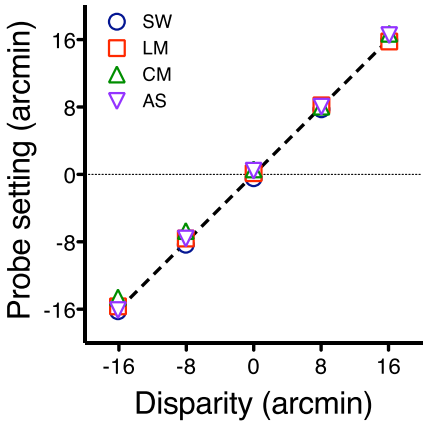

Figure 5. Results for the binocular practice task in Experiment 1, individual data for all four observers. The graph shows binocular probe settings for matching the perceived depth of a vertical bar with depth defined by binocular disparity. The dashed line indicates a perfect match between the disparity of the bar and the probe. Error bars are *SEM*.

the transparent stimulus window in one eye's image. The width of the monocular region was selected from 5, 10, 15, or 20 arcmin.

Monocular regions were presented on the nasal or temporal side of the binocular surface, in the left or right eye, at each of the four widths, producing $2 \times 2 \times 4 = 16$ different conditions. Each condition was repeated five times, in a unique random order for each observer.

Procedure

Stimuli were viewed from a distance of 2 m in a dark laboratory. Observers were seated in front of the mirror stereoscope with head position fixed using a chin and forehead rest. Prior to the experiment, observers practiced the depth-matching task in a training program that used unrelated stimuli. The practice task was necessary, as two of the four observers were not experienced in psychophysical or stereoscopic tasks. In the training task, two black binocular bars (height: 3°, width: 13 arcmin) separated by 45 arcmin were positioned in the center of the screen. One bar was always at zero disparity; the disparity of the other was varied (0, 8, and 16 arcmin, near and far disparity). These disparities were selected to be different from the predicted depth of the monocular regions used in the experiment, but within the same range. A small, black binocular probe dot (diameter: 5 arcmin) was positioned 75 arcmin above the bar with variable disparity, and observers were asked to match the depth of the bar by adjusting the binocular disparity of the probe with the keyboard. The probe moved in 0.4 arcmin steps, and observers were allowed unlimited time to move the probe. In the first version of the task, observers received feedback. The probe changed color from black to orange when it was moved within ±4 arcmin of the bar's disparity. If the probe was moved within ±0.9 arcmin of the disparity of the bar, it turned green, indicating that the depth of the probe was matched to the bar. Following the training with feedback, observers completed the same task without feedback; measurements were made for each disparity three times in random order. The practice data for each observer is shown in Figure 5. All observers were able to accurately match the disparity-defined depth of the vertical bar with the probe for near and far disparities.

After the training task, observers were shown examples of the experimental stimuli in both the nasal and temporal configuration (two examples of each). Observers were asked to freely describe what they saw in the stereograms. All observers spontaneously described the percepts for both conditions, which can be seen in the example stimuli in Figure 4. A temporal monocular region is perceived as an occluded surface located at a depth behind the binocular surface. In the nasal condition, observers did not perceive any depth for the monocular region itself; however, they reported a phantom occluding surface visible at the edge of the binocular figure, on the same side as the monocular region. After observers reported these percepts, the experimental task was explained. In the temporal condition observers were instructed to match the depth of the furthest region of the pattern (i.e., the monocular region), and in the nasal condition they were instead asked to match the depth of the phantom occluding surface.

Observers matched the depth in each condition five times in a random order. The probe (diameter: 5 arcmin) appeared 75 arcmin above the binocular surface at the edge of the figure, which contained either the temporal or nasal monocular region. The disparity of the probe was adjusted using the keyboard. To improve accuracy, observers were instructed to move the probe slightly behind and then slightly in front of the perceived depth of the monocular region (temporal) or phantom occluder (nasal), before matching the depth exactly. Each image was on the screen for a minimum of 4 s before observers could progress to the next trial, the maximum duration was unrestricted.

Results and discussion

The mean probe settings are shown separately for each observer and experimental condition in Figure 6. Observers adjusted the disparity of the binocular probe to match the perceived depth of the monocular region in the temporal condition, and the depth of the phantom occluding surface in the nasal condition. Positive probe settings indicate near disparities, and

