Current Biology 21, 492–496, March 22, 2011 ©2011 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2011.02.011

Report

Non-Bayesian Contour Synthesis

Barton L. Anderson,[1,]* Judit O'Vari,[1] and Hilary Barth[2]
[1]University of Sydney, Department of Psychology, Sydney, NSW 2006, Australia
[2]Wesleyan University, Psychology Department, Middletown, CT 06459, USA

Summary

Recent research has witnessed an explosive increase in models that treat percepts as optimal probabilistic inference [1–13]. The ubiquity of partial camouflage and occlusion in natural scenes, and the demonstrated capacity of the visual system to synthesize coherent contours and surfaces from fragmented image data, has inspired numerous attempts to model visual interpolation processes as rational inference [9–13]. Here, we report striking new forms of visual interpolation that generate highly improbable percepts. We present motion displays depicting simple occlusion sequences that elicit vivid percepts of illusory contours (ICs) in displays for which they play no necessary explanatory role. These ICs define a second, redundant occluding surface, even though all of the image data can be fully explained by an occluding surface that is clearly visible. The formation of ICs in these images therefore entails an extraordinarily improbable co-occurrence of two occluding surfaces that arise from the same local occlusion events. The perceived strength of the ICs depends on simple low-level image properties, which suggests that they emerge as the outputs of mechanisms that automatically synthesize contours from the pattern of occlusion and disocclusion of local contour segments. These percepts challenge attempts to model visual interpolation as a form of rational inference and suggest the need to consider a broader space of computational problems and/or implementation level constraints to understand their genesis.

Results and Discussion

A fundamental challenge in recovering scene geometry involves overcoming conditions of occlusion and camouflage that obscure the boundaries of objects. Under conditions of camouflage, the visual system generates coherent representations of objects by synthesizing fragmented image data into illusory contours (ICs) [14–16], surfaces [9, 17, 18], and volumes [19]. The ubiquity and importance of these problems has inspired models that treat interpolation processes as forms of optimal probabilistic inference. One class of models has focused on understanding how local edge elements are linked to form coherent paths. These models treat contour synthesis as a form of optimal Bayesian inference driven by the statistics of local edge elements [10–12]. Although such approaches can account for a range of data involving contour interpolation, they do not account for the formation of illusory contours and surfaces that arise from the abrupt termination of thin contours [20–22]. The terminations of thin contours do not generate signals that can be used to define either the local orientations of an IC at the points of occlusion or its interpolation path. Hence, the ICs generated by the termination of thin contours cannot be understood with either the same mechanisms or image statistics that have been used to explain the contours formed by linking local edge elements into coherent paths.

Other attempts to model perception as ideal probabilistic inference have emphasized the role of global scene geometry and have argued that the formation of illusory contours, surfaces, and volumes are based on inferences derived from the *genericity* of different scene interpretations [9, 13, 23]. In such models, interpretations that entail accidental or coincidental alignments of the viewer [9, 23] and/or scene variables [13] are assigned lower probabilities than configurations that occur more generically (i.e., frequently). The resulting (or at least dominant) percept corresponds to the scene interpretation that occurs most generically.

These two general types of probabilistic models—one that operates in the local domain of oriented contour segments and the other that operates in the domain of global scene structure—both assert that completion phenomena can be understood as the most probable scene interpretation. Here, we report new forms of illusory contours and surfaces that are difficult to reconcile with either class of model.

Consider a motion sequence depicting the progressive occlusion (deletion) and disocclusion (accretion) of the central segment of a thin contour by a surface that has the same color as the contour's background (Figure 1A). In such conditions, the only information about the occluding surface is the pattern of accretion and deletion of the occluded contour. There is a virtually infinite family of occlusion events that could generate this image sequence: the occluding surface could originate from either side of the contour, and it could possess any number of different combinations of shape and speed profiles (Figures 1B and 1C). These ambiguities remain when four such elements are arranged to form a square (see Movie S1). When this image sequence is viewed, observers have difficulty determining whether the accretion and deletion of the contours arise from surfaces that originate from inside or outside of the square. When the occluding surfaces are colored differently from the background, the occlusion events are no longer ambiguous, and it is apparent that the pattern of accretion and deletion of the square figure is caused by four opaque disks that originate outside the square (Movie S2). Remarkably, however, this display evokes a very strong percept of an additional illusory surface that appears *inside* the square at exactly the same points of occlusion as the visible occluding discs. This percept instantiates one of the most improbable interpretations possible: two simultaneously present surfaces (one of which is plainly visible) that occlude and disocclude the same contours at exactly the same positions and times.

