The spatiotemporal properties of visual completion measured by response classification

Jason M. Gold   Department of Psychological & Brain Sciences, and Department of Cognitive Science, Indiana University, Bloomington, IN, USA

Erin Shubel   Department of Psychology, University of California, San Diego, CA, USA

A constant problem faced by the visual system is the identification of partly occluded objects within the visual scene. Recent experiments have demonstrated that the visual system engages in a process of visual completion, where the hidden parts of objects are filled into the visual representation. Recent experiments have also suggested that there may be a time course to this completion process. Here, we examined the spatiotemporal properties of visual completion by having observers classify figures defined by either luminance-defined or illusory contours and then correlating their decisions with externally added spatiotemporal visual noise. This "response classification" technique allowed us to derive a spatiotemporal correlation map (a "classification movie") that revealed the locations used by observers at each point in space and time during the stimulus presentation. We found that observers gradually became more influenced by noise at locations corresponding to illusory contours across the first 175 ms of stimulus presentation. Our results are consistent with the idea that there is a time course to the completion process on the order of ~175 ms.

Keywords: response classification, reverse correlation, visual completion

Introduction

One of the most fundamental problems faced by the visual system is the segregation of the visual scene into discrete surfaces and objects. The visual system receives patterns of light upon the retinae and must organize and segment this information into surfaces and objects during subsequent stages of processing. Surface and object boundaries typically correspond to physical discontinuities produced by changes in properties such as luminance and spectral content, and these discontinuities can be used by the visual system to segregate the visual scene into discrete parts. However, the visual system is also thought to fill in boundaries where parts of objects are either missing or obscured (e.g., Ringach & Shapley, 1996; Sekuler & Palmer, 1992). In the case of modal completion, the visual system is thought to interpolate illusory boundaries missing from an occluding surface (Kanizsa, 1979; Petry & Meyer, 1987). For example, in the outer columns of Figure 1B, most people see illusory contours forming a thin or fat "square" obscuring the inner quarters of four complete circles. In the case of amodal completion, the visual system is thought to interpolate occluded boundaries behind surfaces (Kanizsa, 1979).

Many of the earlier approaches to the study of illusory and occluded contour formation were subjective in nature and relied on observers to draw or describe what they saw or to rate the intensity of perceived contours (Kanizsa, 1979). Although these approaches are very effective in describing an observer's phenomenology, they also have several shortcomings. First, subjective report techniques assume that

observers are able to accurately report their percepts. However, this is often not the case (Nisbett & Wilson, 1977). Second, subjective techniques say little or nothing about whether observers actually use interpolated contours when discriminating among objects. To address these issues, several recent studies have focused on developing performance-based measures to determine whether observers actually rely upon occluded and illusory contours when performing various visual discrimination tasks.

For example, Sekuler and Palmer (1992) used a primed matching paradigm to determine whether occluded patterns prime an incomplete or complete representation of the stimulus. They found that when observers were asked to report whether two complete test patterns were the same or different, they responded "same" more quickly when they had been primed by an occluded version of the patterns rather than an incomplete version in which the occluder had been removed. More recently, Ringach and Shapley (1996) used a shape discrimination task in which they rotated the inducers (the corners) of Kanizsa squares to create "fat" and "thin" stimuli similar to those shown in Figure 1. They varied the rotation and the distance between inducers and found that discrimination performance was similar for figures defined by real, illusory, or occluded contours but was far worse for a fragmented version of the stimulus in which all of the inducers faced in the same direction.

Both of these studies offer evidence that the completed representations of illusory and occluded contours can affect performance in the same ways that complete objects can. However, in both of these experimental paradigms, properties of completed contours were inferred from their effects
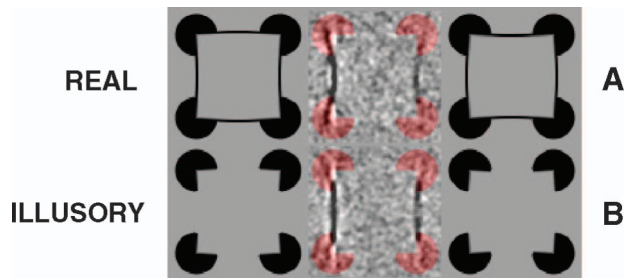
Figure 1. Examples of fat and thin Kanizsa square stimuli (outer columns) and average classification image results (center column) for the real and illusory conditions described in Gold et al. (2000).

on performance or their similarity to other stimuli. Although both support the idea that completion does occur and that it affects processing in various visual tasks, the techniques do not provide information about the spatial properties of the interpolated boundaries. An additional step toward revealing the underlying processes of visual completion is to measure the specific locations in images defined by either real or illusory contours that observers use to perform a task.

