

Amodal completion and the perception of depth without binocular correspondence

Benoit A Bacon, Pascal Mamassian[¶]

Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, Scotland, UK; e-mail: pascal@psy.gla.ac.uk

Received 15 September 2001, in revised form 27 May 2002, published online 16 August 2002

Abstract. Half-occlusions and illusory contours have recently been used to show that depth can be perceived in the absence of binocular correspondence and that there is more to stereopsis than solving the correspondence problem. In the present study we show a new way for depth to be assigned in the absence of binocular correspondence, namely amodal completion. Although an occluder removed all possibility of direct binocular matching, subjects consistently assigned the correct depth (convexity or concavity) to partially occluded ‘folded cards’ stimuli. Our results highlight the importance of more global, surface-based processes in stereopsis.

1 Introduction

The study of depth perception since Barlow et al (1967) and Julesz (1971) has largely focused on binocular disparities and the correspondence problem. Recent studies, however, have shown that there is more to stereopsis than simple point-by-point matching (Gillam and Borsting 1988; Nakayama and Shimojo 1990; Shimojo and Nakayama 1990, 1994; Anderson 1994, 1998, 1999; Anderson and Nakayama 1994; Liu et al 1994; Gillam et al 1999; Gillam and Nakayama 1999).

Da Vinci stereopsis (Nakayama and Shimojo 1990; Shimojo and Nakayama 1990, 1994) shows that monocularly presented (unpaired) objects can be perceived in depth, but only if their presence in one retinal image and absence in the other can be reconciled with the geometry of occluding surfaces. In the same vein, Gillam and Nakayama (1999) used the geometry of occluding surfaces to elicit the perception of ‘phantom’ surfaces in depth in the absence of binocular correspondence. In the latter study, illusory rectangles appear in depth when parts of the ‘further’ objects (vertical lines) are erased to suggest a ‘phantom’ occluder. This shows that visual (modal) completion of illusory contours can elicit a depth percept in the absence of binocular correspondence.

In the present paper we extend these findings and show a new way for depth to be assigned in the absence of binocular correspondence, namely amodal completion. When an object in the visual scene is partially occluded by another, the shape of its occluded parts and consequently its complete form, can be inferred through amodal completion (Kanizsa and Gerbino 1982; Pessoa et al 1998). This is achieved through surface (Nakayama et al 1995) or volume (Tse 1999) completion, a process in which an internal representation is constructed with or without an intermediate stage of contour completion (Kellman and Shipley 1991; Yin et al 1997, 2000; Tse 1999).

Since amodal completion depends upon constructed representations, and since it can be achieved monocularly, then direct point-by-point matching of corresponding features should not be necessary to perceive depth in amodally completed objects. Indeed, we show that amodal completion based on monocular information allows the visual system to use contours that are not physically visible but are *interpolated* to assign stereoscopic depth correctly.

[¶] Author to whom correspondence should be addressed.

2 Methods

Four participants, three males and one female, aged between 27 and 37 years old and employed at the University of Glasgow took part in the experiment. Two potential participants had to be rejected as they could not perceive depth in our baseline condition. Participant BB is one of the authors; the others were experienced psychophysical observers but were naïve as to rationale and aims of the experiment.

Simple stereograms, which can be seen in figure 1, were designed specifically for the experiment. They consist of a black-and-white configuration or ‘object’ and a grey configuration or ‘occluder’, both in front of a random-dot background.

