Multiple Object Tracking through temporal gaps created by the fading of objects

Harry Haroutioun Haladjian[1,2], Zenon W. Pylyshyn[1], & Allan Kugel[1]

[1]Rutgers Center for Cognitive Science, New Brunswick, NJ; [2]Department of Psychology, Rutgers University, New Brunswick, NJ

VSS 2009 Naples, Florida

1. Abstract

In three experiments, we examine whether the encoding of object location is used in Multiple Object Tracking. Observers were asked to track four target discs among eight identical distractors on a display where the same random-dot texture was used for object surfaces and display background. Stereoscopic glasses were used to create two display conditions: 3D (where objects appeared to float in front of the background texture) and 2D (where objects appeared on the background texture). In the 2D displays, disks were only visible while they moved and became indistinguishable from the background when they stopped. In 75% of the trials, the objects halted movement mid-trial for one, two, or four seconds.

Experiment 1 used textured discs with no borders. During the pauses, the discs would appear to dissolve into the background in the 2D condition (because their texture is identical to the background) but remained distinct in the 3D condition. This produced significantly lower tracking performance only in the 2D trials with the longest pause; as expected, no decline was observed in the 3D condition.

Experiment 2 was identical to the first, except the discs had a white border during the entire trial, allowing the discs to remain distinct during halts. In this case there was no effect of pause duration.

Experiment 3 used the same 2D display as Experiment 1, except that in half of the trials object borders briefly flashed "on" before halting. Here, there was an effect of pause duration in both flash and non-flash conditions (decreased performance with longer duration).

These experiments found that objects that disappear without an abrupt offset are more difficult to track, indicating that object locations are not always encoded and used to continue tracking after a gap in visibility. This suggests that the tracking mechanism may not encode location information unless cued by abrupt changes in the visual scene.

2. Background

According to Visual Indexing theory, the success of Multiple Object Tracking (MOT) does not rely on the encoding and continual updating of object location or the extrapolation of motion trajectories (Pylyshyn, 1989, 2001).

How location information is used in MOT is still not well understood. For example, when objects disappear abruptly, there is a tracking advantage for objects that reappear in the loci of disappearance instead of the extrapolated trajectory (Keane & Pylyshyn, 2006). This suggests location encoding may occur during tracking, but this may be due to an abrupt event, such as moving behind an occluder or sudden offset.

In the following experiments, we examine **whether or not location encoding occurs during MOT when the object disappearance is gradual** (i.e., by fading into the background texture).

3. Method

Participants:

- Rutgers students received course credit or renumeration for one 60-min session (only subjects who could see the stereo display participated).

Apparatus & Stimuli:

- Experiments were programmed in C with OpenGL (Linux OS) and presented on a 17" CRT color monitor with viewing angle of ~ 33.6° x 26.4°.
- A random-dot grid of (random) grayscale values was used for the object and background textures in order to create two display conditions:
  1) **Stereo (3D) display:** objects appeared to float in front of the background texture (using eDimensional shutter glasses and disparity); and
  2) **Mono (2D) display:** objects appeared directly on the background texture (no disparity code).
- Objects were circles with ~2.3 cm diameter (~2.1° viewing angle) and the same random-dot texture as the background. A white border 3 pixels thick was used for target flashing and the moving stimuli in Experiment 2.
- Objects moved smoothly in straight lines until collision with the screen border or within ~2° of another object, and bounced in an "elastic" manner; objects never occluded one another.

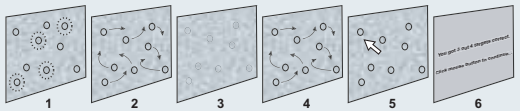


Figure 1. Experimental set-up. Panel 1: initial display of identical circles with flashing targets. Panel 2: objects move without overlap. Panel 3: a mid-trial pause allows borderless objects to fade into background. Panel 4: movement resumes after a specified duration. Panel 5: observer selects targets with mouse. Panel 6: Tracking feedback screen. (Note: black circle borders not present in actual experiments.)

