Deriving behavioural receptive fields for visually completed contours

Jason M. Gold, Richard F. Murray, Patrick J. Bennett and Allison B. Sekuler

The visual system is constantly faced with the problem of identifying partially occluded objects from incomplete images cast on the retinae. Phenomenologically, the visual system seems to fill in missing information by interpolating illusory and occluded contours at points of occlusion, so that we perceive complete objects. Previous behavioural [1–7] and physiological [8–12] studies suggest that the visual system treats illusory and occluded contours like luminance-defined contours in many respects. None of these studies has, however, directly shown that illusory and occluded contours are actually used to perform perceptual tasks. Here, we use a response-classification technique [13–20] to answer this question directly. This technique provides pictorial representations – 'classification images' – that show which parts of a stimulus observers use to make perceptual decisions, effectively deriving behavioural receptive fields. Here we show that illusory and occluded contours appear in observers' classification images, providing the first direct evidence that observers use perceptually interpolated contours to recognize objects. These results offer a compelling demonstration of how visual processing acts on completed representations, and illustrate a powerful new technique for constraining models of visual completion.

Address: Department of Psychology, University of Toronto, 100 St George Street, Toronto, Ontario M5S 3G3, Canada.

Correspondence: Allison B. Sekuler
E-mail: sekuler@psych.utoronto.ca

Received: **16 March 2000**
Revised: **10 April 2000**
Accepted: **10 April 2000**