The stimulus in figure 1a represents the baseline condition. When left unfused, each image can be interpreted either as three small white rectangles and three small black rectangles *amidst* grey bars or as two larger, juxtaposed surfaces of opposite

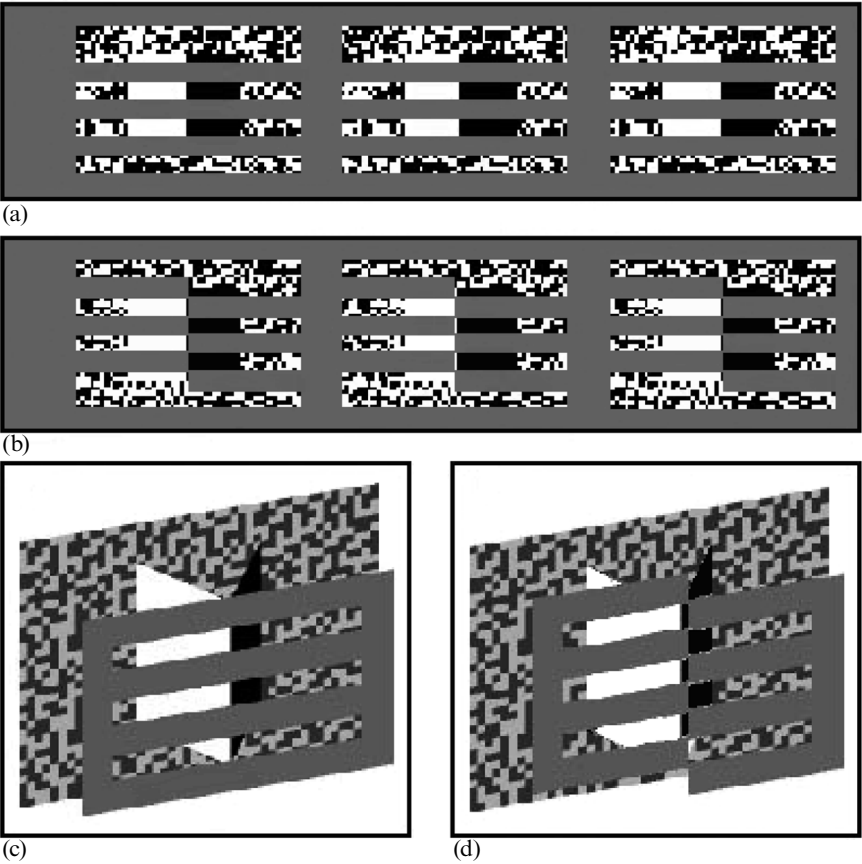


Figure 1. The stimuli created for the experiment. In both (a) and (b), divergent fusers should fuse the left pair of images and convergent fusers should fuse the right pair. (a) Stereograms of the stimulus when correspondence is available and the object is stereoscopically placed behind the occluder (baseline condition). In this particular pair, the edge where black and white meet was shifted by a cumulative crossed disparity of 3.85 arcmin, which creates a convex percept. (b) The exact same stimulus pair (identically shaped and positioned black-and-white object), this time presented behind the ‘crooked occluder’ (experimental condition). Notice how the central edge, the only portion of the image that carries convexity information, is now hidden in a mutually exclusive manner in the two eyes. Note also that the occluding bars are thicker than the space between them, so that no end-point matching is available. (c) Schematic representation of the percept elicited by the stimulus pair in (a). (d) Schematic representation of the percept expected from the stimulus pair in (b).

luminance *behind* grey bars. The latter, a more generic interpretation (Nakayama and Shimojo 1992) which depends upon amodal completion, is usually dominant.

In figure 1a, the black-and-white object is stereoscopically placed behind the grey occluder. The edge where black and white meet was shifted with crossed disparities to elicit the percept of a convex ‘folded card’, illustrated in figure 1c. If the edge was shifted with uncrossed disparities, it would then elicit the percept of a concave card.

The stimulus shown in figure 1b and schematised in figure 1d represents the experimental condition. The black-and-white object is identical in shape and proportion to the one in the baseline condition. The occluder, however, is different. Indeed, the grey horizontal bars have been split at their mid-point and then the two halves have been vertically shifted one relative to the other. It can be seen that this ‘crooked’ occluder hides the edge where black and white meet in a mutually exclusive manner in the two eyes. It therefore does not allow any point-by-point matching of the edge, the only portion of the black-and-white area that provides information about the convexity or concavity of the object. In other words, it removes binocular correspondence. It should be noted that the occluded segments were thicker than the visible segments, so that no end-point matching was possible.