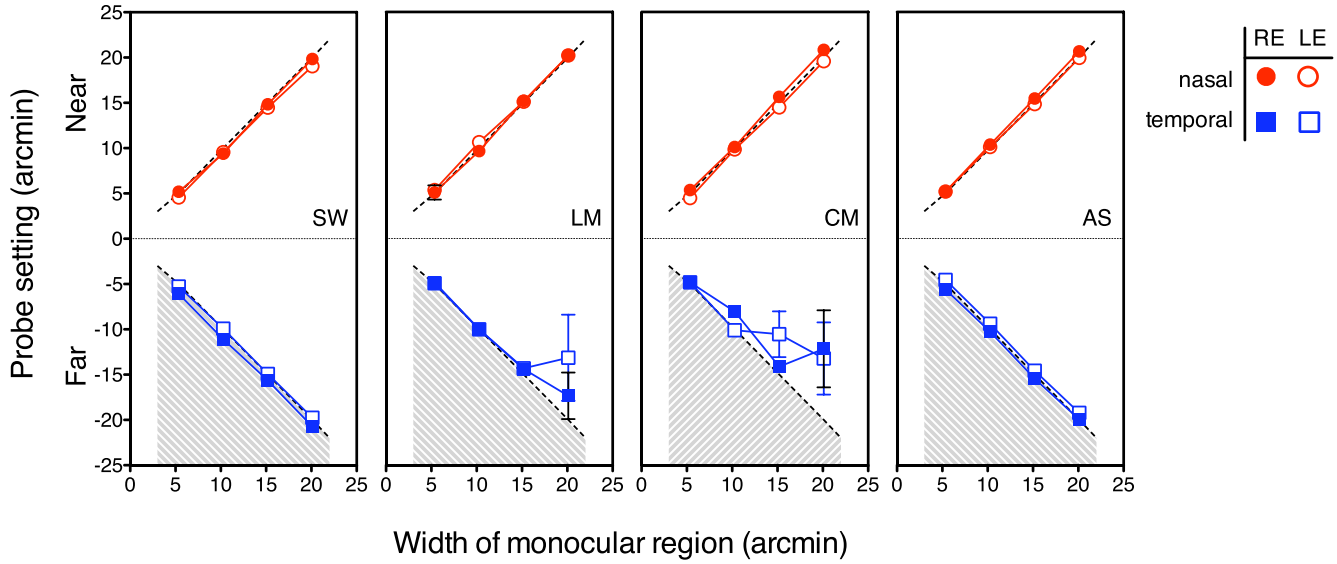

Figure 6. Results for Experiment 1, individual data for all four observers. The graph shows binocular probe settings for temporal (blue squares) and nasal (red circles) conditions as a function of the width of the monocular region. Filled symbols are for the monocular region in the right eye, open symbols for the left eye. The shaded region indicates the depth constraint region for the temporal condition. Error bars show *SEM*.

negative numbers are used for far disparities. Thus the phantom occluder (nasal) was always seen in near depth relative to the binocular surface, and the temporal monocular region was always perceived behind the binocular surface.

In the temporal condition, the mean probe settings closely followed the predictions of the minimum depth constraint (indicated by the dashed line) in the tested range of 5 to 20 arcmin. Similarly, in the nasal condition the mean probe settings matched the predicted depth of the phantom occluder. In both cases, perceived depth of the monocular region (temporal) and the phantom occluder (nasal) increased in direct proportion to the width of the monocular region. This was the case for all observers, across all conditions. The precision of depth matching for the phantom occluder (the nasal condition) was as great as in the practice task in which depth was defined by binocular disparity (compare the standard error bars in Figure 5 with the nasal condition in Figure 6). For the temporal condition, there was greater variability in the settings of the probe for two of the naive observers (LM and CM) for the larger widths of the monocular region. This difference in precision on the matching task may be related to experience; SW (author) and AS are experienced stereo and psychophysical observers and LM and CM did not have prior experience. All observers were right-eye dominant, but there was no systematic difference in the depth perceived when the monocular region was in the left versus the dominant right eye.

Experiments 2A and 2B

In Experiments 2A and 2B we extended the results of Experiment 1 using different stimuli to address three issues. Firstly, in Experiment 1, the image width of each monocular half-image was different in the two eyes because the monocular region was added to only one eye's view. As size disparity is a cue to slant, we wanted to ensure that the depth of the phantom occluder was not determined by the width difference. Secondly, we wanted to confirm that the nasal monocular region was seen at the depth of the binocular surface, and that only the phantom occluder was seen in depth. And finally, we used a different random texture pattern (randomly oriented lines) in Experiment 2 to generalize the results to a new stimulus set. As in Experiment 1, the important feature of this texture is that it is a nonrepeating pattern, thus camouflage of a nasal monocular region against the binocular texture is not possible.

In Experiment 2A, we equated the image width in the left and right eye by adding a monocular region to each eye's view. As in Experiment 1, observers matched the depth of the phantom occluder in the nasal condition, and the monocular region of the pattern in the temporal condition. In Experiment 2B we confirmed that the nasal monocular region was seen at the same depth as the binocular surface, by asking observers to match the depth of the nasal region of the pattern instead of the phantom occluder. In both Experiments 2A and 2B, observers matched the depth of the monocular region in the temporal condition, thus

