

MEASURING SIZE OF A NEVER-PRESENT OBJECT: VISUAL OBJECT FORMATION THROUGH SPATIOTEMPORAL INTERPOLATION

Tandra Ghose¹, Janelle Liu² and Philip J. Kellman³

¹Psychology II, University of Kaiserslautern, Germany,

²Columbia University, New York, U.S.A.,

³Department of Psychology, U.C. Los Angeles, U.S.A

Tandra Ghose <tandraghose@gmail.com>, Phil Kellman <Kellman@cognet.ucla.edu>

Abstract

Spatiotemporal interpolation (STI) refers to perception of complete objects from fragmentary information across gaps in both space and time. No information about shape and size of the object is present in any static frame. Palmer, Kellman & Shipley (2006) found that STI for both illusory and occluded objects produced performance advantages in a discrimination paradigm. Here we report psychophysical studies testing whether STI produces representations that include metric properties of objects. By using an interleaved staircase method, we found that length can be accurately recovered for illusory triangles specified by sequential partial occlusions of background elements in their paths (the STI condition). In the control condition, three moving dots located at the vertices provided the same spatial and timing information as in the STI condition but they did not induce perception of interpolated contours or a coherent object nor accurate length judgments.

In ordinary perception, most objects are partially visible due to occlusion by other objects. For a moving observer, moreover, the pattern of occlusion changes with viewpoint; as a result, a full view of the object may never be present at any instance. For interacting with the objects in the environment, the moving observer has to recover the shape and dimensions, in many cases, of an object that is never fully present in view. Previous research shows that observers can recover and discriminate between the shapes of dynamically occluded objects when their visible fragments satisfy certain spatial and temporal relations, formalized as *spatiotemporal reliability* (Palmer, Kellman, & Shipley, 2006). However, previous research has not shown whether metric properties can be effectively recovered from such dynamically occluded objects.

The process of spatiotemporal interpolation relies both on spatial extraction of object features as well as an hypothesized representation, the dynamic visual icon, that not only preserves fragmentary information over time, but spatially updates the positions of previously seen fragments, allowing them to be appropriately connected to subsequently extracted object fragments (Palmer, Kellman & Shipley, 2006). Because this mechanism of spatiotemporal interpolation includes within it a positional updating mechanism, it could provide the information needed for specifying metric properties in the object representations that result. If metric information can be given through unit formation from motion, either or both of these processes may contribute. Besides testing the possibility of metric representations, the research reported here was designed to distinguish the role of these two kinds of processes. This was accomplished by comparing contour sensitive STI displays with control displays that did not support STI but contained the possibility of common fate information (albeit common fate of elements that were not simultaneously present).

Method

To ensure *spatiotemporal interpolation* we chose kinetic illusory shapes as the stimuli for this experiment that were constructed in a way such that the visual information necessary for interpolation in static spatial displays was never present at any given time. Thus, the question of metric properties may take the form: "What is the length of a side of an object that is never present in the stimulus?"

The experimental design included three crossed within-subject variables: display-type (illusory triangle or dots), target base-length (9.0 cm and 12.0 cm) and staircase direction (2up-1down or 2down-1up). Each observer completed 2 blocks of 4 interleaved staircases in a given direction with a five minute break in between the blocks. In each block, trials from each staircase were presented in random. Fourteen undergraduate students of the University of California, Los Angeles with normal or corrected-to-normal vision participated in this experiment.

Figure-1 shows the displays for the two conditions used in this experiment- the spatiotemporal interpolation (STI) condition and the control condition. In both, elements were defined by sequential partial occlusion of black squares that were placed in the path of a white triangle oscillating on a white background. In the STI condition, the requirement for truly spatiotemporal interpolation was ensured by allowing no more than one vertex of the triangle to be visible at any moment. This design prevented observers from recovering the length of the triangle's base simply by spatial interpolation between two vertices. In the control condition, dots were placed at the vertices of virtual triangles with the same dimension as the ones in the STI condition. Although the dots provided the same spatial and timing information as the triangle vertices in the STI condition, they did not induce perception of interpolated contours or a coherent object.