Gold, Murray, Bennett, and Sekuler (2000) used the response classification technique to provide a performance-based measure of the spatial properties of the completion process. Response classification (Ahumada, 2002; Ahumada & Lovell, 1971; Murray, Bennett, & Sekuler, 2002) typically involves presenting one of two images that have been corrupted by random luminance noise and having observers identify which of the two images was presented (although response classification can be used with many different tasks and numbers of stimuli). When the noise is high in contrast, it will cause the observer to make classification errors. By correlating the contrast of the noise at each pixel location with the observer's responses across many trials, it is possible to derive a correlation map (called a "classification image") that shows which areas of the stimulus the observer used to perform the task. Gold et al. adapted the fat/thin Kanizsa discrimination task used by Ringach and Shapley (1996), using a fixed rotation of the inducers but varying the contrast of the stimuli across trials. The left and right columns of Figure 1 show thin and fat stimuli, respectively, for two conditions of their experiment. In the real condition (Figure 1A), thin black lines between the inducers physically define the contours; in the illusory condition (Figure 1B), only the inducers appear. The center column of Figure 1 shows the classification images (averaged over all observers and smoothed by a small convolution kernel) for the real and illusory conditions obtained by Gold et al. The dark areas in the classification images represent a negative correlation between pixel contrast at that location and a "thin" response; the white areas represent a positive correlation between pixel contrast and a thin response (the red transparent inducers in the figure have been superimposed as landmark references). In other words, the classification images indicate the locations in the stimuli that observers were using when asked to discriminate among objects defined by real or illusory contours. Figure 1 shows that

observers used the regions in between the inducers to perform the task in both conditions, although there was no stimulus information present between the inducers in the illusory condition. These results show that observers use areas that correspond to perceptually filled-in regions of the stimulus when discriminating among patterns. The similarity between the classification images for illusory and luminance-defined contours also suggests that real and interpolated contours are treated similarly by the visual system. However, it is worth noting that, strictly speaking, a classification image does not necessarily reflect the properties of an observer's visual representation of a pattern. What it does reflect is the spatial strategy or "template" used by an observer to recognize a pattern or set of patterns. Using the properties of classification images to make inferences about the spatial characteristics of visual completion requires the assumption that an observer's template is matched to the visual representation of the patterns they are attempting to recognize.

Despite the similarities in the effects on performance and observers' spatial strategies in tasks involving illusory and luminance-defined contours, the spatiotemporal properties of real and illusory contours may be quite different. The visual system receives all contour information about a "real" object from the moment of presentation, whereas the physical representation of a partly occluded object that falls upon the retina is incomplete. At some point in time, the visual system forms a completed representation of the pattern that is then used in object discrimination and other tasks.

Several studies have attempted to measure the temporal properties of the completion process by using either subjective report methods or variants of the performance-based techniques described above. For example, Reynolds (1981) asked observers to report whether Kanizsa triangle figures followed by a set of masks appearing at the same locations as the inducers at various stimulus onset asynchronies (SOAs) appeared as a complete triangle figure defined by illusory contours or as a set of disconnected inducing elements. Reynolds found that the local masking elements disrupted the perception of illusory contours with SOAs below about 100 ms. Beyond this SOA, observers reported seeing a completed triangle.

A performance-based variant of this kind of approach was used by Sekuler and Palmer (1992) in a primed matching experiment involving occluded primes but with variable SOAs. As the duration of the priming stimulus increased, the effects of the occluded prime became more similar to the effects of the complete prime. At a prime duration of 400 ms, the reaction times for occluded and complete primes were virtually identical. By tracing the progression of priming over time, they concluded that the completion process takes between 200 and 400 ms.

Similarly, Ringach and Shapley (1996) adapted their 2AFC Kanizsa square discrimination task by adding two types of backward masking patterns: a "local" mask to determine the amount of time exposed to inducers that is needed to begin the completion process and a "global" mask to determine the amount of time needed to form a completed representation. By varying the duration of exposure to inducers before a

local mask appeared over the inducers, they found that ~117 ms is needed to elicit the interpolation process and affect discrimination performance. To interrupt further completion of the figure, they presented a later global mask at varying times after the local mask. They found that performance improved when the delay between the local and global masks was ~140–200 ms. From these results, they proposed that the completion process consists of two phases: an initial local feature detection stage that lasts ~117 ms and a second global feature integration stage that lasts an additional ~140–200 ms.

More recently, in an experiment designed to measure the effects of motion on the time course of visual completion, Murray, Sekuler, and Bennett (2001) asked observers to determine whether a complete, occluded, or fragmented rotating rectangle was longer vertically or horizontally. Vertical–horizontal discrimination performance was measured for a range of stimulus durations, and they found that discrimination performance for the occluded stimuli matched performance for the complete stimuli beyond ~75 ms, suggesting that observers required ~75 ms to form a completed representation of the occluded stimulus.

Each of these studies provides estimates of the time course of visual completion inferred by interrupting processing at various points in time and then comparing the effects on performance to the effects of "complete" stimuli. These techniques offer valuable assessments of the time required for visual completion to affect performance in visual discrimination tasks. But what are the spatial characteristics of the interpolated contours over the time it takes to generate a completed internal representation that then affects performance? The specific relationship between these two aspects of the completion process remains unclear. To measure this relationship, both spatial and temporal properties of interpolated contours would need to be simultaneously examined.

The purpose of this study was to use the response classification technique to obtain a more comprehensive measure of the spatiotemporal properties of visual completion. By using noise that varies not only over space but over time as well (dynamic noise), a classification image "movie" that reveals the locations that observers use to perform the task at each point in space and time can be created, that is, a spatiotemporal classification image (Neri & Heeger, 2002; Xing & Ahumada, 2002). If there is a time course to visual completion, we would expect to see an increase in the use of the locations that correspond to interpolated contours over time. Moreover, the classification movies would allow us to observe the spatiotemporal dynamics of the completion process.