This illusory percept is extremely robust, and its perceptual vividness depends on simple low-level image properties. It can be observed over a broad range of contrasts, contrast polarities, speeds, and shapes of both the occluded and occluding

*Correspondence: barta@psych.usyd.edu.au

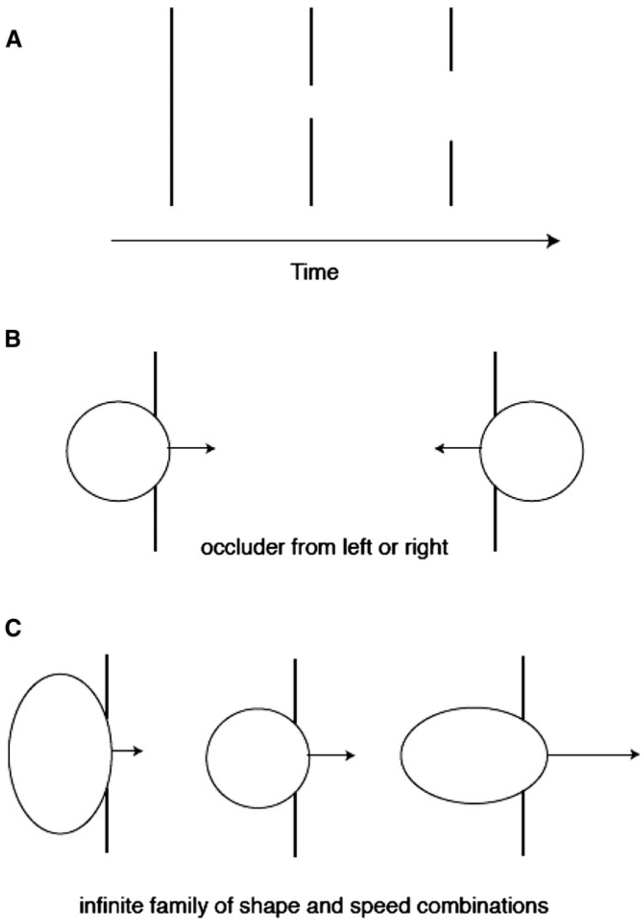

Figure 1. An Illustration of the Family of Ambiguities Generated by the Partial Occlusion of a Single Contour

(A) A time series representation of an occlusion event in which the central segment of a stationary contour is progressively occluded by a surface that is identical in color to the background.

(B) An illustration of two possible solutions to the occlusion event, which show that the direction of motion of the occluding surface is ambiguous. The shape of the occluding surface is shown in outline. The arrows indicate the direction of movement, and the length of the lines depicts the occluding surface's speed.

(C) An illustration showing how the same rates of occlusion can be generated by a variety of different shapes and speed profiles (only one possible direction of motion is shown).

figure (see, e.g., Movies S2–S4). We performed two pairs of experiments to evaluate the factors that modulate the vividness of the illusory figure (see Experimental Procedures). In the first pair of experiments, observers rated the vividness of the illusory surface by manipulating a visual analog scale while either the luminance of the occluding disk or the luminance of the occluded square was systematically varied (holding the luminance of the background and the other variable [disk or square luminance] constant). In the second pair of experiments, observers matched the perceived vividness of the illusory surface generated by a standard display and a comparable display that possessed the inverse luminance relationships (i.e., the standard display's photographic negative). Two matching experiments were performed. In one experiment, the square and background colors of both patterns were fixed, and the disk luminance of the test display was varied. Observers adjusted the luminance of the occluding disk of the match display until it elicited an illusory surface that was equal in perceived strength to the test display. In the other experiment, the disk and background colors of both patterns were fixed and the frame luminance of the test display was varied. Observers adjusted the luminance of the occluded square of the match display until it elicited an IC that was equal in perceived strength to the test display. The results of both sets of experiments reveal that the vividness of the IC depends on the relative edge contrasts of the occluded contour and the occluding discs: the vividness of the IC decreases monotonically as the edge contrast of the occluding contour is increased and increases monotonically as the edge contrast of the occluded contour is increased (Figures 2 and 3). These data suggest that the vividness of the IC is proportional to a relative measure of edge strength between the occluded contour and occluding surface.