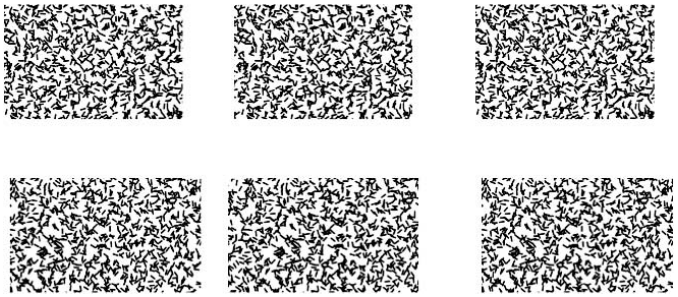

Figure 7. Example of the stimuli for Experiment 2A and 2B. (top row) Crossed fusion of the middle and left figures demonstrates a temporal monocular region—the monocular strip is perceived behind the binocular surface. Crossed fusion of the middle and right figures demonstrates a nasal monocular region—a phantom aperture is perceived to the left and right of the binocular surface (Labels are reversed for divergent fusion). (bottom row) Stimuli with a wider monocular region produce greater depth in both conditions.

this condition was identical in the two experiments. The temporal condition was included in both experiments because half of the observers completed the experiments in the reverse order (Experiment 2B first, Experiment 2A second).

Methods

Observers

The same four observers from Experiment 1 participated in Experiment 2A, and three of these observers also participated in Experiment 2B. The order of the experiments was counterbalanced across observers: LM and CM completed 2B first, SW completed 2A first, and AS completed only 2A.

Stimuli

The stimuli in both Experiments 2A and 2B were pseudorandom line patterns of even density (e.g., Gillam & Blackburn, 1998). Example stimuli for free-fusion are shown in Figure 7. The patterns were composed of 2,400 black short line segments (4.5 [h] × 0.9 [w] arcmin) with random orientation drawn from a uniform distribution [0–360 degrees]. The even density of the pattern was produced by allocating the same number of lines per subsection of the image area—six lines were drawn per 9 × 9 arcmin region of the stimulus. Within each region, the placement of the lines was random (x and y coordinates of each line were drawn from a uniform distribution). The stimulus was presented through a transparent window on the screen as in Experiment 1, to control how much of the pattern was visible in each eye's image. The binocular region of the line pattern was 1.5 (h) × 2.2 (w) deg. Monocular regions were added using the same method as in Experiment 1; however, they were always added to the image in both eyes. For example, in the nasal condition, a monocular region was added to the right side of the left eye's image, and to the left side in the right eye's image. This means the width of the stimuli was always identical in the two eyes, but each eye saw a unique vertical strip of the pattern on one side of the binocular surface that was invisible to the other eye. A unique random line pattern was generated for each trial in the experiment.

Procedure

The experimental procedure was the same as in Experiment 1. As the monocular region was now on both sides of the binocular surface, the probe appeared above the left side for half of the trials, and the right side for the other half. Observers were instructed to match the side of the figure that the probe appeared above. As in Experiment 1, prior to starting the experiment observers were shown examples of the stimuli and asked to describe what they saw (see Figure 7). All observers described the nasal monocular region as though looking through a transparent "window" (aperture) that was in the foreground, with the line pattern perceived at the depth of the screen. In the temporal condition they saw the two temporal strips of the pattern at a depth behind that of the binocular surface, as though the binocular region was occluding a surface behind it.

Monocular regions were presented on the nasal or temporal side of the binocular surface, at each of the four widths (5, 10, 15, 20 arcmin), resulting in $2 \times 4 = 8$ different conditions. Each condition was repeated six times, in a unique random order for each observer.

Results and discussion

The mean probe settings in Experiment 2A are shown separately for each observer in Figure 8. In the nasal condition observers matched the depth of the phantom occluder (aperture), and in the temporal condition they matched the depth of the monocular region of the random line texture. The phantom aperture (nasal) was always seen in near depth relative to the binocular surface, indicated by the positive probe settings for all observers in this condition, and closely followed the predicted depth. The temporal monocular region was always perceived behind the binocular surface, specified by negative probe settings. In the temporal condition the mean probe settings closely followed the minimum depth constraint (indicated by the dashed line) in the tested range of 5 arcmin to 20 arcmin, for all four observers. There is no eye condition