Methods

Observers

Two men and one woman served as observers in the experiment. Their ages were 28, 31, and 27, respectively. One

of the men (M.S.) was new to psychophysical tasks, whereas the other two participants were experienced psychophysical observers. Both male observers were naive as to the purposes of the experiment. The third observer was an author and was thus aware of the experimental hypotheses. All observers had corrected-to-normal vision.

Apparatus

Stimuli were displayed on a Sony Trinitron Multiscan G520 CRT monitor that displayed 1024 × 768 pixels at a frame rate of 85 Hz. The display subtended a visual angle of 20.93 × 15.91 deg from the viewing distance of 100 cm. Luminance calibration was performed with a Minolta LS-100 photometer, and a 1779 element lookup table was generated from the calibration data to linearize the CRT (Tyler, Chan, Liu, McBride, & Kontsevich, 1992). The luminance values from this lookup table were used in constructing the stimuli on each trial. Display luminance ranged between 0.92 and 145.14 cd/m^2; average (background) luminance was defined as 49.64 cd/m^2. MATLAB (version 5.2.1) and the Psychophysics Toolbox (Brainard, 1997) were used to generate the stimuli and control the experiment.

Stimuli

Fat and thin Kanizsa figures (the "signals") were produced by rotating the corners of Kanizsa squares by ±1.75 deg (see outer columns of Figure 1). Each inducer spanned a radius of 0.34 deg of visual angle and the center of each was 1.36 deg from the center of adjacent inducers, giving a support ratio (the ratio of the diameter of a single inducer to the distance between the center of adjacent inducers) of 0.25. Signals were negative in contrast (i.e., darker than the background) on a uniform gray background (49 cd/m^2). In the illusory condition, only the inducers appeared; in the real condition, a thin black line connected the inducers. On each trial, 43 unique frames of high-contrast Gaussian white noise ($\mu = 0$, $\sigma = 25\%$ contrast) were generated. Each noise field subtended 2.13 × 2.13 deg (100 × 100 pixels), which covered the entire area of the signal.

Procedure

All sessions were conducted in a dark room. Viewing distance was stabilized with a chinrest. Each trial began with 1 s of fixation on a 2 × 2 pixel central fixation point. Next, a fat or thin Kanizsa square appeared in the center of the screen embedded in 43 successive frames of noise (approximately 12 ms/frame for a total of 508 ms). The final frame was followed by a 36-ms blank screen, after which high-contrast, noise-free versions of both possible signals and their corresponding keypress responses

appeared above and below fixation. This selection screen remained visible until the observer made a keypress response, after which the next trial automatically began. Auditory feedback was given after each trial in the form of a high (correct) or low (incorrect) beep. Signal contrast was varied across trials according to a staircase procedure to maintain approximately constant performance throughout the experiment (for details, see Threshold estimation section). The sequence of noise fields generated for each trial was saved for later analysis.

Each observer completed approximately 1,000 trials over the course of a 1-hr session. Observers completed between one and three sessions per day. The mean number of days each observer took to complete a condition was 26 days. A blocked design was used (i.e., observers participated in only one condition within a given session, and all sessions for a given condition were completed before switching to a new condition). Observers C.B. and M.S. completed the illusory condition before the real condition; E.S. completed the real before the illusory condition. Observers E.S. and M.S. completed 30,000 trials in each condition. C.B. completed 50,000 trials in the illusory condition and 30,000 trials in the real condition. Only the first 30,000 trials from the illusory condition are included in these analyses for purposes of comparison with the other observers and conditions.

Threshold estimation

A two-down, one-up adaptive staircase procedure was used to vary the signal contrast across trials. The staircase shifted through a series of logarithmically spaced signal contrast levels according to the observer's response accuracy. A Weibull function was fit to the data to estimate the contrast level that yielded 71% correct performance. Bootstrap simulations (Efron & Tibshirani, 1993) were used to produce confidence intervals for the threshold estimates (500 simulated experiments per threshold).

Ideal observer

For the task and stimuli used in this experiment, the ideal decision rule is to compare both of the noise-free signals with the noisy stimuli in each frame of a given trial and choose the signal that yields the higher overall cross-correlation (Green & Swets, 1966; Tjan, Braje, Legge, & Kersten, 1995). As the ideal observer is only constrained by the physical availability of information, a comparison of human to ideal performance gives a measure of the proportion of the information used by a human observer. A human observer's efficiency is the ratio of ideal to human contrast energy threshold, where contrast energy is the squared contrast integrated across the signal over space and time (Pelli, 1990; Tjan et al., 1995).

Ideal observer performance in the fat/thin Kanizsa task was estimated through Monte Carlo simulations (5,000 trials per condition). The ideal observer thresholds were calculated as described above for the human observers.