The illusory surfaces that appear in these displays cannot be understood with either of the probabilistic models discussed above. Local edge interpolation models cannot explain these surfaces [10–12] because there are no oriented edge elements at the points of occlusion that are aligned with the direction that the illusory contours form. Indeed, any visible orientation at the contour terminators would have the orientation of the (visible) disks, which are approximately orthogonal to the local orientation of the illusory contours that form at the contour terminators. The illusory surfaces are also not consistent with theories that model percepts as the most generic scene interpretation [9, 13, 23] because the formation of the illusory surface entails positing two occluding surfaces to account for the occlusion of a single contour segment at precisely the same points in space and time.

For the displays presented herein, the main puzzle is to understand why the visual system constructs additional (illusory) surfaces when there are clearly visible occluding surfaces that account for all of the image data. Ecologically, the simultaneous occlusion of the contour terminators at the precisely same moments in space and time is extremely improbable. This suggests that these percepts either represent a nonoptimal output arising from the particular way that neural tissue implements occlusion computations, and/or that they reflect a compromise solution resulting from the need to solve multiple computational problems simultaneously.

There are two aspects of the illusory surfaces that must be explained by any model. First, the illusory figures are generated relatively locally and appear to emerge in response to the local accretion and deletion of the occluded contour segments. This can be observed by covering three of the four occluding disks in Movies S2–S4: a vivid illusory surface can be seen to appear along each side of the occluded square on the opposite side of the visible occluding disks. Second, the visibility of the illusory figure monotonically decreases as the strength (contrast) of the "true" occlusion interpretation increases. This suggests that the visible occluders suppress the formation of the illusory figure and that this suppression is proportional to the strength (contrast) of the visible occlusion signal. The challenge generated by these displays is to explain why this suppression exhibits a slowly saturating dependence on contrast that allows such an improbable percept to form. Some insight into this question can be gained by considering the broader perceptual consequences of the alternative, vigorous suppression that eliminates the alternative occlusion solution as soon as the occluding disks were perceptually visible. If the suppression we observed for the illusory figure arises from a competition between local units tuned to different directions of motion,