EXPERIMENTAL SET-UP:

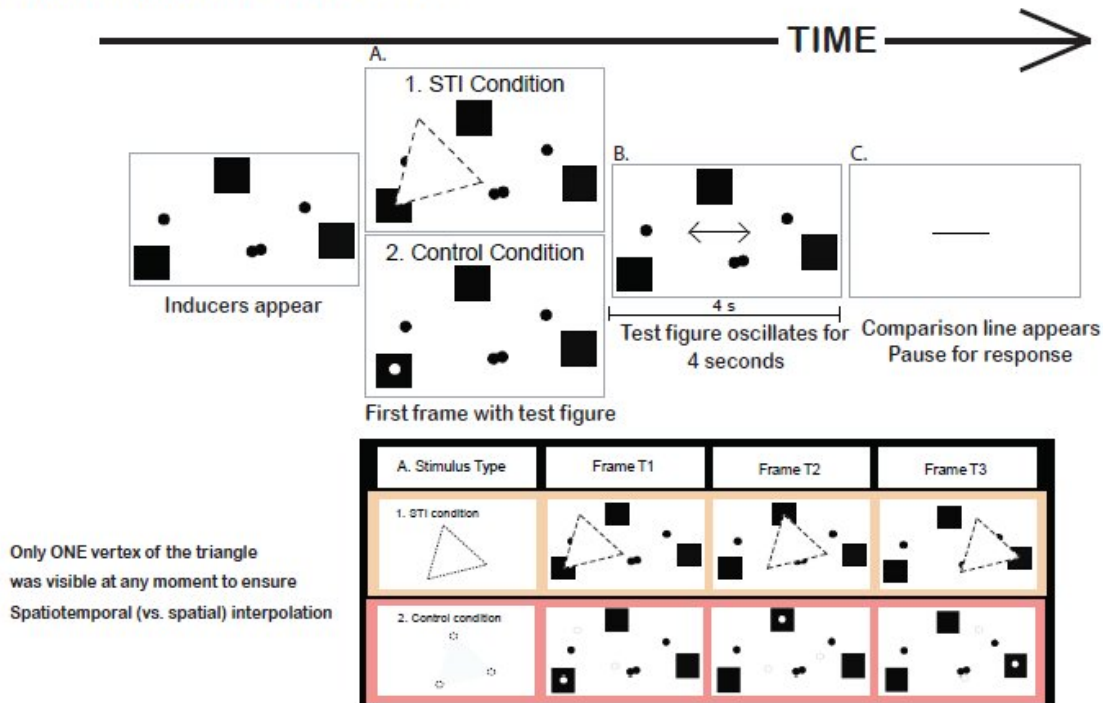


Figure-1: Timeline for an experimental trial and the two kinds of displays used in this experiment.

The illusory equilateral triangle was presented with two baselengths at 9.0cm and 12.0cm (4.5° and 6° of visual angle respectively or 120 and 160 pixels respectively). The triangles were rotated by 15° from the horizontal direction of motion to ensure unambiguous motion information. (When a straight edge is oriented parallel to the direction of motion, its motion is ambiguous because the component of motion along its length produces no changes on the retina.) As the triangle translated across the screen, black inducing elements (squares with sides of 4.5cm/60pixels) were sequentially occluded by vertices of the triangle. More inducing elements in shape of small dots (1.5cm/20pixels in diameter) were placed in fixed positions in order to refresh the memory about the speed of triangle (as shown in Figure-1). Two small dots (1.5cm in diameter, each) were placed close together on the triangle's base so as to identify which side was the target base. These dynamic displays created a clear representation of an illusory triangle despite the little information actually given.

In the Control condition, the dot stimuli contained the same black inducing elements, but the illusory triangle was specified by three dots (1.5cm in diameter) located at each vertex of the underlying equilateral triangle (Figure 1). As in the STI condition, only one dot was visible at any instant. The presence of dots instead of vertices with tangent discontinuities prevented contour interpolation based on the principles of spatiotemporal reliability. No definite form was perceived in the moving display; however, since the dots moved together, they could group together into a visual unit as expected by the Gestalt grouping cue of common fate (assuming common fate can operate with elements not simultaneously visible). In order to estimate the extent between the dots at the base of the virtual triangle, the observer would have to use cognitive strategies based on speed and time instead of spatiotemporal reliability.

The points of the underlying psychometric function that were sampled in the experiment were fitted with cumulative Gaussian using the PSIGNFIT toolbox (Wichmann & Hill, 2001) to fit psychometric functions to data points. The toolbox returned the threshold (the 50% point) and the slope at the mean for the best fitting cumulative Gaussian for the data points. The mean gives the measure of the point of subjective equality (PSE) or the measure of length at which the observer is 50% likely to make "longer" judgment. The slope of the psychometric function defined an *imprecision* measure for each stimulus. The greater the slope, the more precisely could the base-length be estimated.