Results and discussion

Performance

Contrast energy thresholds and efficiencies (ideal/human energy thresholds) are plotted for each observer in each condition in Figure 2. For all observers, thresholds were
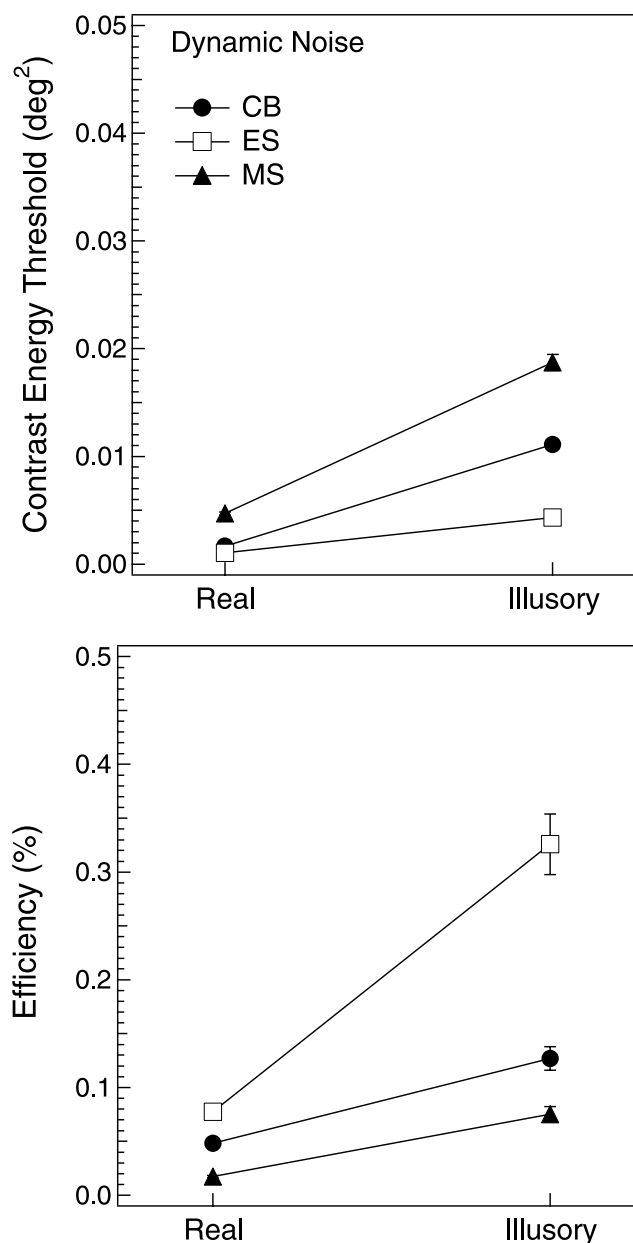

Figure 2. Thresholds (top panel) and efficiencies (bottom panel) for the fat/thin discrimination tasks in which the externally added noise was temporally dynamic. Error bars correspond to ±2 *SE*.

lowest in the real condition. However, efficiencies were actually greatest in the illusory condition. This surprising reversal between the real and illusory conditions nicely illustrates how comparing human to ideal performance unconfounds physical constraints on performance from psychological inefficiencies. One possible explanation as to why efficiencies were greater in the illusory than the real condition is that observers may have been using templates that were better matched to the signal in the locations where the signal was present in the illusory than the real condition. This possibility is addressed with the classification image analyses below. However, a second possible process that could explain the pattern of relative efficiencies across conditions is that there was greater temporal summation in the illusory condition than the real condition. With dynamic Gaussian noise, averaging across *N* frames (samples) of noise will reduce the expected variance of the averaged noise by a factor of *N*. Thus, temporal averaging in this task would effectively increase the signal-to-noise ratio, reducing the signal contrast necessary to obtain threshold performance. If the visual system integrates information across longer periods when identifying patterns defined by illusory contours, it would serve to boost efficiency in the illusory condition.

To test whether part of the difference in efficiencies across conditions was due to greater temporal summation in the illusory condition, we measured efficiencies for the same stimuli but in static rather than dynamic noise. All procedures and stimuli remained identical to those described in the main experiment, with the exception that one frame of noise (instead of 43 frames) remained on the screen for the duration of each trial. The same three observers completed 500 trials in each condition. If it is the case that there is greater temporal summation in the illusory condition, we would expect observers to benefit more from dynamic noise in the illusory condition than in the real condition. That is, the relative difference in thresholds between static and dynamic noise would be greater in the illusory condition than in the real condition.

The resulting contrast energy thresholds and efficiencies in static noise are plotted for each observer in each condition in Figure 3. The corresponding ratios of static to dynamic contrast energy thresholds are shown in Figure 4. All ratios are above unity because thresholds were higher with static noise in all conditions, indicating a general improvement in performance with dynamic noise. However, for all observers, the threshold ratio is larger for the illusory condition than for the real condition, indicating that observers did in fact benefit more from dynamic noise in the illusory condition.