Published: **19 May 2000**

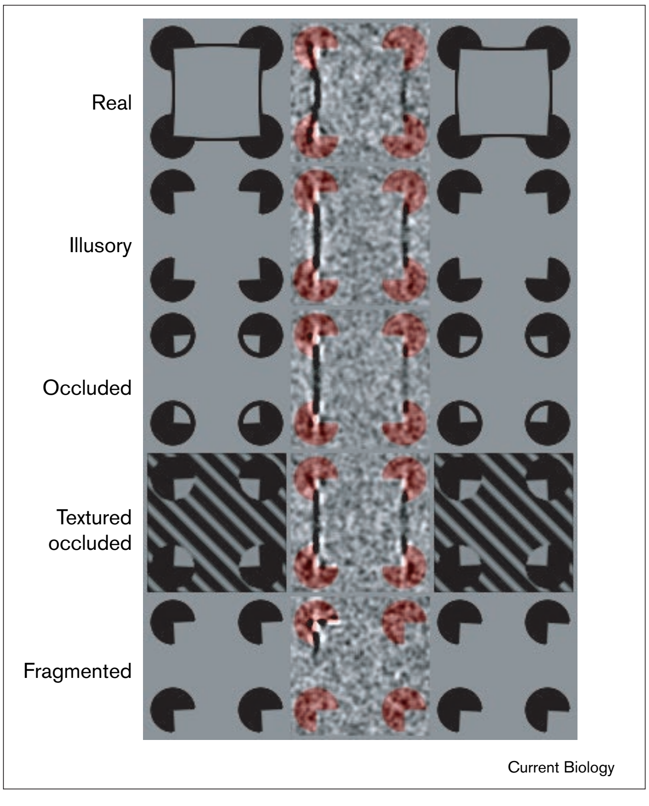
Current Biology 2000, 10:663–666

0960-9822/00/$ – see front matter

Results and discussion

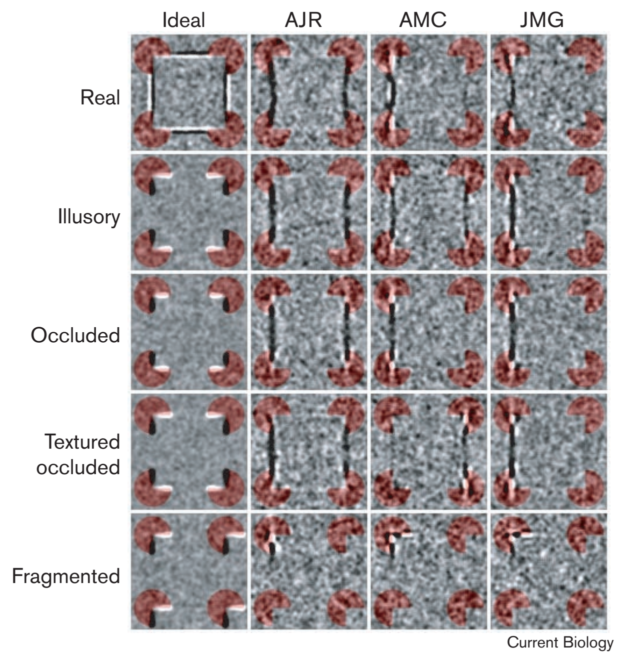
We used a shape discrimination task originally developed to measure spatial and temporal properties of illusory and occluded contours (sometimes referred to as 'modal' and 'amodal' contours, respectively) [4,5,21]. Three observers discriminated between 'fat' and 'thin' stimuli created by slightly rotating the inducers (that is, corners) of a Kanizsa square (Figure 1). Stimuli were observed in five conditions. Stimuli in the Real condition had thin parabolic contours interpolated between adjacent inducers (Figure 1a). Stimuli in the Illusory condition showed only the inducers, but between the inducers observers perceived illusory contours—subjective luminance edges in regions where luminance is physically uniform (Figure 1b). Stimuli in the Occluded condition were similar to those in the Illusory condition, but had a thin ring around the perimeter of each inducer (Figure 1c). This arrangement gave the appearance of a fat or thin square viewed through four holes in an occluding surface. Consequently, the square was defined in part by four occluded contours perceived as lying behind an opaque surface. Stimuli in the Textured Occluded condition were similar to those in the Occluded condition, but the inducers were embedded in a sine wave grating (Figure 1d). This strengthened the impression that the area around the inducers was an occluding surface. Stimuli

Figure 1



Stimuli and average classification images for each condition of the shape discrimination experiment. Each row corresponds to a different condition. The left and right columns show Thin and Fat stimuli, respectively. In the experiment, each corner inducer was rotated by ±1.75°. The inducers of the stimuli shown in the figure have been rotated by ±3.5° for clarity. The middle column shows smoothed average classification images, combining data from three observers. Red inducers have been superimposed on each classification image.

Figure 2



Smoothed classification images – behavioural receptive fields – for each observer in each condition of the experiment. Each row corresponds to a different condition and each column to a different observer. The first column shows the classification images for the ideal observer. The remaining columns show the classification images for three human observers.

in the Fragmented condition were similar to stimuli in the Illusory condition, but all the inducers faced in the same direction (Figure 1e). This arrangement prevented completion of the square, so observers did not perceive illusory or occluded contours (see Materials and methods).

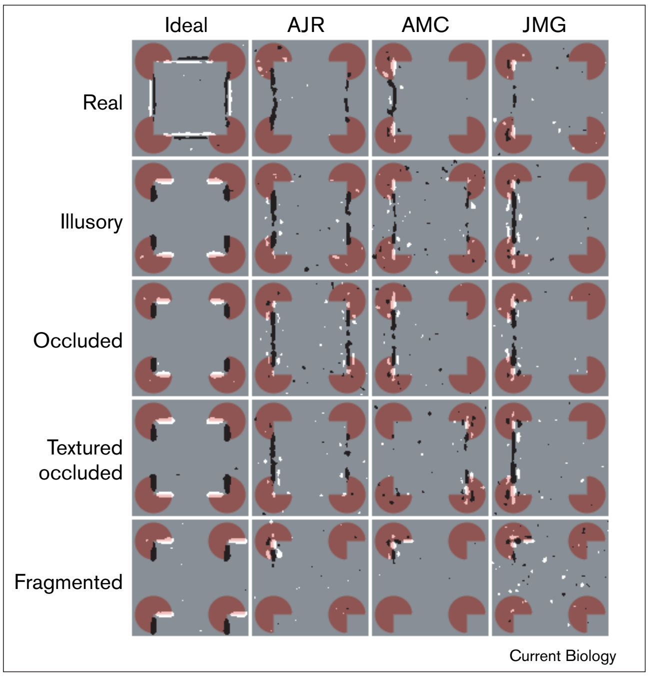
To determine the spatial locations that observers used to discriminate between fat and thin stimuli in each condition, we used a response-classification technique that has previously been applied to auditory [13,14] and visual [19,20] detection, vernier acuity [15,17], and letter discrimination [16,18]. The technique works as follows. Consider a task where observers must discriminate between two signals, S1 and S2. On each trial, either S1 or S2 is presented in luminance noise (resembling the 'snow' on a detuned television), and the observer's task is to state which of the two signals was presented. The signal contrast is adjusted across trials so that observers maintain a criterion level of performance (for example, 75% correct). On many trials, the noise will cause observers to make classification errors; on some trials, the noise is distributed in such a way as to make S1 look more like S2, or to make S2 look more like S1. To determine which image features make the observer more likely to respond 'S1' or 'S2', we find the correlation, across all trials, between the contrast at each pixel in the noise fields and the observer's responses. The resulting map is called a classification image, and it shows which image locations affected observers' responses. That is, it shows which image locations observers used to perform the task. In this sense the classification image is a behavioural receptive field (see Materials and methods for additional computational details). Unlike the stimuli used in past experiments that have measured classification images, patterns with illusory or occluded contours evoke percepts that have no corresponding physical attributes in the stimulus. Thus, if a classification image showed that observers' decisions had been affected by noise in locations where illusory or occluded contours were perceived, it would indicate that observers were using perceptually interpolated contours to perform the task. In other words, such a result would provide direct evidence that observers were basing their decisions on a perceptually completed representation of the stimulus, and it would specify the location of interpolated contours in that representation. To test these possibilities, we applied the response-classification technique to the fat/thin discrimination task.