In both the baseline and the experimental conditions, the black-and-white object was stereoscopically placed behind the grey occluder. We predicted that in these instances of ‘logical occlusion’ depth would be well perceived, even in the experimental condition when binocular correspondence was not available.

Two additional conditions were created to control for monocular or alternative strategies that the observers could use to solve the scene. In these conditions (condition 3, straight occluder; condition 4, crooked occluder), the black-and-white object was stereoscopically placed in front of the occluder. This put stereoscopic information and T-junction information in conflict and we predicted that this ‘illogical occlusion’ would severely impair performance, even with the straight occluder when binocular correspondence was available (Nakayama et al 1989, 1995).

All four conditions and the expected performance of the observers in each is summarised in table 1.

Table 1. The four tested conditions and the expected results.

Condition	Expected performance
1 With correspondence (straight occluder) Logical occlusion (object behind)	Baseline condition (as in figures 1a and 1c). Depth should be well perceived.
2 Without correspondence (crooked occluder) Logical occlusion (object behind)	Experimental condition (as in figures 1b and 1d). Can amodal completion allow the perception of depth without binocular correspondence?
3 With correspondence (straight occluder) Illogical occlusion (object in front)	Control condition. T-junction information is expected to prevail over stereo information. Performance should be poor.
4 Without correspondence (crooked occluder) Illogical occlusion (object in front)	Control condition. No depth should be perceived.

The observers sat 1 m from the monitor, placed their chin on a chin rest and viewed the stimulus pairs through a modified Wheatstone stereoscope. In each trial, they had to indicate whether the black-and-white object appeared to be convex or concave (2AFC) by pressing a key.

They were asked to complete four blocks of 288 trials, two blocks with the straight occluder (with correspondence) and two blocks with the crooked occluder

(without correspondence). Each block contained 144 'object stereoscopically behind the occluder' (logical occlusion) and 144 'object stereoscopically in front of the occluder' (illogical occlusion) trials, presented in random order. Each stimulus pair was presented for 500 ms (only slightly more noisy results were obtained at 250 ms).

The stimuli were created with the PsychToolbox (Brainard 1997; Pelli 1997) for MATLAB and were presented on a 21-inch monitor by an Apple G4 computer. The non-occluded (random-dots) surface subtended $3.08 \text{ deg} \times 2.02 \text{ deg}$ of visual angle. The occluding horizontal bars were 0.257 deg in height, which was larger than the area visible between them (0.224 deg).

The black-and-white area subtended $1.54 \text{ deg} \times 1.28 \text{ deg}$ of visual angle. It was presented with a crossed or an uncrossed disparity of 9.62 arcmin, which placed it 47 mm in front of or behind the grey occluder respectively.

In half the trials, the left part of the object was white, as in figure 1, while in the other half, the left part of the object was black.

Convexity or concavity of the black-and-white area was introduced by shifting the edge (where the black and the white surfaces meet) by 0.965, 1.93, 2.89, or 3.85 arcmin in opposite directions in each eye. These shifts created cumulative disparities of 1.93, 3.85, 5.78, or 7.70 arcmin and the binocular edge therefore appeared 9, 19, 28, or 37 mm in front of or behind the fixed sides of the black-and-white area.

The dependent variable, the 'proportion convex', refers to the proportion of times that the observer reported a percept of convexity, out of 32 repetitions of a given condition. The distributions of 'proportion convex' as a function of edge disparity (9 steps) were fitted with cumulative Gaussians. The slopes at the point of subjective equality and confidence intervals were obtained by Monte Carlo resampling (Wichmann and Hill 2001a, 2001b).