These data suggest that temporal summation was greater for stimuli defined by illusory contours. One possible explanation for this is that the visual system integrates information over the initial completion period in the illusory condition but halts this summation process earlier when discriminating among objects defined by luminance contours, which offer more information from the point of stimulus
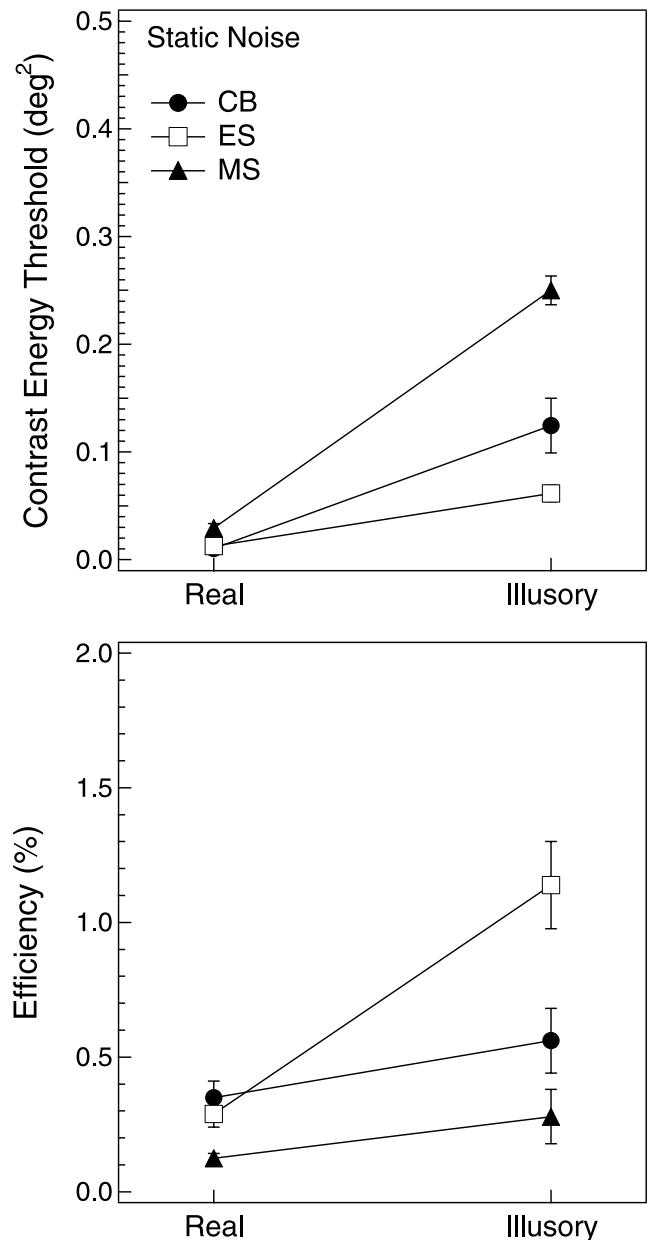

Figure 3. Thresholds (top panel) and efficiencies (bottom panel) for the fat/thin discrimination tasks in which the externally added noise was temporally static. Error bars correspond to ±2 *SE*.

onset. Further experiments are necessary to examine more closely the idea that the visual system integrates information across longer periods when identifying patterns defined by illusory contours.

Classification images

For each observer, the noise fields shown during the experiment were classified, averaged, and combined across trials according to the combination of signal and response on each trial to make a classification movie for each condition.
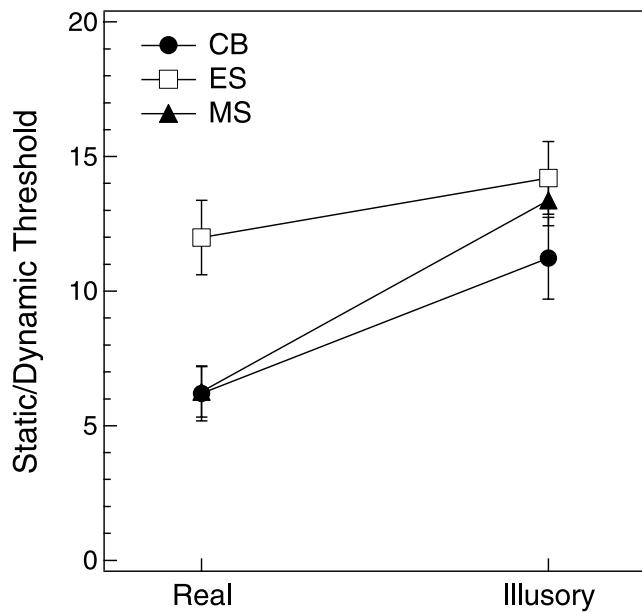
Figure 4. The ratio of static/dynamic thresholds for the fat/thin discrimination tasks. Error bars correspond to ±2 *SE*.

Specifically, for each trial, all 43 noise fields were classified according to the identity of the signal (S_{thin} or S_{fat}) and the response (R_{thin} or R_{fat}) made by the observer on that trial. The noise fields at each frame within each stimulus–

response bin were then averaged and combined according to Equation 1:

$$C = (\overline{S_{thin}\,R_{thin}} + \overline{S_{fat}\,R_{thin}}) - (\overline{S_{thin}\,R_{fat}} + \overline{S_{fat}\,R_{fat}}). \quad (1)$$

The resulting classification movie shows how the correlation between pixel contrast at each stimulus location and the observer's responses changed over the course of the 43 stimulus frames. If there is a time course for visual completion, we would expect to see features gradually emerge between the inducers in the illusory condition over time. The classification movies in the real condition serve as a control to test the possibility that any gradual changes seen in the illusory condition simply reflect the time course of normal visual information processing between the inducers.

