

Object perception from cross-sectional images with in-situ and ex-situ displays

Bing Wu, Roberta Klatzky, George Stetten

A widely-used method to display 3D imaging data is to present it as a sequence of cross-sectional images. The 3D structure is then to be perceived by mentally integrating information over images. The current research evaluates the effectiveness of two types of displays, the Sonic Flashlight (SF) and conventional viewing, on the spatiotemporal integration of information across cross-sectional images. The SF creates an *in-situ* visualization that portrays imaged data in their appropriate point of origin in space, enabling perceptual processes to be directly brought to localize and perceive the imaged anatomy in 3D space (Stetten, Chib, & Tamburo, 2000; Stetten, 2003). Alternatively, conventional displays show data *ex-situ* at locations external to its source. In the following experiments, we investigate how human processing enables the construction of 3D representations from 2D images when using *in-* and *ex-situ* displays.

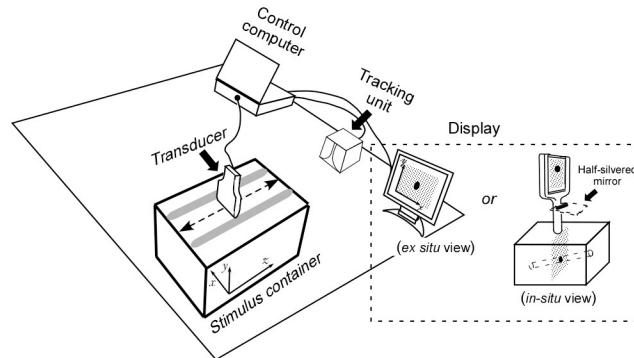


Figure 1. The *ex-situ* and *in-situ* display of cross-sectional images.

Experiment 1: Perception of familiar forms from cross-sections

Experiment 1 was to examine subjects' ability to perceive patterns from cross-sectional images.

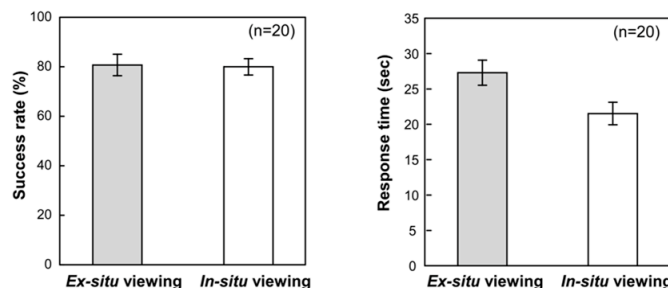
Stimuli & Task: Subjects explored a hidden target by moving an imaging transducer over it. Successive cross-sectional views were displayed *in-situ* at the location of the target or *ex-situ* at a displaced display. Subjects were informed that the target was a digit or an English letter in upper or lower case. The task was to identify the target as accurately and quickly as possible.

Figure 2. The setup of Experiment 1.

Results: (a) The superiority of *in-situ* viewing was shown by faster identification speed at comparable accuracy levels.

(b) A common type of error with *ex-situ* viewing was to mis-report the configuration of features that were otherwise perceived correctly.

(c) Subjects' performance was influenced by the cross-sectional complexity of stimuli (curves > lines; diagonals > straight lines; asymmetry > symmetry).



(a)

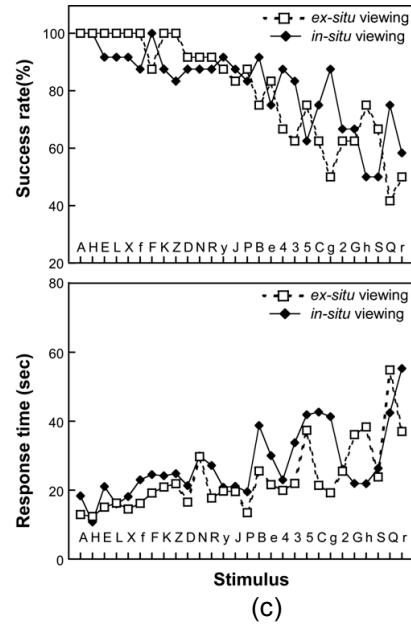
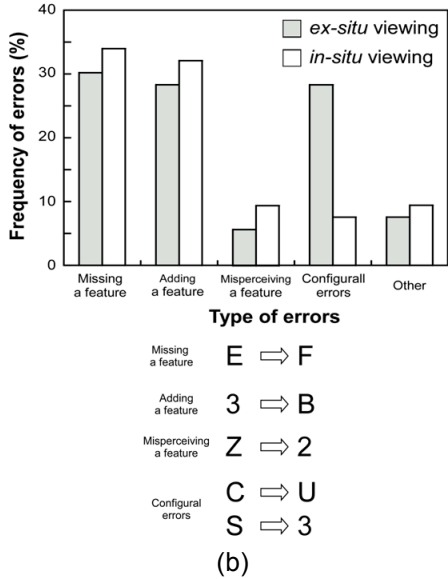
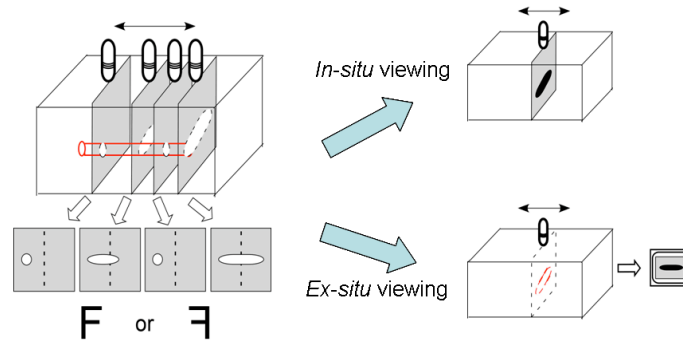