3 Results

Results for the four observers tested are shown in figure 2. All four observers excelled in the baseline condition, when binocular correspondence was present and the object was stereoscopically placed behind the occluder (left graphs, solid dots). In the experimental condition, when the object was still stereoscopically placed behind the occluder but the binocular correspondence was removed by the crooked occluder (right graphs, solid dots), the participants, especially BB, FG, and ML were still very efficient at discriminating between convexity and concavity. In other words, the participants could perceive depth without binocular correspondence even though no segments of the central edge offered the possibility of matching. This shows that the visual system can use contours that are not physically visible in the image but rather are interpolated, or completed, on the basis of monocular information to assign depth. The amodal completion of the black-and-white object allowed the perception of depth without binocular correspondence.

In the control conditions, when the object was stereoscopically placed in front of the occluder, thus creating conflict between disparity and T-junction information, performance was greatly impaired both with correspondence (left graphs, open dots) and without correspondence (right graphs, open dots).

In these two conditions, the first three observers could not report convexity or concavity accurately: BB was at chance level while FG and ML showed only residual discrimination accompanied by a bias for convexity. In other words, T-junction information preceded over disparity information, as was expected from previous reports (Nakayama et al 1989, 1995). Since the monocular information presented was identical, the negative results in the control conditions ensure that no monocular cues could have provided alternative strategies to solve the task.