Figure 5 is a QuickTime movie that summarizes the results of the response classification analysis for all three observers in both the illusory and real conditions. When the QuickTime movie is played, each panel shows an individual observer's spatiotemporal classification movie for a particular condition (indicated in the lower left corner below each panel). Each frame in the QuickTime movie advances through the spatiotemporal classification movies by ~12 ms (indicated in the lower right corner below each panel). For the purposes of visualization, the classification movies have been smoothed over space and time by a
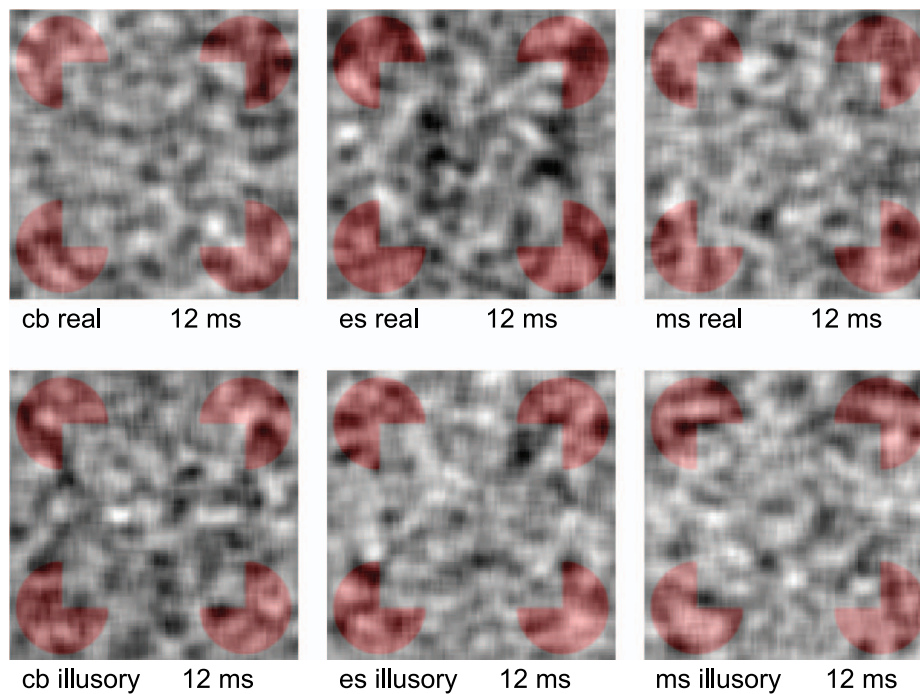
Figure 5. Classification movies for each observer in each condition. Click on the image to view the movies. Each panel corresponds to a single observer's spatiotemporal classification image for a particular condition (indicated in the bottom left corner of each panel). The bottom right corner of each panel shows the corresponding number of milliseconds that have passed during the stimulus presentation. Each movie has been convolved with a 5 × 5 × 5 spatiotemporal convolution kernel, and red inducers have been superimposed as signal landmarks.

linearly ramped $5 \times 5 \times 5$ pixel spatiotemporal convolution kernel. The brightness of each pixel in the classification movies represents the correlation between the noise contrast at that location in space and time and the observer's responses across trials. Specifically, the dark areas in the classification movies indicate a negative correlation between pixel contrast at that location and the observer's thin responses, and the white areas indicate a positive correlation between pixel contrast and the observer's thin responses. Unrotated red inducers have been superimposed on the classification movies to provide a reference to the general position where the signals appeared within the noise.

The classification movies shown in Figure 5 share several features with the images obtained with static noise by Gold et al. (2000) (see Figure 1). Most notably, with the exception of one observer (M.S.), contours appear in the space between adjacent inducers in the illusory as well as the real condition (although not at all points in time, as discussed below). Recall that there is no information about the signal in the areas between the inducers in the illusory condition, yet we see contours in these regions of the classification images, indicating that noise in these locations affected observers' responses. There are also similarities in observer strategies between the two studies, as indicated by the tendency of observers to rely on the vertical sides of the figure in both conditions.

Of greater interest here, though, are the temporal changes that can be seen in the classification movies. Although one observer (M.S.) did not show any influence of noise in the regions between the inducers during the entire stimulus duration (discussed in more detail below), the remaining two observers (C.B. and E.S.) showed a gradual increase over time in the influence of noise in the vertical regions between the inducers in the illusory condition. This gradual change takes place over the course of the first ~200 ms, a time period that is similar to the time course estimated by Ringach and Shapley (1996) using backward masking in a nearly identical task. In contrast, there is a far less gradual change in the influence of noise in the vertical regions between the inducers in the real condition during this time period for both of these observers.

We quantified the effect of time on the use of the spatial locations between the inducers for these two observers by measuring the cross-correlation between each frame of each classification movie and the parts of the template used by an ideal observer in the real condition that fall between the inducers. If observers made increasing use of the regions between the inducers over time, the correlation with the ideal template in these locations from the real condition should also increase accordingly.

For our task, the ideal template is simply the difference image produced by point-by-point subtracting the thin from the fat noise-free images. It can be shown that this difference image is also what would be produced by computing a classification image for a noiseless observer using an ideal decision rule after an infinite number of trials (Gold, Sekuler,

Figure 6. Illustration of the template used in the cross-correlation analysis (see text for details).

& Bennett, 2004). We multiplied this template by a masking cross pattern that preserved the values of the template between the inducers and set the remaining pixels to zero contrast. We then blurred the masked template with the same kernel used to smooth the human classification images to make the human and ideal templates comparable. The resulting template is shown in Figure 6. Finally, this template was cross-correlated (i.e., point-by-point multiplied and summed) with each frame of the real and the illusory smoothed classification movies for observers C.B. and E.S.