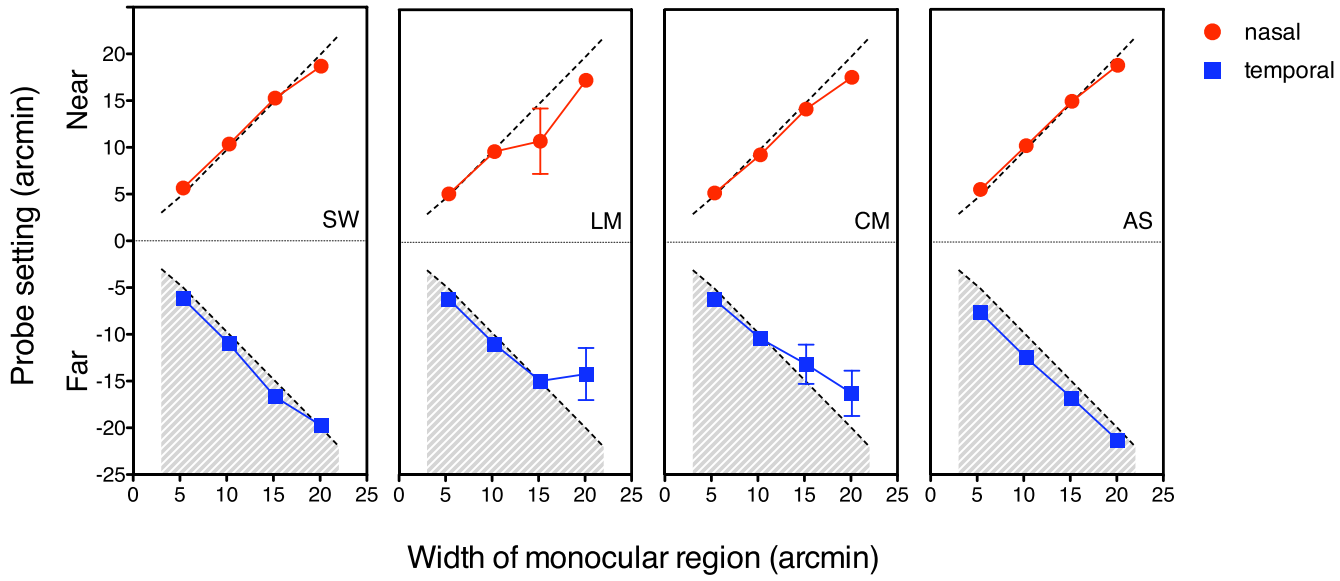

Figure 8. Results for Experiment 2A, individual data for all four observers. The graph shows binocular probe settings for temporal (blue squares) and nasal (red circles) monocular regions as a function of the width of the monocular region. Observers matched the depth of the monocular lines when temporal, and the depth of the phantom aperture when nasal. The shaded region is the depth constraint region for the temporal condition. Error bars show *SEM*.

in Experiment 2 as a monocular region was always presented to both eyes, and probe settings for matching the monocular region on the left or right side of the binocular figure were combined for analysis.

In Experiment 2B (Figure 9), observers matched the depth of the monocular region of the line pattern in both temporal and nasal conditions (i.e., not the phantom occluder). They were instructed to match the depth of the (monocular) lines at the far edge of the texture in both nasal and temporal conditions. In this case, the nasal probe settings were at zero disparity for two observers, demonstrating that the nasal monocular region was perceived at the same depth as the binocular surface. The probe settings for the third observer (LM)

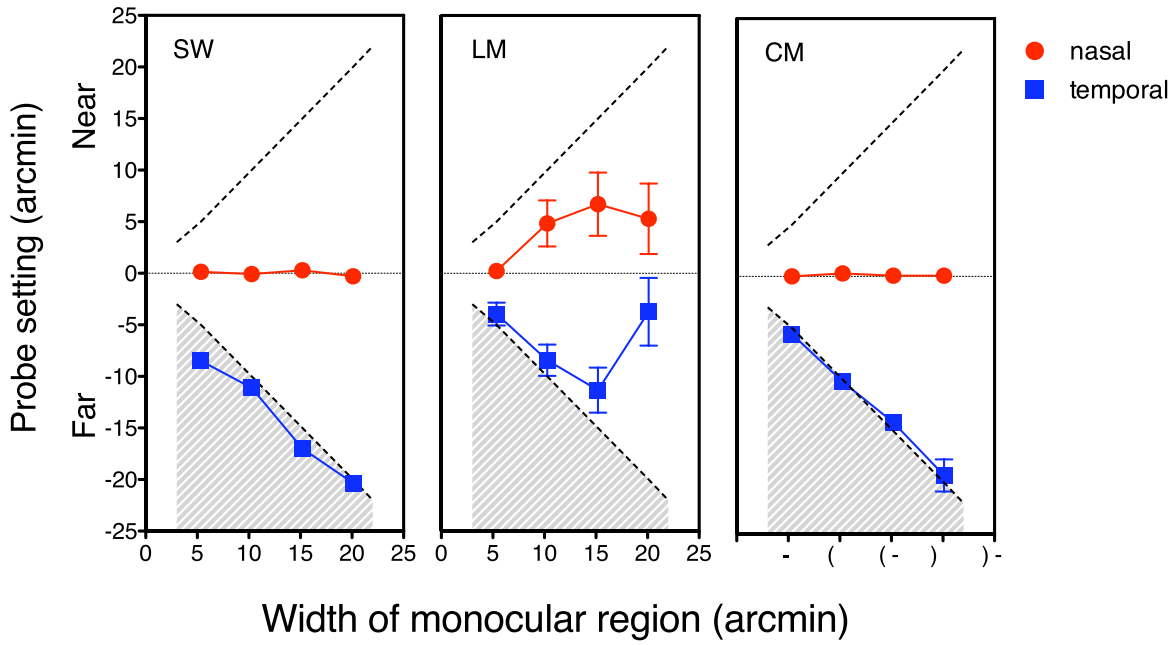

Figure 9. Results for Experiment 2B, individual data for all observers. Observers matched the depth of the monocular lines in both nasal (red) and temporal (blue) conditions. The shaded region is the depth constraint region for the temporal condition. Error bars show *SEM*.

have greater variability; the mean for this observer is only at zero disparity when the monocular region is 5 arcmin wide. For wider monocular regions (above 5 arcmin) the mean probe settings for this observer remain around 5 arcmin. The temporal condition in Experiment 2B was identical to that in Experiment 2A, and again the monocular region was perceived as behind the monocular surface, reflected in the negative probe settings for all observers. The greater variability in probe settings for observer LM in the temporal condition in Experiment 2B (particularly for the widest monocular region of 20 arcmin) compared to Experiment 2A may be an effect of practice because this observer completed Experiment 2B first.