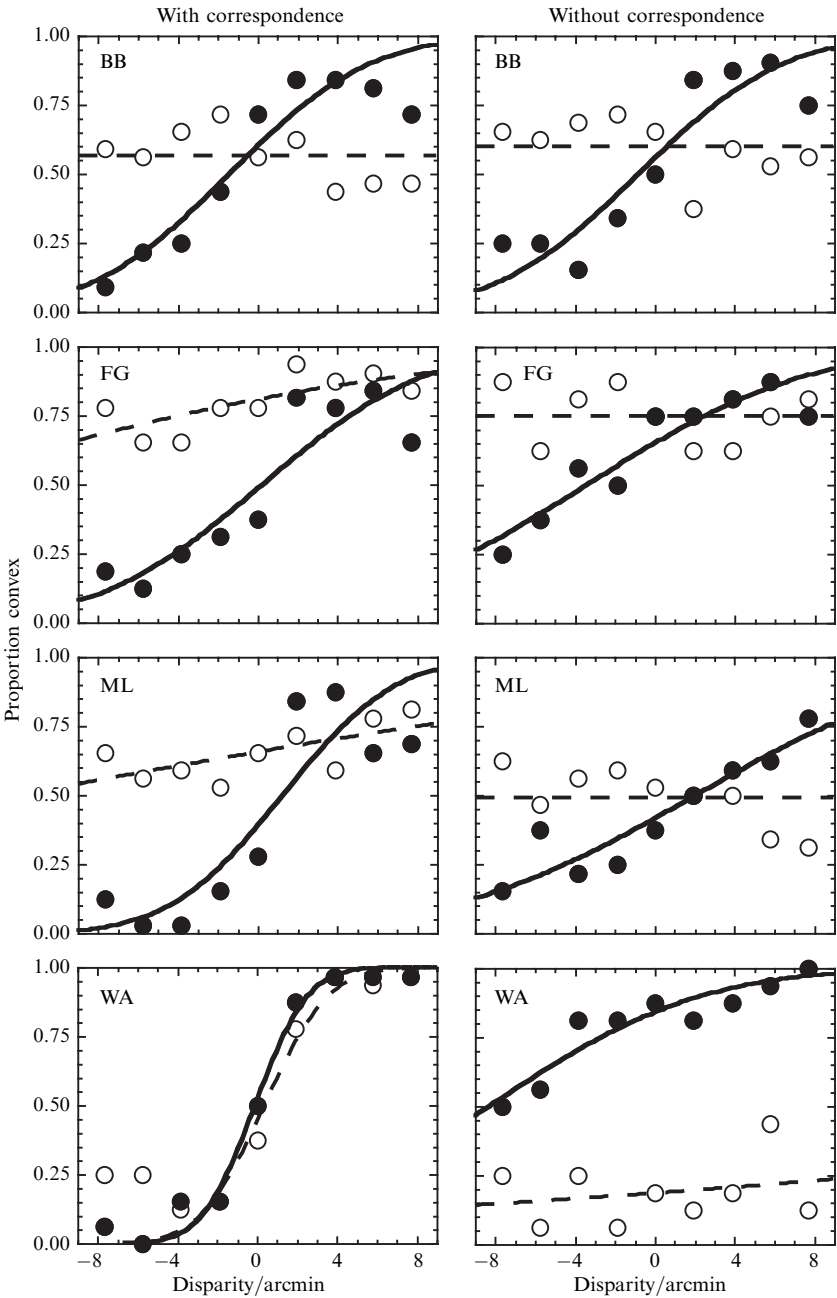


Figure 2. The results. Left graphs: with correspondence (straight occluder); right graphs: without correspondence (crooked occluder). Filled dots: object stereoscopically placed behind the occluder (logical occlusion); open dots: object stereoscopically placed in front of the occluder (illogical occlusion). All four observers perceived depth in the baseline condition (straight occluder, logical occlusion), as shown by slopes significantly different from zero (BB: 3.42 ± 0.354 ; FG: 3.26 ± 0.402 ; ML: 4.26 ± 0.450 ; WA: $7.65 \pm 0.816 \text{ deg}^{-1}$). All observers similarly perceived depth in the experimental condition (crooked occluder, logical occlusion), that is without correspondence, as shown by slopes significantly different from zero (BB: 3.43 ± 0.348 ; FG: 2.50 ± 0.384 ; ML: 2.36 ± 0.402 ; WA: $2.78 \pm 0.480 \text{ deg}^{-1}$). In the control conditions where the objects were placed stereoscopically in front of the occluders (illogical occlusions), the observers could not reliably assign depth (with the exception of participant WA, see a discussion of her results in the text).

The results of the fourth observer differed significantly from the others, especially when correspondence was available and the object was stereoscopically placed in front of the occluder (left graph, open dots). WA, a very experienced psychophysical observer with exceptional stereoacuity and vast experience of stereoscopic displays, showed an almost perfect performance in this condition. She seemed to value disparity cues far above T-junction information and was therefore able to report convexity or concavity accurately. She indeed reported verbally that she could clearly see three independent black-and-white wedges in front of the occluder, while the other subjects reported that the four conditions did not yield qualitatively different percepts.

WA also complained that, when correspondence was removed by the crooked occluder, she “didn’t know what to fuse” and that she believed she might have been “fusing the edge with the side of the occluder”. If such a strategy is adopted, the disparities thus created will always elicit a convex percept, as shown in figure 3.

Accordingly, WA shows a very strong bias for convexity in the experimental condition (without correspondence, logical occlusion) and reported that she perceived either “convexity or nothing”. She also reported that, when the percept was unclear, she pressed the concavity key to compensate, which explains her bias for concavity in the without correspondence, illogical occlusion condition.

In summary, the pattern of results for WA in the experimental condition demonstrates that relying on local disparity information by matching individual segments of the object without regard to the whole is misleading.

It should be noted that very similar results were obtained when the black-and-white area was horizontally translated to less central positions, or when the black and the white areas were made to be of unequal sizes. Results were also unaffected by relocation of the fixation point to different positions on the frontoparallel plane or along the line of sight.

A posteriori analysis revealed an interesting effect. In all four participants, convexity was most likely to be perceived when white was on the left and concavity was more likely to be perceived when black was on the left. The effect was significant in two participants (BB: $z = 4.36$; FG: $z = 3.42$, $p < 0.001$). This is consistent with demonstrations showing that observers assume by default that light is coming from above-left (Sun and Perona 1998; Mamassian and Goutcher 2001).