The results of this cross-correlation analysis are shown in Figure 7. The top panel of Figure 7 shows the results for observer C.B., and the bottom panel shows the results for observer E.S. Each graph plots the cross-correlation between the ideal template and the human classification movie as a function of time. The real condition is shown in the filled symbols and the illusory condition is shown in the open symbols. The error bars on each symbol correspond to ± 2 *SE*, estimated through bootstrap simulations (Efron & Tibshirani, 1993). Specifically, a series of 30,000 trial experiments were simulated for each observer's classification movie in each condition by randomly sampling (with replacement) from the trials used to generate each classification movie. Because of the large amount of time necessary to run these simulations, only 50 simulated experiments were conducted for each classification movie (this took approximately 12 days to complete running on four G4 PowerMac computers simultaneously). According to Efron and Tibshirani (1993), 50 simulated experiments are a sufficient number to obtain a reasonable estimate of standard error through bootstrapping methods.

These data show that there was a gradual increase in correlation for both observers in both conditions during the initial period of the stimulus presentation. However, both observers show a more gradual increase in correlation to an initial peak for the illusory condition. For both observers, the initial increase in correlation in the real condition peaked
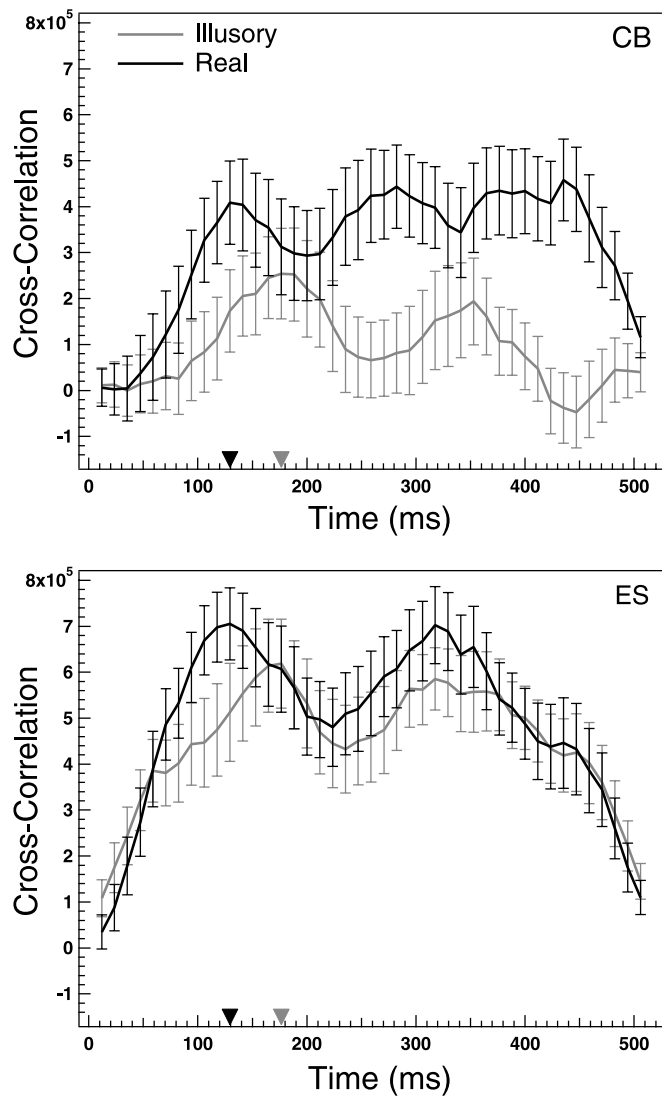
Figure 7. The results of the cross-correlation analysis for observers C.B. (top panel) and E.S. (bottom panel) in the fat/thin discrimination tasks. Each point corresponds to the correlation between the template shown in Figure 6 and each successive frame in the smoothed classification movies depicted in Figure 5. Error bars correspond to ±2 *SE*.

at ∼129 ms (indicated by the black triangle marker in each figure), whereas the initial increase in correlation for the illusory condition peaked later at ∼176 ms (indicated by the gray triangle marker in each figure). Although these results are consistent with the idea that observers were increasingly influenced by the regions in between the inducers in the illusory condition during the first ∼176 ms of the stimulus presentation, it is worth pointing out that similar results could also be obtained with a kind of process that does not involve a gradual change in visual processing over time. For example, a process that involves only the use of frames beyond ∼129 ms but was subject to a large amount of nonuniformly distributed (e.g., Gaussian) temporal uncertainty would predict a gradual increase in the influence of

noise during the first part of the classification movie. However, this kind of model would also be forced to claim that real and illusory contours are subject to different amounts of temporal uncertainty. This is because (a) the increase in the influence of noise between the inducers is more rapid and reaches ceiling earlier in the real than the illusory condition, and (b) both observers made use of the regions corresponding to the physical edges of the inducing elements in the illusory condition earlier in time than the regions corresponding to the locations of the illusory contours.