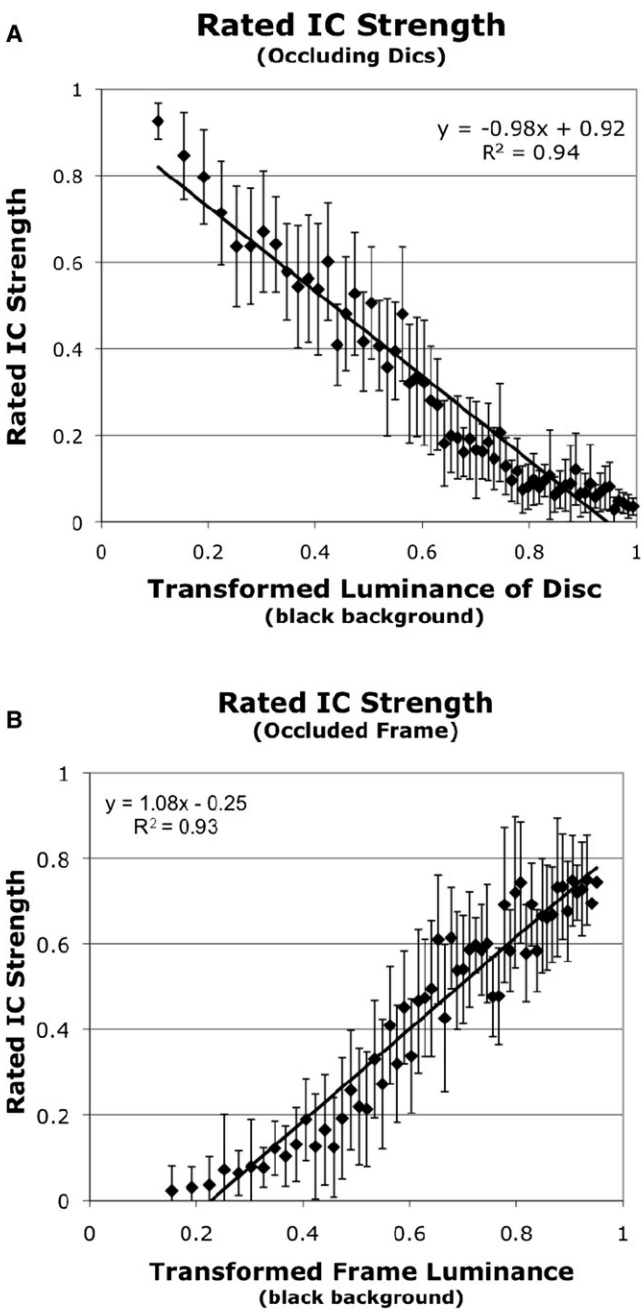

Figure 2. The Results of the Rating Experiments

(A) Observers rated the clarity of the illusory contour as a function of the luminance of the occluding disk, for fixed background and frame luminances.

(B) Observers rated the clarity of the illusory contour as a function of the luminance of the occluded frame for fixed background and disk luminances. The data reveal that the perceptual vividness of the illusory contour increases monotonically as the luminance of the occluding disk is decreased (A), or the luminance of the occluded frame is increased (B). The luminance values of the varied parameter were transformed according to a power law that maps linear luminance values into approximately perceptually equal intervals (see Methods). Data represent the average ratings for 3 observers, and the error bars represent 95% confidence intervals of the standard error of the mean.

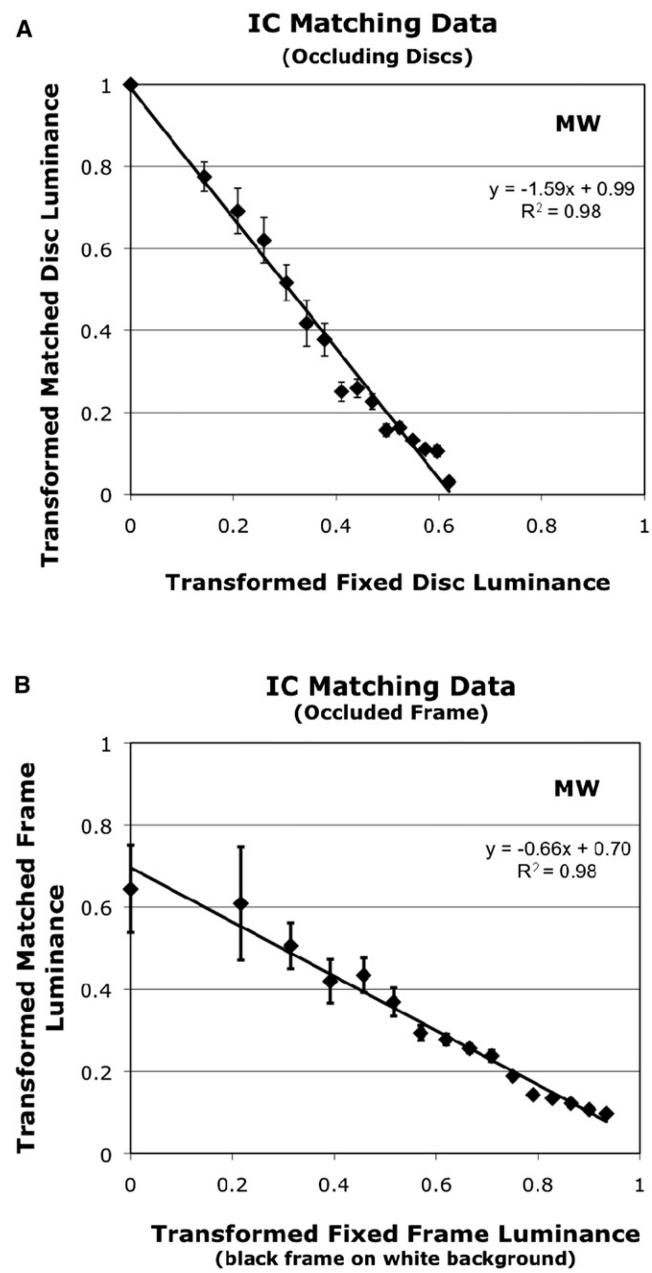

Figure 3. Results of the Matching Experiments

(A) Observers adjusted the luminance of the occluding disk of one illusory contour display (black background, white occluded frame) to match the perceived vividness of an illusory contour display of its photographic negative (white background, black occluding frame).