General discussion

We have demonstrated quantitative depth arising from monocular regions attached to binocular surfaces. A monocular region on the temporal side of the binocular surface was perceived to be located behind the binocular surface. This is consistent with occlusion geometry, in which part of the background is hidden from one eye's view by a nearer surface (Figure 1a). The depth of temporal monocular regions closely followed the minimum depth constraint. However, in our experiments a nasal monocular region was not perceived in near depth; this would only be expected if the conditions for monocular camouflage were satisfied (Figure 1c), which was not the case for our stimuli. Instead, nasal monocular regions were perceived at the same depth as the binocular surface accompanied by a phantom occluder at a near depth, geometrically accounting for the absence of the monocular region in the other eye (see Figure 1d). The depths of the phantom surfaces were quantitative, and closely followed the depth predicted by the geometry illustrated in Figure 1d. This is the first time that quantitative depth has been measured for a phantom occluder in this form of da Vinci arrangement. Importantly, the results of Experiment 2 demonstrate a qualitative dissociation between nasal and temporal monocular regions. Temporal monocular regions appear at a far depth, whereas nasal monocular regions appear at the same depth as the binocular surface, with a phantom surface perceived in near depth. We do not believe that existing models of the processing of monocular regions in binocular depth perception can account for this combination of results. We elaborate on the issues below.

Several stereo models incorporate the depth of monocular regions (e.g., Assee & Qian, 2007; Grossberg & Howe, 2003; Hayashi, Maeda, Shimojo, & Tachi, 2004; Watanabe & Fukushima, 1999). They model the situations shown in Figures 1a (occlusion) and 1b (aperture), but not 1c (camouflage) and 1d (phantom). Although Assee and Qian (2007) illustrate the phantom resolution to refute the idea that nasal monocular regions are necessarily invalid, and it is mentioned as a possibility by Watanabe and Fukushima (1999), neither of these authors include it in their models. In general, the models match regions of binocular texture, locate unpaired points that correspond to monocular regions, and then assign the monocular regions to the depth of the background surface, following the observation of Julesz (1971). It is assumed that monocular regions represent part of a continuous background surface. In our experiments, temporal monocular regions are perceived behind the occluding binocular surface, but there is no background surface. Assee and Qian (2007) develop the most physiologically plausible model of da Vinci stereopsis, but like other models theirs requires the presence of a background surface for locating monocular regions in depth. These authors recognize that this requirement creates a difficulty in explaining the depth of monocular regions when the background surface is featureless, as is the case in our stimuli. They deal with this "atypical" situation by attributing the depth to Panum's limiting case or double fusion of a monocular line with two binocular lines. Given this claim and also given that no current models attempt to explain the phantom occluder, it is worth considering matching in our stimuli, and the degree to which disparity based on matching processes can account for our results.

It can be seen from Figure 1d that the location of the phantom occluder in depth could be specified by two matches: (a) between the outer edge of the monocular texture (a luminance edge) and the outer edge of the binocular texture in the other eye (also a luminance edge) and (b) between the edges of the binocular texture in both eyes, only one of which is a luminance edge. This would imply a form of double matching, since the edge of the binocular texture in the eye without the monocular region is involved in both matches. However, given that one of the matches is not between luminance edges it would be an unusual form of Panum's limiting case. It should also be noted that matching of the entire vertical border is not necessary as shown in the example stereogram of Figure 10. Here a phantom occluder is still perceived for a nasal monocular region that is half the height of the binocular surface.

Although the above form of Panum's limiting case could account for the depth of the temporal monocular region, it does not account for the perception of the nasal region. In this case the depth of the phantom edge is perceived as near (a possible form of double matching as described above), but the monocular texture is separated from it and seen at the depth of the

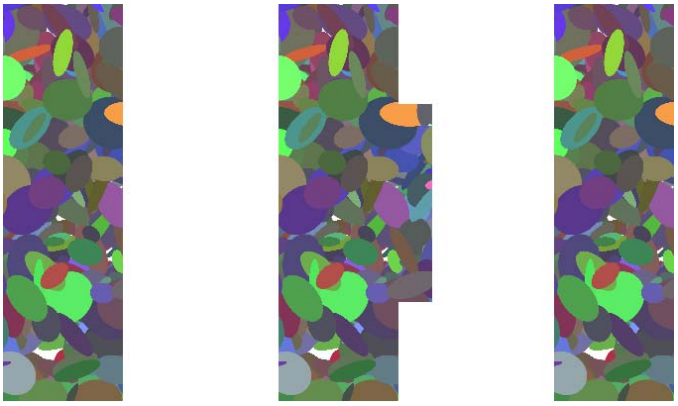

Figure 10. Example stereogram showing that a phantom contour is perceived even when the monocular region is half the height of the binocular surface. Crossed fusion of the middle and left figures demonstrates a nasal monocular region—a phantom occluder is perceived next to the binocular surface. Crossed fusion of the middle and right figures demonstrates a temporal monocular region—the monocular strip is perceived behind the binocular surface. The stimuli are similar to those used in Experiment 1.

binocular surface. It is clear that Panum-type matches are interpreted differently on the nasal and temporal sides of a binocular surface. This dissociation can only occur because the geometry of occlusion is incorporated into the interpretation via mechanisms that are not yet understood.