The resulting classification images for all five conditions (Real, Illusory, Occluded, Textured Occluded, and Fragmented) for each observer are shown in Figures 2 and 3. Each image in Figure 2 is the result of 10,000 trials, and has been smoothed by computing a weighted average of adjacent pixels. The brightness of each pixel in Figure 2 indicates the correlation between noise contrast at that pixel and the observer's 'thin' response. Pixels that are brighter than mean grey indicate a positive correlation: positive-contrast noise (brighter than the background) in these locations biases observers to respond 'thin', and negative-contrast noise (darker than the background) biases observers to respond 'fat'. Conversely, pixels in Figure 2 that are darker than mean grey indicate a negative correlation. Figure 3 shows which pixels in Figure 2 are significantly different from zero ($p < 0.001$): pixels significantly greater than zero are white, and pixels significantly less than zero are black. Inducers, in red, have been superimposed on the classification images to provide positional reference frames. The leftmost column of Figures 2 and 3 shows classification images for the ideal observer (a hypothetical observer that uses a statistically optimal decision rule [22]) after a simulation of 10,000 trials. The ideal observer's classification images show where physical stimulus information resides for the fat/thin discrimination task in each condition. As one would expect, the ideal observer uses locations between inducers only in the Real condition. In the remaining conditions, the ideal observer only uses the edges of the inducers themselves. Locations between the inducers are not used in these conditions, because these locations are not informative. The remaining columns in Figures 2 and 3 show classification images for human observers in each condition, and the middle column in Figure 1 shows the average classification image across observers. Like the ideal observer, human observers used regions that correspond to physical luminance edges in the Real and Fragmented conditions (although human observers relied on only one or two luminance edges in both conditions). In the remaining conditions, however, human and ideal observers used very different locations. When making judgements about illusory and occluded contours, human observers used regions between adjacent inducers (where illusory and occluded contours subjectively appeared), even though these conditions contained no physical information in these locations.

The fact that each observer's classification images were nearly identical in the Real, Illusory, Occluded and Textured Occluded conditions is a powerful demonstration of the perceptual reality of illusory and occluded contours. Observers used illusory and occluded contours to perform the fat/thin discrimination task in the same way that they used luminance contours in the Real condition. This is especially surprising in the Occluded and Textured Occluded conditions, where observers perceived occluded boundaries as being hidden behind another surface. These results validate previous behavioral [1–7] and physiological [8–12] studies that have provided indirect evidence that observers use illusory and occluded contours to perform perceptual tasks. For instance, Ringach and Shapley [5] showed that observers are better at discriminating between fat and thin stimuli in the Illusory and Occluded conditions than in the Fragmented condition, and they argued that this is because observers use the illusory and occluded contours to discriminate between fat and thin stimuli, just as they would use luminance-defined contours. Another possibility, however, is that observers base their perceptual decisions solely on the luminance-defined edges of the inducers, but make more efficient use of these edges when they are perceptually organized into a single square. To address this possibility, Ringach and Shapley conducted a control experiment showing that performance worsens when the illusory contours are disrupted by a mask. However, this manipulation is difficult to interpret, as it may also change the perceptual organization of the stimulus. Our classification images directly show that even in the unmodified fat/thin stimulus, image locations at illusory and occluded contours affect observers' responses. This rules out the possibility that observers only use physical luminance edges, and provides convincing evidence that they use illusory and occluded contours constructed by the visual system.