Although two of the observers showed the above pattern of results, the third observer (M.S.) appears to have adopted a very different strategy in this task. First, this observer used the horizontal edges of the top two inducers to perform the task, a strategy no other observer has adopted in this task when measured with the response classification technique. Second, this observer used only the physically defined edges of the inducers and not the regions in between the inducers to perform the task, thus showing no evidence of engaging in visual completion. Although it is not clear why this observer adopted a different strategy, the pattern of results may be consistent with recent work by Pillow and Rubin (2002). They argue that the preferential use of vertical contours exhibited by most observers is related to the fact that each vertical illusory contour in a centrally fixated Kanizsa square is processed by a single hemisphere of the brain, whereas each horizontal illusory contour is processed by both hemispheres of the brain. Pillow and Rubin hypothesize that difference in hemispheric processing results in stronger interpolation for vertical than horizontal contours because the completion process recruits only a single hemisphere to join vertical points of occlusion but must coordinate the activity between two hemispheres to join horizontal points of occlusion. Such an account would predict that an observer who used the horizontal edges of the occluders to perform the task would exhibit weaker interpolation and thus make less use of the horizontal regions in between the inducers. This is in fact what we see with observer M.S. However, this observer also makes little or no use of the regions in between inducers in the real condition, a result that is inconsistent with Pillow and Rubin's finding that real contour discrimination is not compromised by cross-hemispheric processing. Thus, it remains unclear whether the lack of evidence for interpolation exhibited by observer M.S. in the illusory condition is due to a process associated with visual completion.

Finally, it is worth noting that the regions between the inducers in C.B. and E.S.'s classification movies in the illusory condition appear to fade in relatively uniformly across the regions in between the inducers during the first ∼176 ms of the stimulus presentation. This result is consistent with a completion process that does not slowly propagate across space from the points of occlusion but rather fills in uniformly across the regions of interpolation. Of course, it still remains a possibility that our measures were simply too temporally coarse to reveal a highly rapid initial filling-in

process that involves propagation from the points of occlusion. Such a process would likely be difficult to measure using the response classification technique, as it has been shown that internal noise greatly reduces the signal-to-noise ratio in classification movies at high temporal frequencies (Xing & Ahumada, 2002). However, at the very least, our results are inconsistent with the existence of a slowly propagating interpolation process.

Conclusions

In the experiments described above, we measured spatio-temporal classification images for discriminating stimuli defined by either luminance-defined or perceptually interpolated contours. We found that the locations corresponding to perceptually filled-in contours had increasingly greater influence on observers' decisions over time, and that this process was more gradual than for luminance-defined contours. These results are consistent with the idea that the gradual changes between the inducers in the presence of perceptually filled-in contours do not simply reflect the time course of normal information acquisition but rather might reflect a time course associated with the completion process itself. Based on the temporal changes observed in the classification movies, we estimate the completion process to be on the order of ~176 ms in duration. An interesting and important question for future research is how dependent this estimate is on the particular properties of the stimuli (e.g., distance between inducers, amount of curvature in the interpolated contours, etc.).

Acknowledgments

Commercial relationships: none.
Corresponding author: Jason M. Gold.
Email: jgold@indiana.edu.
Address: Department of Psychological and Brain Sciences, 1101 East 10th Street, Indiana University, Bloomington, IN, 47405.

References

Ahumada, A. J., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision,* *2*(1), 121–131, http://journalofvision.org/2/1/8/, doi:10.1167/2.1.8. [PubMed] [Article]

Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America, 49,* 1751–1756.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap.* New York: Chapman & Hall.

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology, 10,* 663–666. [PubMed] [Article]

Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science, 28,* 167–207.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Kanizsa, G. (1979). *Oganization in vision.* New York: Praeger.

Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision, 2*(1), 79–104, http://journalofvision.org/2/1/6/, doi:10.1167/2.1.6. [PubMed] [Article]

Murray, R. F., Sekuler, A. B., & Bennett, P. J. (2001). Time course of amodal completion revealed by a shape discrimination task. *Psychonomic Bulletin & Review, 8,* 713–720. [PubMed]

Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience, 5,* 812–816. [PubMed] [Article]

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3–24). Cambridge, MA: Cambridge University Press.

Petry, S. J., & Meyer, G. E. (1987). *The perception of illusory contours.* New York: Springer-Verlag.

Pillow, J., & Rubin, N. (2002). Perceptual completion across the vertical meridian and the role of early visual cortex. *Neuron, 33,* 805–813. [PubMed] [Article]

Reynolds, R. I. (1981). Perception of an illusory contour as a function of processing time. *Perception, 10,* 107–115. [PubMed]

Ringach, D. L., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal

boundary completion. *Vision Research, 36,* 3037–3050. [PubMed]

Sekuler, A. B., & Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General, 121,* 95–111.

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research, 35,* 3053–3069. [PubMed]

Tyler, C. W., Chan, H., Liu, L., McBride, B., & Kontsevich, L. (1992). Bit-stealing: How to get 1786 or more grey levels from an 8-bit color monitor. In B. E. Rogowitz (Ed.), *Human vision, visual processing, and digital display III* (pp. 351–364). Bellington, WA: SPIE.

Xing, J., & Ahumada, A. J. (2002). Estimation of human-observer templates in temporal-varying noise [Abstract]. *Journal of Vision, 2*(7), 343a, http://journalofvision.org/2/7/343/, doi:10.1167/2.7.343.