General Procedure:

- Subjects sat approximately 60 cm from the monitor in a darkened room.
- 12 identical circles appeared at random locations and 4 were identified as targets by blinking outlines; then they began moving around the screen.
- In 75% of the trials, circles stopped movement for either 1, 2, or 4 seconds and then resumed movement.
- At the end of the trial, faint outlines appeared on the circles and subjects selected the targets using a mouse pointer; tracking feedback was provided.
- Data were analyzed in SPSS (ANOVA & LSD post-hoc tests; "subject id" was treated as a random variable); subjects with poor performance (near chance) or with incomplete data were excluded from the analysis.

Figure 2. Example of the random-dot texture with white borders (partial screen shot from the experiment display).

4. Results

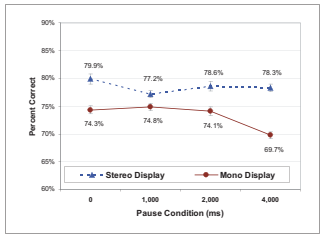
EXPERIMENT 1 (n=22)

Two display types: stereo (3D) & mono (2D).
Four pause conditions: 0, 1, 2, & 4 seconds.

Results

- ANOVA showed interaction between display and pause duration (F=4.3, p<.01), and main effects for display type (F=54.0, p<.001) and pause durations (F=5.1, p<.01).
- Stereo display showed significant difference between no pause and 1-s pause.
- Mono display showed significant difference between the 4-s pause and all other pauses.
- Tracking performance in trials with extended pause durations was further explored in Experiment 1b (for mono display only).

Experiment 1 Tracking Results



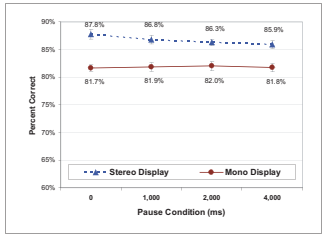
EXPERIMENT 2 (n=20)

Identical to Experiment 1, except objects had a constant white border throughout the trial.

Results

- ANOVA showed an advantage for stereo (F=17.8, p<.001), with no interactions.
- There was no effect of pause duration on tracking performance.

Experiment 2 Tracking Results



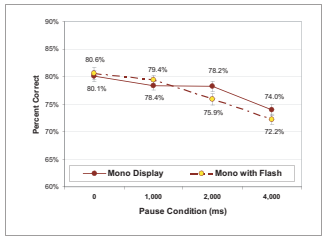
EXPERIMENT 3 (n=21)

Two display types (mono only): ***with flash*** or ***without flash*** (the object-borders flashed "on" for 15 ms immediately before pauses).
Four pause conditions: 0, 1, 2, & 4 seconds.

Results

- Overall ANOVA showed an effect only for pause duration (F=14.9, p<.001).
- For the no-flash trials, significant differences between 4-s & all other pause conditions (replicating mono condition in Experiment 1).
- For the flash trials, significant differences between: no-pause & 2-s, no-pause & 4-s, 1-s & 2-s, 1-s & 4-s, and 2-s & 4-s.

Experiment 3 Tracking Results



Note: All error bars represent average within-subjects variability.

4. Results (Cont'd)

EXPERIMENT 1b (n=15)

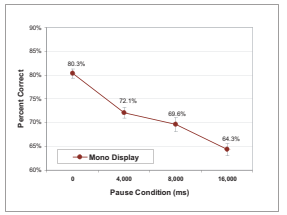
One display type: mono.

Four pause conditions: 0, 4, 8, & 16 seconds.

Results

- Overall ANOVA showed a main effect of pause duration (F=22.9, p<.01), with a continuous and significant decrease in tracking performance.

Experiment 1b Tracking Results



5. Conclusion

- Objects that disappear gradually were more difficult to track during longer disappearances (4 or more seconds), indicating that object locations were not encoded effectively and used to continue tracking after a gradual gap in visibility.
- This suggests that simply remembering locations does not benefit tracking -- objects must remain visible for properties, such as location, to remain attached to the object.
- Abrupt changes to tracked objects may reinforce location-encoding, but further experiments testing this hypothesis are necessary.

6. References

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology, 52*(4), 346-368.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*(1), 65-97.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1-2), 127-158.

Special thanks to Ryan McKendrick and Robyn White for assistance with data collection. For information about this and other projects at the Visual Attention Lab:

http://ruccs.rutgers.edu/finstlab/
haladjian@ruccs.rutgers.edu