These data also demonstrate how the response classification technique can be used to probe an observer's strategy in a perceptual task. The classification images in Figure 2 reveal that there are individual differences in the regions observers used to perform each task. For example, Figure 2 shows that observer AJR used both vertical edges of the stimulus in all but the Fragmented condition, whereas observers AMC and JMG tended to use only the left vertical edge. The classification images also reveal strategies common to all observers. For example, all observers relied almost exclusively on the vertical contours in the Real, Illusory, Occluded and Textured Occluded conditions (perhaps because the stimuli were described as 'fat' and 'thin' and not 'short' and 'tall'). Furthermore, in the Fragmented condition all the observers relied almost exclusively on the top-left inducer. Although it is not clear why our observers relied more on the left or top-left regions of the stimulus, this result may be related to other left and top-left biases found in native English readers (for example [23,24]). In any case, these examples demonstrate how the response classification technique can reveal differences in strategy across conditions and observers that cannot be detected by gross measures of performance (for example, percent correct).

Figure 3



The classification images from Figure 2 after being submitted to a statistical test. Each image shows the pixels that reached statistical significance ($p < 0.001$) in the corresponding classification image from Figure 2. Pixels significantly greater than zero are white; pixels significantly less than zero are black. The number of pixels that reached statistical significance greatly exceeded the number expected if observers' responses were uncorrelated with the noise. Each classification image consisted of 10,000 pixels, so with a criterion of $p < 0.001$, 10 pixels would be expected to reach significance by chance alone. The mean number of significant pixels across all human classification images was 327, and ranged from 139 pixels (observer AMC, fragmented) to 539 pixels (observer AJR, occluded).

Finally, these results provide the first objectively derived behavioural receptive fields of illusory and occluded contours. Because classification images reveal the shapes of interpolated contours, the response classification technique provides a method of testing spatio-temporal theories of completion. For example, we are currently using this technique to examine the claim that completion changes over time [2,5], and to determine the conditions under which local or global processes dominate completion [25,26].

Materials and methods

Stimuli

In all stimuli, the inducers were rotated 1.75° clockwise or counterclockwise. The radius of each inducer was 0.32° of visual angle, and the distance between the centres of adjacent inducers was 1.34°. The entire stimulus subtended a $2.0 \times 2.0°$ square, occupying 100×100 pixels on the display monitor. The background luminance was 27 cd/m^2, and the stimuli were negative in contrast (that is, darker than background luminance). The Gaussian noise was static and white (that is, the value at each pixel was sampled from an independent Gaussian distribution), and had a root-mean-square contrast of 14%. A new noise field was generated on each trial and the noise covered the entire stimulus.

Procedure

On each trial, a fat or a thin stimulus was randomly chosen and shown embedded in noise. Each trial consisted of an initial 500 msec fixation period, followed by the stimulus and noise for another 500 msec, and then a blank screen until the subject responded with a keypress. Auditory feedback was given in the form of a high- or low-pitched tone. The fixation point remained visible throughout the trial. We used QUEST [27] to vary the stimulus contrast to maintain 75% correct performance throughout each session. The signal contrast set by QUEST varied between 6% (Real condition) and 16% (Occluded and Fragmented conditions). The order of the conditions was randomized across observers. Each condition consisted of 10,000 trials, and was completed in five to eight sessions over the course of six to ten days. In each session, observers viewed stimuli from a single condition. At the beginning of each condition, each observer was familiarized with noise-free, high-contrast versions of the stimuli. Observers also received an initial 25 trials with high-contrast stimuli before beginning the experiment.

Classification images

In each of our tasks, there are two stimuli (S1, S2) and two responses (R1, R2), giving four possible stimulus–response categories for each trial: S1R1, S1R2, S2R1 and S2R2. It can be shown [17] that the correlation between the luminance at each pixel and the observer's response is found by averaging the noise fields in each stimulus–response category across trials, and computing the classification image, $C = (S1R1 + S2R1) - (S1R2 + S2R2)$. The smoothed classification images (Figures 1,2) were computed by convolving the classification image with a 5×5 convolution kernel (the outer product of $(1, 2, 3, 2, 1)^T$).

