The Time Course Of Visual Completion Measured By Response Classification

Erin Shubel & Jason M. Gold *Indiana University, Bloomington*

1 BACKGROUND

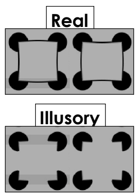
Several studies have provided evidence that people will rely upon both occluded & illusory contours when performing various visual discrimination tasks (e.g., Murray et al., 2001; Ringach & Shapley, 1996; Sekuler & Palmer, 1992). Results from these studies have also suggested that it takes the visual system a certain amount of time to build up a completed representation.

Recent studies have used a technique called *response classification* to pinpoint the locations in images that people are using when they are asked to discriminate amongst objects defined by either illusory or occluded contours (Gold et al., 2000; Murray, 2002). Here, we used this technique with dynamic noise to examine the spatiotemporal aspects of the completion process.

2 METHODS

Observers
- Two males & one female
- Ages 28, 31, & 27, respectively
- Two were naïve

Stimuli
- "Fat" & "thin" Kanizsa figures
- Inducers rotated by ± 1.75°
- Inducer radius: 16 pixels (0.34°)
- Distance between centers of adjacent inducers: 64 pixels (1.36°)
- Support ratio: 0.25



- Signals were negative in contrast.
- Background was uniform grey (49 cd/m²).

- 30,000 trials/condition/observer
- Blocked by condition (order was counterbalanced)

Noise Fields
- High contrast Gaussian white noise fields
  σ = 25% contrast
- 100 x 100 pixels (2.13° x 2.13°)
- 43 unique frames/trial
- Duration: ~12 ms/frame
  Total duration: ~508 ms/trial
  (85 Hz monitor)

Staircase
Signal contrast was varied across trials with an adaptive staircase to maintain 71% correct performance throughout the experiment.

Ideal Observer
Ideal observer performance was measured in each condition by Monte Carlo simulations (10,000 trials/condition). The ideal rule for these tasks & stimuli is to maximize the cross-correlation between the signal + noise combination & the noise-free signals (Green & Swets, 1966; Tjan et al., 1995).

3 PERFORMANCE

Contrast energy thresholds & efficiencies are plotted for each observer in each condition in Figures 1 & 2, respectively.

Efficiency is defined as the ratio of the ideal to human contrast energy thresholds in each condition.

Figure 1.

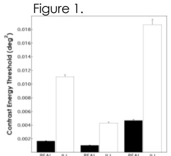
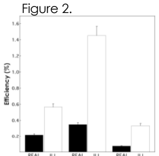


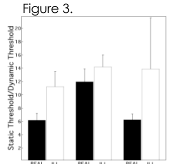
Figure 2.



Note that, for all observers, efficiency was greater in the Illusory condition. To test whether part of this difference was due to greater temporal summation in the Illusory condition, we measured efficiency for the same stimuli in static noise. If so, we would expect observers to benefit more from dynamic noise in the Illusory condition than in the Real condition. The ratio of static to dynamic contrast energy thresholds for each condition & observer is shown in Figure 3.

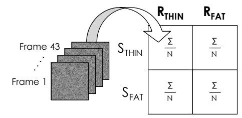
These data show that observers did in fact benefit more from dynamic noise in the Illusory condition, indicating temporal summation was greater for these stimuli.

Figure 3.



4 CLASSIFICATION IMAGES

Noise fields were classified, averaged, & combined in order to make a single classification movie for each observer. Specifically, for each trial, all 43 noise fields were classified according to the response made by the observer on that trial. After the noise fields from all the trials were classified, the noise fields at each frame within each stimulus-response bin were averaged & the averages were combined according to the following formula:

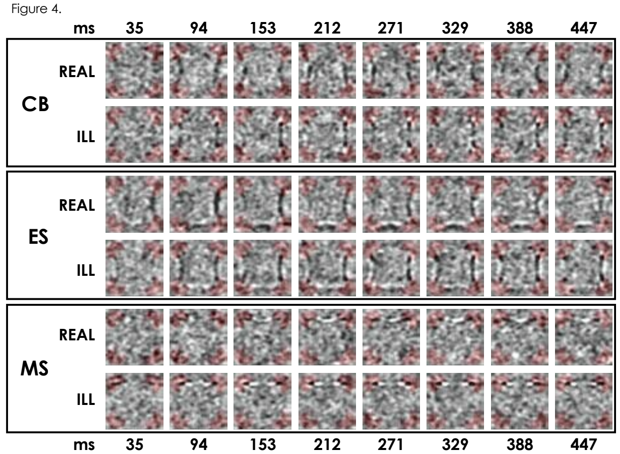


$$C = (\overline{S_{THIN}\mathbf{R_{THIN}}} + \overline{S_{FAT}\mathbf{R_{THIN}}}) - (\overline{S_{THIN}\mathbf{R_{FAT}}} + \overline{S_{FAT}\mathbf{R_{FAT}}})$$

The resulting classification movie shows how the correlation between pixel contrast & 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. The classification images 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 processing of real contours.