Figure 3. The results of Experiment 1. (a). The mean success rate achieved (left) and the mean correct response time (right) with two displays; (b). Common error patterns in subjects judgments; (c). Effects of stimulus complexity on subjects' performance.

Experiment 2: Perception of the spatial relationships among pattern elements

Experiment 1 suggests that individual features can be encoded *ex-situ*, but the configuration is not treated as a whole. We used mental rotation to probe for configural vs. feature-based processing.

Stimuli & Task: The hidden pattern was a pre-cued letter (F, y, J, or R) in one of four possible orientations (0°, 90°, 180°, or 270°). Subjects judged from cross-sectional views whether it was normal or mirror-imaged as accurately and quickly as possible.



Task: To judge if the letter (F, y, J, or R) is normal or mirror-reversed.

Figure 4. The experimental setup of Experiment 2.

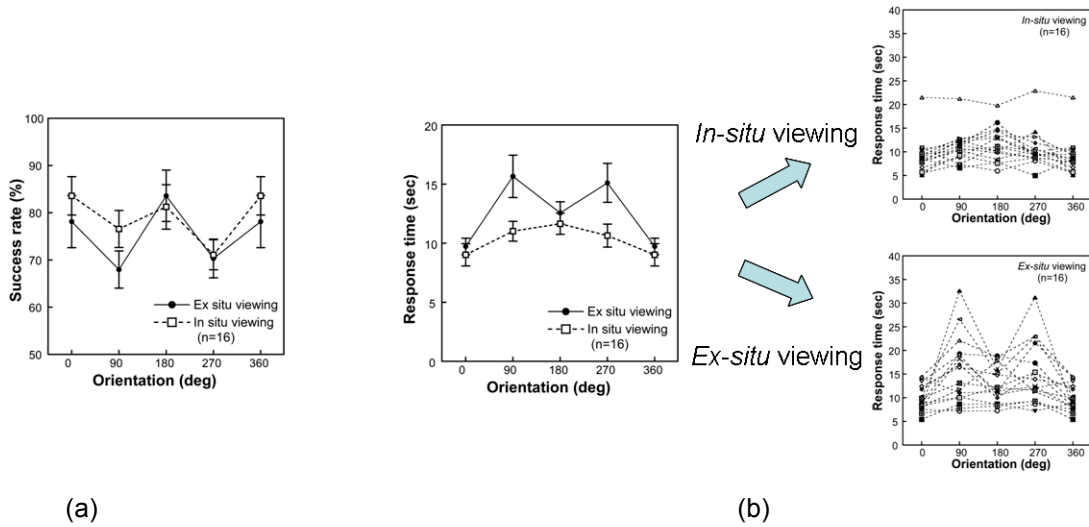


Figure 5. The results of Experiment 2. The mean success rate (a, left) and the mean correct response time (b, right) as a function of pattern orientation.

Results: Data from *in-situ* viewing are consistent with mental transformation of a wholistic configuration. The *ex-situ* data are consistent with the hypothesis that rotated letters are processed feature by feature, not as configural wholes (Jolicoeur, 1990).

Experiment 3: Effects of 2D pattern complexity on subjects' performance

Experiment 3 pursued the implication of Experiment 1 that difficulty for both viewing conditions increased with cross-sectional complexity, but *in-situ* viewing was superior regardless of complexity. The stimuli used were multi-segment objects similar to those described in Palmer (1977).