(B) Observers adjusted the luminance of the occluded frame of one illusory contour display (black background, light occluded frame) to match the perceived vividness of an illusory contour display of its photographic negative (white background, dark occluding frame). The results reveal that the strength of the illusory contour is a monotonically decreasing function of (A) the contrast of the occluding disk relative to the surround, and (B) and as a function of the contrast of the occluded frame. Results shown for one observer, other observers' data were very similar. The error bars are 95% confidence intervals of the mean.

vigorous suppression would favor a winner-take-all solution. Although this would lead to a more sensible interpretation of these motion displays, it could impair our ability to perceive transparent motion in a broad range of other visual contexts. There is now a plethora of data that documents our ability to perceive transparent motion, but this perceptual ability comes at a cost: the sensitivity for detecting a moving surface is lowered by the presence of a second moving (transparent) surface [24–26]. If this reduction in sensitivity arises from a competitive suppression between motion sensors tuned to different direction of motion, it could explain why our illusory figure is not perceptually eliminated until the contrast of the visible occluding figure is quite high.

The preceding analysis suggests that the highly nongeneric percepts observed in our displays may reflect a compromise solution between two distinct computational problems: the computation of occlusion relationships from the accretion and deletion of local contour segments and the need to represent transparent motion. The accretion and deletion of local contour segments is ambiguous, and the cause of the occlusion cannot be uniquely derived from either the local pattern of accretion or deletion or from the local orientation of the contour terminators. This ambiguity initiates the formation of two local occluding contours that have (approximately) orthogonal orientations and opposite directions of motion. These contour segments subsequently compete for dominance as a function of their visual strength (contrast) and direction of motion. If this interpretation is correct, then we would expect that the illusory figures reported herein exist because of the need to perceive motion transparency. This, in turn, suggests that these displays should evoke percepts of transparency. Although some observers do not spontaneously report a percept of transparency for the displays in Movies S1–S4, this percept can be readily evoked by translating four dots over the occluding disks in a manner consistent with the motion of the illusory figure inside of the square (see Movies S5–S7). Importantly, this display exhibits the same contrast dependence we observed in our original display: the illusory contour fades in strength as the contrast of the disks is increased or as the contrast of the occluded square is decreased. This suggests that the moving dots do not cause the formation of the illusory figure or the percept of transparency but rather merely reveal its presence.

Models that treat percepts as optimal rational inference are motivated by the recognition that our sensory experience must provide a pragmatically useful and biologically competitive representation of our environment. Such models begin by attempting to articulate the perceptual abilities that are critical to our survival, which are then analyzed to derive optimal solutions [1–13] utilizing probabilistic tools such as Bayes' theorem. Percepts are putatively explained as forms of rational inference when there is a close correspondence between optimal performance derived from computational considerations and observers' performance on a specified perceptual task. There are a number of problems that are raised by treating percepts as optimal Bayesian inference. First, these models assume that it is possible to understand vision as a set of circumscribed computational problems that were subjected to the pressures of natural selection to achieve optimal solutions. Even if this view is accepted, there is no a priori method for identifying the computational problems for which optimal solutions were selected; they can only be guessed and justified on the basis of their putative importance for our survival. Second, Bayesian models of perception are often deemed explanatory when they succeed in replicating a perceptual phenomenon or data set. It should be noted, however, that there is currently no method for identifying phenomena or data that can be taken as counterfactuals to Bayesian models. The set of Bayesian models is infinite; at most, only a particular combination of priors, likelihoods, and utility functions can be rejected. This renders the perception as Bayesian inference claim untestable at best, and to the extent that it is always possible to find some combination of priors, likelihood, and utility that can generate any data set, it becomes meaningless. Our description of our findings as non-Bayesian should therefore be understood as an assertion about claims that mechanisms underlying *contour synthesis* instantiate a process of ideal rational inference. We suggest that any attempt to model the phenomena reported herein as products of Bayesian inference will have to either consider a broader space of computational problems for which these percepts emerge as adaptive by-products and/or consider implementation level constraints that prevent the system from achieving optimal forms of inference in displays of this kind.