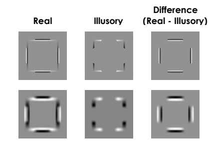
For purposes of visualization, the classification images were smoothed over space & time by a spatiotemporal convolution kernel (9x9x9). Samples of the classification movie taken at 59 ms intervals are shown in Figure 4 for each observer in each condition.

Figure 4.



5 CORRELATION ANALYSIS

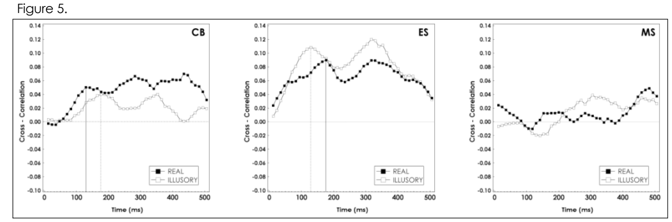
To quantify the changes in the classification images, we compared the human observers' classification movies with a classification image derived from the ideal observer's classification templates in the Illusory & Real conditions. Ideal observer classification templates (shown at right) were generated for both the Real & Illusory conditions. The top figure shows the raw classification templates, & the bottom figure shows the templates after being smoothed by the same spatiotemporal convolution kernel as was used with the human classification images. These templates indicate the regions of the stimuli that contain task information for each task. In the Real condition, there is information around the entire perimeter of the square. In the Illusory condition, there is only information at the edges of the inducers. By subtracting the Illusory template from the Real template to form a *Difference* template, we are left with only the regions between the inducers -- those representing the locations corresponding to the illusory & real contours.



For each observer & each condition, every frame in the human classification movie was cross-correlated with the Difference template. Given the temporal changes that we observed in the classification images, we would expect to find a more gradual increase in correlation for the Illusory than for the Real condition for observers CB & ES.

Figure 5 shows the correlations between the human classification images & the Difference template at each frame in time.

Figure 5.



One observer (CB) shows the predicted effect: the initial peak in correlation with the Difference template is around 120 ms in the Real condition, whereas the correlation increases until around 200 ms in the Illusory condition. Surprisingly, the second observer (ES) shows the opposite effect -- the Illusory correlation peaks first, at 120 ms, & the Real condition is most highly correlated with the Difference template at 200 ms. Closer inspection of ES's classification images shows a shift over time from using the right edge of the square to using the bottom edge of the square in the Real condition. This change in strategy could have served to reduce the correlations at the earlier points in the stimulus presentation. The third observer (MS) shows no consistent effect in either direction. This is most likely because he was not using the illusory contours in the Illusory condition & relied almost exclusively on the edges of the inducers in the Real condition.

6 SUMMARY, CONCLUSIONS, & FUTURE DIRECTIONS

We had observers discriminate between fat & thin Kanizsa squares in spatiotemporal noise, & found that illusory contours emerge in observers' classification images as early as 47ms & are fully present by about 120 ms. These results are relatively consistent with previous estimates of the time course of completion, e.g. Ringach & Shapley (1996): 120-170 ms; Sekuler & Palmer (1992): 100-200 ms; Murray et al. (2001): 46-114 ms.

Although observers participated in 30,000 trials per condition, the classification movies are still quite noisy. This is consistent with the results of Xing & Ahumada (2002), who found that the signal strength in spatiotemporal classification images is highly limited by internal noise at high noise frame rates. Therefore, one possible way to reduce the noisiness of our classification images would be to increase the duration of each noise frame. An alternate approach would be to present the stimuli in static noise for varying stimulus durations & interrupt completion with a mask (Ringach & Shapley, 1996). We are currently exploring these possibilities.

7 REFERENCES & ACKNOWLEDGEMENTS

Brainard, DH. (1997) The Psychophysics Toolbox, *Spatial Vision* 10:433-436.
Gold JM, Murray RF, Bennett PJ, & Sekuler AB (2000). Deriving behavioral receptive fields for visually completed contours. *Current Biology*, 10:663-666.
Green DM & Swets JA (1966). *Signal Detection Theory & Psychophysics*. New York: Wiley.
Murray RF (2002). Perceptual organization & the efficiency of shape discrimination. PhD thesis, University of Toronto.
Murray RF, Sekuler AB, & Bennett PJ (2001). Time course of amodal completion revealed by a shape discrimination task. *Psychonomic Bulletin & Review*, 8(4):713-720.
Ringach DL & Shapley R (1996). Spatial & temporal properties of illusory contours & amodal boundary completion. *Vision Research*, 36:3037-3050.
Sekuler AB & Palmer SE (1992). Perception of partly occluded objects: A microgenetic analysis. *JEP: General*, 121:95-111.
Tjan BS, Braje WL, Legge GE, & Kersten D (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, 35(21):3053-3069.
Xing J & Ahumada AJ (2002). Estimation of human-observer templates in temporal-varying noise. *Journal of Vision*, 2(7):Abstract 343.

Thanks to Chad Bouchard & Michael Schulbaum for their tireless participation as observers.