Overall we think it likely that our results are attributable to an interaction between binocular matching and the implementation of occlusion geometry. These two forms of constraint are involved to different extents in the depth perceived in experimental stimuli with monocular regions. At one extreme is the phantom rectangle (Gillam & Nakayama, 1999; Grove, Gillam, & Ono, 2002; Kuroki & Nakamizo, 2006; Mitsudo, Nakamizo, & Ono, 2005), which generates quantitative depth based on monocular regions despite a complete absence of disparities in the stimulus. At the other extreme are stimuli initially thought to demonstrate depth from monocular regions where the depth has since been attributed to matched disparate features (Liu, Stevenson, & Schor, 1994, 1997). There is a final category in which the depth is clearly due to an interaction. For example, in monocular gap stereopsis depth is seen at the monocular gap even if there is no disparity at the stimulus edges, but the magnitude and precision of the depth is influenced by the disparity there (Pianta & Gillam, 2003). Although our stimuli are very different from this we believe they also involve an interaction.

There is another example in the literature that may involve an interaction between matching and occlusion geometry, in which the presence of monocular regions influences the perception of a binocular surface (Tsirlin, Wilcox, & Alison, 2010). In this case the surface was a white rectangle surrounded by a binocular frame of random dot texture. Inside the frame and adjacent to the white rectangle was a square of random dot texture with near disparity. The white rectangle was visible in the monocular half-images and had luminance-defined edges, thus it was not a phantom surface. However, the addition of monocular regions of texture to the display influenced the percept of the white rectangle in depth. When monocular strips of texture were added to the nasal side, the white rectangle appeared in near depth, in front of the binocular frame. When monocular regions were added to the temporal side, the white rectangle appeared behind the binocular frame. The depth of the white rectangle in both cases was quantitative, as measured by a depth-matching task. However, because this depth was also predicted by disparity in the stimulus (either by edge-matching or the size disparity of the white rectangle), one side of the binocular frame was removed in an additional experiment to control for binocular matching. In this experiment the depth was no longer quantitative. This suggests that edge-matching was involved in the quantitative depth perceived in the original display. In a final experiment using only occlusion conditions, quantitative depth was restored in the stimuli missing one edge of the binocular frame when the random-dot square was placed at the same disparity as the binocular frame. In this case the white rectangle appeared slanted, which suggests the involvement of size disparity that was also present. The interaction between matching and occlusion geometry is complex in these stimuli, because in all of the experiments the introduction of monocular regions also created a size disparity in the white rectangle, and a position disparity of the white rectangle relative to the binocular frame. Thus it is difficult to separate the contribution of matching and occlusion geometry to the depth perceived.

Our results also suggest an interaction between occlusion geometry and edge matching. However, our stimuli are very different because the phantom surface is not defined in the monocular half-images, and is only present with binocular viewing and when the monocular region of texture is added to the nasal side. Although the white rectangle is referred to as an "illusory occluder" (Tsirlin, Wilcox, & Allison, 2010), it has luminance-defined edges that are visible in each monocular half-image. Unlike our phantom occluder, the white rectangle is visible in both temporal and nasal conditions in their experiments. The phantom surface in our stimuli is only visible with camouflage geometry. Thus we have demonstrated that a phantom occluding surface with quantitative depth is employed by the visual system to resolve nasal monocular regions when

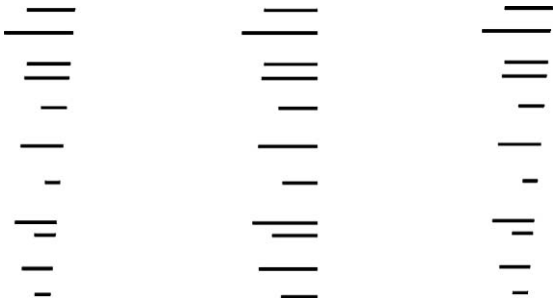

Figure 11. Stimuli from Gillam and Grove (2004). A constant width is added to the lines in one eye's view. With crossed fusion of the left pair, an oblique phantom occluding edge is perceived at the right edge of the lines slanted in near depth, and the lines appear flat. With crossed fusion of the right pair the lines appear at different slants, with no occluder. Labels are reversed for divergent fusion.

camouflage is prevented, as proposed by Assee and Qian (2007).