Experimental Procedures

Observers

Three experienced observers participated in the rating experiments, and two experienced observers participated in the matching experiments, one of which was one of the authors (J.O.).

Displays

The image sequences were presented on a CRT monitor (1600 × 1024 resolution, 75 Hz). The lookup table values were linearized and calibrated with a Minolta CS-100 chromameter. The luminance range of the display was 0 to 75.2 cd/m^2. The stimuli were viewed from a distance of 65 cm in a dark room for all experiments.

Two displays types were used. One type contained a stationary light diamond frame (a square rotated 45 degrees), which was occluded by four disks that oscillated over both the frame and a homogenous black background. (The displays were identical to those presented in the Movies S1–S7, but were rotated 45 degrees for purposes of subsequent experiments not reported here. This had no impact on the visibility of the illusory figures but was done for technical convenience). For the rating experiments, each side of the diamond was 5.7 cm long and had a line thickness of 0.5 mm. The circular occluding disks had a diameter of 6 cm. The disks on opposite sides of the diamond oscillated toward and away from each other in counter phase with a period of 5.33 s with a constant speed of 1.0 cm/sec. A small phase offset of 0.4 s between the two pairs of opposing disks was introduced that caused the illusory figure to appear to deform in shape. This phase offset had no discernible impact on the visibility of the illusory figure; it only changed its perceived shape. The second display type was identical to that described above but had the inverse luminance profile (a dark diamond frame, light discs, and a white background).

Rating Experiments

Two experiments were performed in which observers were instructed to rate the perceived strength of the illusory figure by manipulating a visual mark along a visual analog scale. In the first experiment, the occluding disk luminance was randomly varied between 0 cd/m^2 and 75.2 cd/m^2 in 1.18 cd/m^2 increments, and the diamond luminance was fixed at the highest luminance the display could output (75.2 cd/m^2). In the second experiment, the luminance of the occluded diamond was randomly varied between 0 and 75.2 cd/m^2 in 1.18 cd/m2 increments, whereas disk luminance was fixed at 8.8 cd/m^2.

Participants were instructed to rate the strength of the illusory contour by manipulating a rating scale at the bottom of the screen (a white line 19.45 cm long with ends labeled "min" [minimum] and "max" [maximum]). Observers used a computer mouse to slide a vertical red tick mark that was coupled to the position of the mouse and pressed a key when they were satisfied with their rating. Before the experiment began, participants were given three practice trials (black, mid-gray, and white disks for experiment 1, or black, mid-gray, and white diamonds for experiment 2) to familiarize them with the range of luminance values and illusory contour strengths they would encounter during the session. Each luminance condition was repeated 20 times for each observer in both experiments.

Matching Experiments

In the experiments 3 and 4, the display was divided into a left and right half. The left half of the screen was similar to the display used in experiments 1 and 2, whereas the right half of the display contained its photometric inverse (white background, light disks, and a dark occluded frame). The length of each side of the diamond was 6.6 cm long with a line thickness 0.5 mm, and disks were 6.2 cm in diameter. The four disks oscillated in the same way as before (pairwise 0.4 s out of phase) but with a period of 1.67 s and a constant speed of 3.1 cm/s. The luminance values of the background and diamond on the left side of the screen were 0 cd/m^2 and 35.4 cd/m^2, respectively, and the disks were randomly varied across trials between 0 and 37.1 cd/m^2 in 2.07 cd/m^2 increments. The luminance values of the background and diamond on the right side of the screen were 75.2 cd/m^2 and 0 cd/m^2, respectively. Participants adjusted the disk luminance on the right half of the screen with a mouse until the percepts of the illusory figures on both sides of the display were perceived to be equal in strength. For the experiment 4, the luminance values of the disks and background were fixed in both halves of the display, and the frame luminance was randomly varied in the match between 0 cd/m^2 and 75.2 cd/m^2 in increments of 4.4 cd/m^2. The luminance values of the background and disk were 0 cd/m^2 and 0.3 cd/m^2 for the left-hand display and 75.2 cd/m^{2}and 69.3 cd/m^2 for right-hand display. Participants adjusted the frame luminance on the right half of the screen using a mouse until the percepts of the illusory figures on both sides of the display were perceived to be equal in strength.