4 Discussion

The present study offers a clear demonstration that amodal completion allows the visual system to assign stereoscopic depth in the absence of binocular correspondence. More precisely, amodal completion of the occluded object, based on monocular information, allows the visual system to use contours which are not visible but are *interpolated* to assign stereoscopic depth. Indeed, without any possibility of direct binocular matching observers consistently perceived the appropriate 3-D configuration of amodally completed surfaces behind a solid occluder.

This supports the view that the visual system interpolates surfaces to construct a 3-D representation of an object (Nakayama et al 1995). It shows that this representation is precise and vivid enough to elicit correct depth percepts in the absence of direct binocular correspondence.

In da Vinci stereopsis (Nakayama and Shimojo 1990) the object is shown monocularly and its counterpart in the other eye is assumed by the visual system to be under the occluder, as close as possible to visibility. As such, the visual system can successfully solve the scene geometry and perceive depth by establishing correspondence between the seen object and the border of the occluder. In our demonstration, using this strategy is misleading, as shown in figure 3. The way in which corresponding elements of our stimuli are occluded in an alternative and mutually exclusive manner in the

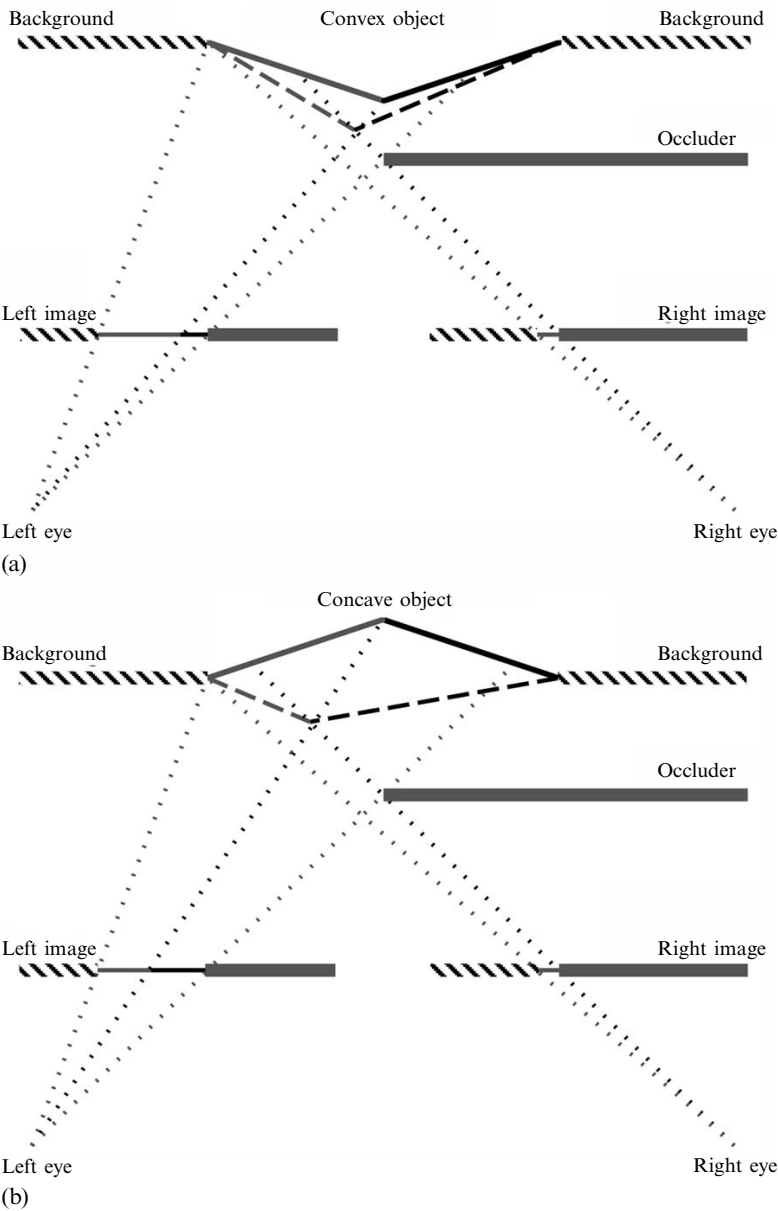


Figure 3. Schematic, top-view representation of the stimuli presented in the experimental condition (crooked occluder). This diagram shows that fusing the object’s black-and-white edge in one eye with the side of the occluder in the other eye gives disparity information that always leads to a convex percept, and that such a strategy is therefore misleading. In (a), the object presented (continuous grey-and-black folded card) is convex and fusing the edge in one eye with the side of the occluder in the other correctly elicits a convex percept (dashed grey-and-black folded card). In (b), the object presented is now concave but fusing the edge in one eye with the side of the occluder in the other eye will elicit a *convex* percept. This further supports our claim that the three observers that performed well in the experimental condition and did not report any qualitative difference in percepts between the baseline and experimental conditions based their judgments not on local cues but on the global perception of an amodally completed black-and-white surface.

two eyes forces the visual system to depend solely on interpolated contours or surfaces, which it is clearly able to do.