There is substantial evidence that edge-matching can influence the depth of intervening texture (McKee & Mitchison, 1988; McKee, Verghese, & Farell, 2004; Mitchison & McKee, 1987a, 1987b). For patterns with ambiguous matches such as sinusoidal gratings, the surface of the texture is perceived at the depth specified by the edge disparities. This is not the case for the stimuli we have investigated, in which the depth is perceived in a phantom contour, instead of in the texture that defines the edges. The presence of the monocular region in this case indicates a depth discontinuity, evidence that the binocular and monocular regions are not part of the same surface. The independence of depth processes for edges and texture could be ecologically useful for dealing with occlusion in natural viewing. Unmatched local regions of texture occur next to occluding surface edges; thus it is adaptive to have an edge-matching process that is independent of the process that assigns depth to the texture.

The contribution of occlusion geometry to the global resolution of monocular regions has been demonstrated in another context involving edge matching. As described earlier, Gillam and Grove (2004) demonstrated that phantom contours can be elicited by ambiguous horizontal disparity. Figure 11 shows a stereogram taken from their paper; in one eye's view the ends of the lines are vertically aligned on both sides, whereas in the other eye the line lengths are truncated along a diagonal on the nasal side. When fused, a slanted phantom occluder is perceived in depth and the lines appear flat. If the phantom occluder were produced purely by edge matching of the vertical edge in one eye and the diagonal edge in the other eye it should persist when the eyes are reversed. However, when the eye's views are reversed, no phantom occluder is perceived, and the individual lines appear at different slants as predicted by local horizontal disparity. This dissociation is important; it shows that the phantom occluder involves global resolution of the scene, which must be consistent with occlusion geometry, and thus it is not simply the result of edge matching. Seeing phantom occluders to account for certain patterns of binocular information raises the same issues as for other forms of subjective contours. It is unclear whether they are always present in particular contexts and only revealed when luminance contours are removed, or whether they are only "created" to account for the situation when luminance contours are absent.

A possible neural mechanism for da Vinci stereopsis has been identified in a class of V2 cells, which respond to disparity-defined edges (von der Heydt, Zhou, & Friedman, 2000). These cells are used as the basis of Assee and Qian's (2007) model of da Vinci stereopsis. Some of these cells are also orientation-tuned, or selective for the direction of the depth step. These results are of interest because this class of cells responds to the depth step in random-dot stereograms, in which a monocular region of texture is necessarily adjacent to the depth step. Thus it is unclear whether the cells are responding to the depth discontinuity given by the disparity change, the presence of the monocular region, or both. Gillam and Borsting (1988) suggested that monocular regions help locate disparity discontinuities, whereas most models view monocular regions as a *consequence* of disparity discontinuities. It would be interesting to know the extent to which monocular regions influence the cells' responses; hopefully this question will be addressed by future physiological research.

Summary and conclusions

We have demonstrated that a phantom occluding surface accounts for the presence of monocular regions on the nasal side of a binocular surface when alternative visual explanations are precluded. This provides experimental confirmation for Assee and Qian's (2007) suggestion that there is a valid interpretation of nasal monocular regions even when monocular camouflage is not possible, and consequently they should no longer be considered ecologically "invalid." Importantly, the perceived depth of the phantom surface was quantitative, and as precise as observers' depth matching for regular stereopsis defined by binocular disparity. Monocular regions on the opposite (temporal) side of the binocular surface were perceived behind the binocular region. In this case, there was no phantom surface and the monocular region of the texture was perceived in depth. Existing models of

stereo vision that include monocular regions are unable to account for the depth of the phantom occluder. We suggest that the visual system uses a combination of occlusion geometry and complex matching to locate edges in depth, and this may involve a form of double-matching in which one of the contours is not luminance-defined.

Keywords: da Vinci stereopsis, monocular regions, illusory contours, depth perception, binocular vision

Acknowledgments

This research was supported by Australian Research Council Discovery Grant DP110104810 awarded to Barbara Gillam.

Commercial relationships: none.
Corresponding author: Susan G. Wardle.
Email: s.wardle@unsw.edu.au.
Address: School of Psychology, The University of New South Wales, Sydney, Australia.

Footnote

[1] It has been claimed that the Nakayama & Shimojo (1990) stimulus involves a form of double-matching (see Assee & Qian, 2007; Gillam, Cook, & Blackburn, 2003).

References

Anderson, B. L. (1994). The role of partial occlusion in stereopsis. *Nature, 367,* 365–368.

Assee, A., & Qian, N. (2007). Solving da Vinci stereopsis with depth-edge-selective V2 cells. *Vision Research, 47,* 2585–2602.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Brooks, K. R., & Gillam, B. J. (2006). The swinging doors of perception: Stereomotion without binocular matching. *Journal of Vision, 6*(7):2, 685–695, http://www.journalofvision.org/content/6/7/2, doi: 10.1167/6.7.2. [PubMed] [Article]

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.

Da Vinci, L. (*c.* 1508). *Manuscript D. Bibliotheque de l'Institut de France*. Figure reprinted in Strong, D. (1979). *Leonardo on the Eye*. New York: Garland.