Acknowledgements

This research was funded by grants to P.J.B. and A.B.S. from the Natural Science and Engineering Research Council of Canada.

References

1. Paradiso MA, Shimojo S, Nakayama K: **Subjective contours, tilt aftereffects, and visual cortical organization.** *Vision Res* 1989, **29:**1205-1213.
2. Sekuler AB, Palmer SE: **Perception of partly occluded objects: a microgenetic analysis.** *J Exp Psych Gen* 1992, **121:**95-111.
3. Dresp B, Bonnet C: **Subthreshold summation with illusory contours.** *Vision Res* 1995, **35:**1071-1078.
4. Rubin N, Nakayama K, Shapley R: **Enhanced perception of illusory contours in the lower versus upper visual hemifields.** *Science* 1996, **271:**651-653.
5. Ringach DL, Shapley R: **Spatial and temporal properties of illusory contours and amodal boundary completion.** *Vision Res* 1996, **36:**3037-3050.
6. Gegenfurtner KR, Brown JE, Rieger J: **Interpolation processes in the perception of real and illusory contours.** *Perception* 1997, **26:**1445-1458.
7. Bockisch CJ: **Orientation discrimination and tilt aftereffects with luminance and illusory contours.** *Vision Res* 1999, **39:**765-776.
8. von der Heydt R, Peterhans E, Baumgartner G: **Illusory contours and cortical neuron responses.** *Science* 1984, **224:**1260-1262.
9. Grosof DH, Shapley RM, Hawken MJ: **Macaque V1 neurons can signal 'illusory' contours.** *Nature* 1993, **365:**550-552.
10. Sheth BR, Sharma J, Rao SC, Sur M: **Orientation maps of subjective contours in visual cortex.** *Science* 1996, **274:**2110-2115.
11. Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RB: **The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging.** *J Neurosci* 1999, **19:**8560-8572.
12. Sugita Y: **Grouping of image fragments in primary visual cortex.** *Nature* 1999, **401:**269-272.
13. Ahumada AJ, Lovell J: **Stimulus features in signal detection.** *J Acoust Soc Am* 1971, **49:**1751-1756.
14. Ahumada AJ, Marken R, Sandusky A: **Time and frequency analyses of auditory signal detection.** *J Acoust Soc Am* 1975, **57:**385-390.
15. Ahumada AJ: **Perceptual classification images from vernier acuity masked by noise.** *Perception* 1996, **25:**18.
16. Watson AB, Rosenholtz R: **A Rorschach test for visual classification strategies.** *Invest Ophthalmol Vis Sci* 1997, **38:**1.
17. Ahumada AJ, Beard BL: **Response classification images in vernier acuity.** *Invest Ophthalmol Vis Sci* 1998, **39:**1109.
18. Watson AB: **Multi-category classification: template models and classification images.** *Invest Ophthalmol Vis Sci* 1998, **39:** 912.
19. Abbey CK, Eckstein MP, Bochud FO: **Estimation of human-observer templates in two-alternative forced-choice experiments.** In *Proceedings of the Society of Photo-optical Instrumentation Engineers.* Edited by Krupinski EA. San Diego: SPIE; 1999:284-295.
20. Ahumada AJ, Beard BL: **Classification images for detection.** *Invest Ophthalmol Vis Sci* 1999, **40:**3015.
21. Rubin N, Nakayama K, Shapley R: **Abrupt learning and retinal size specificity in illusory-contour perception.** *Curr Biol* 1997, **7:**461-467.
22. Green DM, Swets JA: *Signal Detection Theory and Psychophysics.* New York: Wiley; 1966.
23. Zelinsky GJ: **Using eye saccades to assess the selectivity of search movements.** *Vision Res* 1996, **36:**2177-2187.
24. Eviatar Z: **Culture and Brain Organization.** *Brain Cogn* 2000, **42:**50-52.
25. Sekuler AB: **Local and global minima in visual completion: effects of symmetry and orientation.** *Perception* 1994, **23:**529-545.
26. Sekuler AB, Palmer SE, Flynn C: **Local and global processes in visual completion.** *Psychol Sci* 1994, **5:**260-267.
27. Watson AB, Pelli DG: **QUEST: a Bayesian adaptive psychometric method.** *Percept Psychophys* 1983, **33:**113-120.