Gillam and Nakayama's (1999) demonstration showed 'phantom' occluders appearing in depth without binocular correspondence. The phantom occluder was created by the visual system to account for the presence of one feature in one eye and its absence in the other eye. Their demonstration therefore involves a binocular comparison of the information available and the creation of a surface as the only way to reconcile the two images. In contrast, our demonstration forces the visual system to process each image monocularly and to interpolate the missing features of the occluded object before any binocular processing.

Panum's (1858) famous limiting case is related to da Vinci stereopsis, to phantom stereopsis, and to the present demonstration. In short, presenting one thin vertical line in one eye and two in the other leads to the perception of two vertical bars, one closer to the observer than the other. The depth effect seems to be dependent on an internal reconstruction of the stimulus configuration that could produce the discrepant retinal images (see Gillam et al 1995). The same can be said of the experiment discussed here, but we stress once again that ours is the only instance where the stimuli have to be reconstructed monocularly (amodally completed) before successful matching can take place.

Our results are complementary to those of Häkkinen and Nyman (1997). They reported that the perception of slant due to horizontal magnification of one retinal image is diminished when an occlusion interpretation is possible. We, on the other hand, have shown that slant can be perceived even though the actual borders that should define horizontal magnification are occluded. They argue convincingly that "the matching of binocularly visible edges is not a sufficient process for determining the three-dimensional structure of the visual world" (page 38). We add that the occlusion of those edges and the elimination of point-by-point binocular matching does not in itself preclude the correct perception of the 3-D structure of the visual world. This is especially true if, as Häkkinen and Nyman (1997) said, local cues can be interpreted according to global context.

Verbal reports to the effect that the four conditions did not result in qualitatively different percepts and in a different subjective evaluation of performance are also of interest. This taps into recent research in visual awareness and consciousness, where a difference is observed between subjective appraisal and objective assessment of performance (Kolb and Braun 1995; Crick and Koch 1998). This implies that the interpolation of surfaces and the resulting 3-D representation can occur at a largely unconscious level if the stimuli are small, degraded, or brief enough. These phenomena concur with the idea that surface interpolation, like figure-ground processing, is a fast, first-stage, low-level process that is not dependent upon direct binocular matching.

Acknowledgments. We thank David Knill and Wendy Adams for their insight in the early stages of the experiment as well as Frederic Gosselin, Barton Anderson, and the anonymous reviewers for their comments on the manuscript. This study was supported by NSERC PDF-242082-2001 to BB and HFSP RG0109/1999-B to PM.