Gillam, B. (2011). The influence of monocular regions on the binocular perception of spatial layout. In L. R. Harris & M. R. M. Jenkin (Eds.), *Vision in 3D environments* (pp. 46–69). New York: Cambridge University Press.

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.

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

Gillam, B., Cook, M., & Blackburn, S. (2003). Monocular discs in the occlusion zones of binocular surfaces do not have quantitative depth—A comparison with Panum's limiting case. *Perception, 32,* 1009–1019.

Gillam, B., & Grove, P. M. (2004). Slant or occlusion: Global factors resolve stereoscopic ambiguity in sets of horizontal lines. *Vision Research, 44,* 2359–2566.

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

Gillam, B. J., & Blackburn, S. G. (1998). Surface separation decreases stereoscopic slant but a monocular aperture increases it. *Perception, 27,* 1267–1286.

Grossberg, S., & Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three-dimensional surface perception. *Vision Research, 43,* 801–829.

Grove, P. M., Gillam, B., & Ono, H. (2002). Content and context of monocular regions determine perceived depth in random dot, unpaired background and phantom stereograms. *Vision Research, 42,* 1859–1870.

Häkkinen, J., & Nyman, G. (1996). Depth asymmetry in da Vinci stereopsis. *Vision Research, 36,* 3815–3819.

Harris, J. M., & Wilcox, L. M. (2009). The role of monocularly visible regions in depth and surface perception. *Vision Research, 49,* 2666–2685.

Hayashi, R., Maeda, T., Shimojo, S., & Tachi, S. (2004). An integrative model of binocular vision: A stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry. *Vision Research, 44,* 2367–2380.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Kaye, M. (1978). Stereopsis without binocular correlation. *Vision Research, 18,* 1013–1022.

Kuroki, D., & Nakamizo, S. (2006). Depth scaling in phantom and monocular gap stereograms using absolute distance information. *Vision Research, 46,* 4206–4216.

Lawson, R. B., & Gulick, W. L. (1967). Stereopsis and anomalous contour. *Vision Research, 7,* 271–297.

Lee, A. B., Mumford, D., & Huang, J. (2001). Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision, 41,* 35–59.

Liu, L., Stevenson, S. B., & Schor, C. M. (1997). Binocular matching of dissimilar features in phantom stereopsis. *Vision Research, 37,* 633–644.

Liu, L., Stevenson, S. B., & Schor, C. M. (1994). Quantitative stereoscopic depth without binocular correspondence. *Nature, 367,* 66–69.

McKee, S. P., & Mitchison, G. J. (1988). The role of retinal correspondence in stereoscopic matching. *Vision Research, 28,* 1001–1012.

McKee, S. P., Verghese, P., & Farell, B. (2004). What is the depth of a sinusoidal grating? *Journal of Vision, 4(7)*:1, 524–538, http://www.journalofvision.org/content/4/7/1, doi:10.1167/4.7.1. [PubMed] [Article]

Mitchison, G. J., & McKee, S. P. (1987a). Interpolation and the detection of fine structure in stereoscopic matching. *Vision Research, 27,* 295–302.

Mitchison, G. J., & McKee, S. P. (1987b). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research, 27,* 285–294.

Mitsudo, H., Nakamizo, S., & Ono, H. (2005). Greater depth seen with phantom stereopsis is coded at the early stages of visual processing. *Vision Research, 45,* 1365–1374.

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

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

Ruderman, D. L. (1997). Origins of scaling in natural images. *Vision Research, 37,* 3385–3398.

Sachtler, W. L. B., & Gillam, B. (2007). The stereoscopic sliver: A comparison of duration thresholds for fully stereoscopic and unmatched versions. *Perception, 36,* 135–144.

Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2010). Monocular occlusions determine the perceived shape and depth of occluding surfaces. *Journal of Vision, 10*(6):11, 1–12, http://www.journalofvision.org/content/10/6/11, doi:10.1167/10.6.11. [PubMed] [Article]

von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research, 40,* 1955–1967.

von Szily, A. (1921). Translated in: W. H. Ehrenstein & B. J. Gillam (1998). Early demonstrations of subjective contours, amodal completion, and depth from half-occlusions: "Stereoscopic experiments with silhouettes" by Adolf von Szily (1921). *Perception, 27,* 1407–1416.

Wardle, S. G., Bex, P. J., Cass, J., & Alais, D. (2012). Stereoacuity in the periphery is limited by internal noise. *Journal of Vision, 12*(6):12, 1–12, http://www.journalofvision.org/content/12/6/12, doi:10.1167/12.6.12. [PubMed] [Article]

Watanabe, O., & Fukushima, K. (1999). Stereo algorithm that extracts a depth cue from interocularly unpaired points. *Neural Networks, 12,* 569–578.

Wheatstone, C. (1838). Contributions to the physiology of vision - Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London, 128,* 371–394.

Zannoli, M., & Mamassian, P. (2011). The role of transparency in da Vinci stereopsis. *Vision Research, 51,* 2186–2197.