Results and Discussion

Figure-2 depicts the average of experimental results for fourteen observers. The two charts show the accuracy and imprecision measures for the STI and the control conditions for the two base lengths. The spatiotemporal interpolation condition (STI) produced a more accurate perception of length for both the 120 (9cm) and 160 (12cm) pixels base-lengths. For base length 120, the subjects perceived an average estimated length of 124.38 pixels (standard error of the mean 7.807) in the STI condition compared to 213.5 pixels (standard error of the mean 17.71) in the control condition. For base length 160 pixels, the subjects perceived an average estimated length of 171.42 pixels (SE=7.71) in the STI condition compared to 213.36 pixels (SE=19.65) in the control condition. The observations were confirmed by analyses using a 2 X 2 repeated measures ANOVA. There were reliable main effects of base length { $F(1,13)=17.02$, $p<.001$ } and display-type { $F(1,13)=12.68$, $p<.005$ } as well as a reliable base length X display-type interactions { $F(1,13)=12.42$, $p<.005$ }. The values differed significantly between the STI and the control conditions both for 160 pixels { $t_{13}=2.1$, $p<.05$ } and 120 pixels { $t_{13}=4.7$, $p<.001$ }. This pattern of results is consistent with the notion that estimation of metric properties such as length is more accurate when form is recovered due to STI than when length has to be recovered through cognitive strategies.

The imprecision measure is an indicator to the sensitivity of the processes involved in the recovery of metric properties on spatiotemporal interpolation. Lower imprecision for STI condition compared to the control condition would indicate that the availability of form from STI produces more precise and accurate results than those achieved through cognitive strategies. The imprecision measure for the baselength of 160 pixels was consistent with this prediction. For baselength of 160 pixels, the imprecision for the STI condition, 21.14 pixels (SE=3.88) was much lower than the control condition 41.07 pixels (SE=6.10). The planned t-test supported the findings that these values differed significantly $\{t_{13}=4.1, p<.001\}$. However, the imprecision results for the STI and control conditions for baselength of 120 pixels did not differ significantly $\{t_{13}=-0.8, p=.41\}$. The imprecision values were 36.88 pixels (SE=5.14) in the STI condition compared to an imprecision of 31.47 pixels (SE=5.03) in the control condition. In a 2X2 repeated measures ANOVA, there were no reliable main effects of base length or $\{F(1,13)=.112, p=.74\}$ and display-type $\{F(1,13)=3.9, p=.06\}$ but there was a reliable base length X display-type interactions $\{F(1,13)=10.98, p<.05\}$. We do not have a good explanation for the imprecision measures for the base length of 120 pixels for the STI condition.

These results show that spatiotemporal interpolation can produce object representations with accurate metric properties. Thus, cognitive inferences about length based on position and timing, as in the dots control condition, were much less accurate. This suggests a special role of perceptual object formation from spatiotemporal interpolation in producing representations of functionally important object properties.

RESULTS:

 STI object
 Dot control
N=14 subjects
160 pixels ~ 6.0 degrees
120 pixels ~ 4.5 degrees

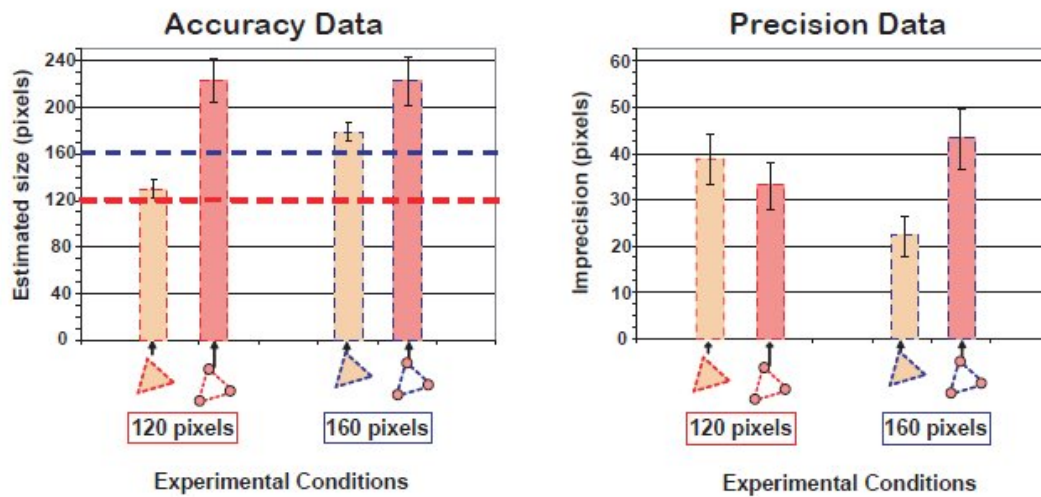


Figure-2: Results as elaborated above

References

- Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General*, 135, 513–541.
- Wichmann, F.A. and N.J. Hill: The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and Psychophysics* **63** (8), 1293-1313 (2001)