Data Analysis

The results of all of the reported experiments generated nonlinear monotonic functions of rated and matched illusion strength. We transformed all luminance values used in our experiments by normalizing all luminance values to the maximum luminance of the monitor (75.2 cd/m^2) and then transforming these data with a power law (exponent = .45, which is the inverse of the typical monitor gamma of 2.2). The power law transformation was done to transform the physically equal luminance increments used in our experiments into (approximately) equal perceived luminance differences.

Supplemental Information

Supplemental Information includes seven movies found with this article online at doi:10.1016/j.cub.2011.02.011.

Acknowledgments

We would like to thank Josh McDermott and Jon Winawer for helpful comments on an earlier version of this manuscript. This work was supported by a grant from the Australian Research Council to B. Anderson.

Received: January 10, 2011
Revised: February 9, 2011
Accepted: February 9, 2011
Published online: March 3, 2011

References

1. Maloney, L.T., and Zhang, H. (2010). Decision-theoretic models of visual perception and action. Vision Res. *50*, 2362–2374.
2. Maloney, L.T. (2002). Statistical decision theory and biological vision. In Perception and the Physical World: Psychological and Philosophical Issues in Perception, D. Heyer and R. Mausfeld, eds. (New York: Wiley), pp. 145–189.
3. Maloney, L.T., and Mamassian, P. (2009). Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. Vis. Neurosci. *26*, 147–155.
4. Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. Annu. Rev. Psychol. *55*, 271–304.
5. Geisler, W.S., and Kersten, D. (2002). Illusions, perception and Bayes. Nat. Neurosci. *5*, 508–510.
6. Yuille, A., and Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? Trends Cogn. Sci. (Regul. Ed.) *10*, 301–308.
7. Weiss, Y., Simoncelli, E.P., and Adelson, E.H. (2002). Motion illusions as optimal percepts. Nat. Neurosci. *5*, 598–604.
8. Feldman, J. (2009). Bayes and the simplicity principle in perception. Psychol. Rev. *116*, 875–887.
9. Nakayama, K., and Shimojo, S. (1992). Experiencing and perceiving visual surfaces. Science *257*, 1357–1363.
10. Geisler, W.S., Perry, J.S., Super, B.J., and Gallogly, D.P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. Vision Res. *41*, 711–724.
11. Geisler, W.S., and Perry, J.S. (2009). Contour statistics in natural images: Grouping across occlusions. Vis. Neurosci. *26*, 109–121.
12. Elder, J.H., and Goldberg, R.M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. J. Vis. *2*, 324–353.
13. Freeman, W.T. (1994). The generic viewpoint assumption in a framework for visual perception. Nature *368*, 542–545.
14. Kanizsa, G. (1979). Organization in Vision (New York: Praeger).
15. Petry, S., and Meyer, G. (1987). The Perception of Illusory Contours (New York: Springer-Verlag).
16. von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. Science *224*, 1260–1262.
17. Anderson, B.L., Singh, M., and Fleming, R.W. (2002). The interpolation of object and surface structure. Cognit. Psychol. *44*, 148–190.
18. Anderson, B.L. (1999). Stereoscopic surface perception. Neuron *24*, 919–928.
19. Tse, P.U. (1999b). Volume completion. Cognit. Psychol. *39*, 37–68.
20. Gillam, B. (1987). Perceptual grouping and subjective contours. In The Perception of Illusory Contours, S. Petry and G.E. Meyer, eds. (New York: Springer–Verlag), pp. 268–273.
21. Ehrenstein, W. (1941). Über Abwandlungen der L. Hermannschen *Helligkeitserscheinung*. Z. Psychol. Z. Angew. Psychol. *150*, 83–91.
22. Grossberg, S., and Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. Percept. Psychophys. *38*, 141–171.
23. Albert, M.K. (2001). Surface perception and the generic view principle. Trends Cogn. Sci. (Regul. Ed.) *5*, 197–203.
24. Lindsey, D.T., and Todd, J.T. (1998). Opponent motion interactions in the perception of transparent motion. Percept. Psychophys. *60*, 558–574.
25. Snowden, R.J., and Verstraten, F.A.J. (1999). Motion transparency: Making models of motion perception transparent. Trends Cogn. Sci. (Regul. Ed.) *3*, 369–377.
26. Curran, W., Hibbard, P.B., and Johnston, A. (2007). The visual processing of motion-defined transparency. Proc. Biol. Sci. *274*, 1049–1056.