Stimuli & Task: Subjects were presented with a target via cross-sectional views and subsequently a corresponding whole pattern on a LCD. The task was to scan the target to build a representation, then discriminate whether the two patterns were the same or different.

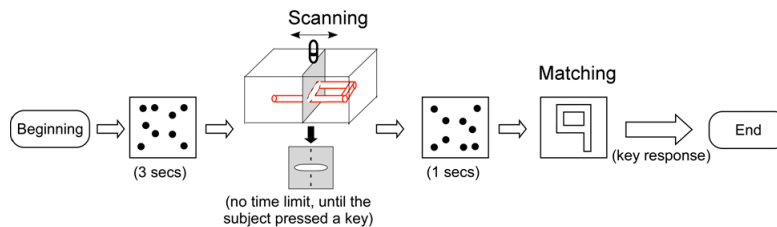


Figure 6. The setup of Experiment 3.

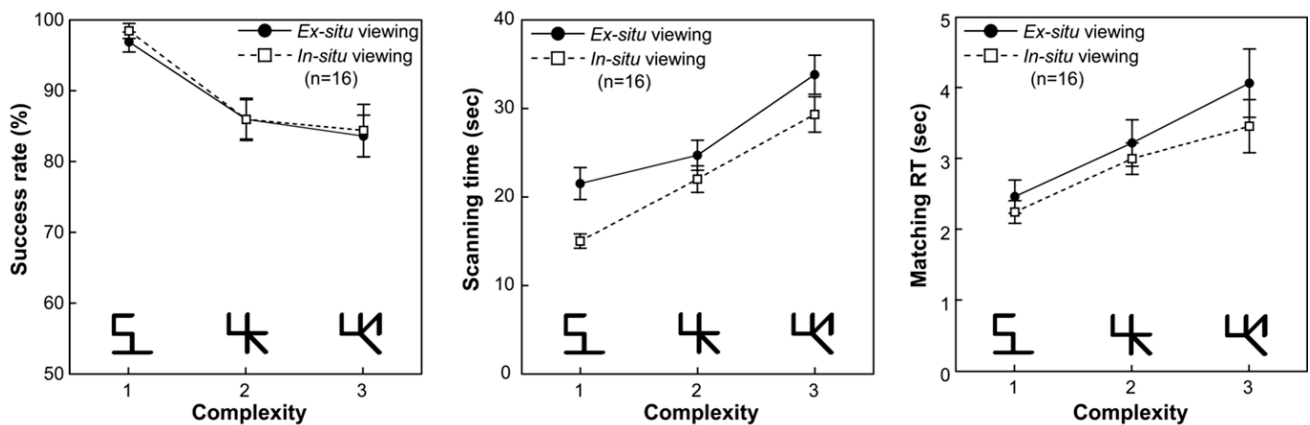


Figure 7. The results of Experiment 3. From left to right: the mean success rate, the mean scanning response time and the mean correct matching response time as functions of pattern complexity

Results: As predicted, *in-situ* viewing resulted in faster formation of a pattern representation, and there was some residual advantage for the matching portion of the trial. Furthermore, a multiple regression analysis predicted the scanning time using two parameters, representing the spatial density of cross sections and the spatial extension of the object segments along the scan ($r^2 = 0.72$ and 0.80 for the *ex-situ* and *in-situ* data, respectively).

Conclusions:

1. People can mentally construct complex patterns from cross-sectional views.
2. Results revealed an advantage for *in-situ* visualization over *ex-situ* display in the construction of mental representations from cross-sectional images. The subjects' ability to accurately integrate information from images relies on co-locating the visual consequences of exploratory action with the action itself, as implemented in the SF display. In contrast, when vision is displaced, features can be encoded, but it is difficult to construct a configuration relating them spatially.
3. Ongoing work is extending this research to fully 3D patterns. Initial results show that superiority of *in-situ* display is being maintained.

Reference:

- Jolicoeur, P. (1990). Identification of disoriented objects: a dual systems theory. *Mind and Language*, 5, pp. 387–410.
- Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9, 441-474.
- Stetten, G., Chib, V., & Tamburo, R. (2000). System for Location-Merging Ultrasound Images with Human Vision, *IEEE Proceedings of the Applied Imagery Pattern Recognition (AIPR) Workshop*, Washington D.C., pp. 200-205, 2000.
- Stetten, G. (2003). System and Method for Location-Merging of Real-Time Tomographic Slice Images with Human Vision, U.S. Patent no. 6,599,247, issue date, July 29, 2003.