References

- Anderson B L, 1994 "The role of partial occlusion in stereopsis" *Nature* **367** 365–368
- Anderson B L, 1998 "Stereovision: beyond disparity computations" *Trends in Cognitive Sciences* **2** 214–222
- Anderson B L, 1999 "Stereoscopic surface perception" *Neuron* **24** 919–928
- Anderson B L, Nakayama K, 1994 "Toward a general theory of stereopsis: Binocular matching, occluding contours and fusion" *Psychological Review* **101** 414–445
- Brainard D H, 1997 "The Psychophysics Toolbox" *Spatial Vision* **10** 433–436
- Crick F, Koch C, 1988 "Consciousness and neuroscience" *Cerebral Cortex* **8** 97–107

-
- Gillam B, Borsting E, 1988 "The role of monocular regions in stereoscopic displays" *Perception* **17** 603–608
- 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, Blackburn S, Cook M, 1995 "Panum's limiting case: double fusion, convergence error, or 'da Vinci stereopsis'" *Perception* **24** 333–346
- 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
- Häkkinen J, Nyman G, 1997 "Occlusion constraints and stereoscopic slant" *Perception* **26** 29–38
- Kanizsa G, Gerbino W, 1982 "Amodal completion: Seeing or thinking", in *Organization and Representation in Perception* Ed. J Beck (Mahwah, NJ: Lawrence Erlbaum Associates) pp 167–190
- Kellman P, Shipley T, 1991 "A theory of object interpolation in object perception" *Cognitive Psychology* **23** 141–221
- Kolb F C, Braun J, 1995 "Blindsight in normal observers" *Nature* **377** 336–338
- Liu L, Stevenson S B, Schor C M, 1994 "Quantitative stereoscopic depth without binocular correspondence" *Nature* **367** 66–69
- Mamassian P, Goutcher R, 2001 "Prior knowledge on the illumination position" *Cognition* **81** B1–B9
- Nakayama K, He Z J, Shimojo S, 1995 "Visual surface representation: a critical link between lower-level and higher-level vision", in *Invitation to Cognitive Science* Eds S M Kosslyn, D N Osheron (Cambridge: MIT Press) pp 1–70
- Nakayama K, Shimojo S, 1990 "Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points" *Vision Research* **30** 1811–1825
- Nakayama K, Shimojo S, 1992 "Experiencing and perceiving visual surfaces" *Science* **257** 1357–1363
- Nakayama K, Shimojo S, Silverman G H, 1989 "Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects" *Perception* **18** 55–68
- Panum P L, 1858 "Physiologische Untersuchungen über das Sehen mit zwei Augen" (Kiel: Homman)
- Pessoa L, Thompson E, Noe A, 1998 "Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception" *Behavioral and Brain Sciences* **21** 723–802
- Pelli D G, 1997 "The VideoToolbox software for visual psychophysics: Transforming numbers into movies" *Spatial Vision* **10** 437–442
- Shimojo S, Nakayama K, 1990 "Real world occlusion constraints and binocular rivalry" *Vision Research* **30** 69–80
- Shimojo S, Nakayama K, 1994 "Interocularly unpaired zones escape local binocular matching" *Vision Research* **34** 1875–1881
- Sun J, Perona P, 1998 "Where is the sun?" *Nature Neuroscience* **1** 183–184
- Tse P U, 1999 "Volume completion" *Cognitive Psychology* **39** 37–68
- Yin C, Kellman P J, Shipley T F, 1997 "Surface completion complements boundary interpolation in the visual integration of partly occluded objects" *Perception* **26** 1459–1479
- Yin C, Kellman P J, Shipley T F, 2000 "Surface integration influences depth discrimination" *Vision Research* **40** 1969–1978
- Wichmann F A, Hill N J, 2001a "The psychometric function I: fitting, sampling and goodness-of-fit" *Perception & Psychophysics* **63** 1293–1313
- Wichmann F A, Hill N J, 2001b "The psychometric function II: Bootstrap-based confidence intervals and sampling" *Perception & Psychophysics* **63** 1314–1329

