The Implicit Processing in Multiple Object Tracking

Hirokazu Ogawa Department of Psychology, Kwansei Gakuin University e-mail
Akihiro Yagi Department of Psychology, Kwansei Gakuin University

Previous research has shown that implicit memory of visual context guides visual attention to a target object in a dynamic scene (Chun & Jiang, 1999). We investigated how attention affects implicit learning of contextual information in dynamic scene using a multiple object tracking (MOT) task. Participants were asked to track five identical targets which moved independently and unpredictably among five identical distractors. In this task, the motion patterns (trajectories) of the target items (Experiment 1) or the distractor items (Experiment 2) were made invariant by repeating them throughout the entire experimental session. The results showed that the invariant motion patterns of the target set improved MOT performance implicitly. In addition, participants demonstrated greater performance when the motion patterns of both target and distractor set were made invariant. This additional facilitation by the distractor set was not observed when only patterns of the distractor set was repeated. These data suggest contextual modulation of attentional tracking and sensitivity to the global motion pattern in a dynamic scene.

Keywords: contextual cueing, implicit learning, multiple object tracking.

Introduction

Visual attention plays a critical part in human conscious perception of the visual world. Observers are often unaware of changes in the visual scene until attention is directed to the changed object (Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1997). Such "change blindness" phenomenon implies that the capacity of the visual representations is severely limited. The role of visual attention is to rapidly prioritize aspects of a complex scene to pick up behaviorally relevant information and ignore irrelevant information. But, given that visual attention can operate in such an efficient manner, one question that arises is — how does the visual system determine where to direct attention?

Chun and Jiang (1998) claimed that visual context is one strong candidate, which exists in almost all visual scenes. They used visual search task in which participants had to detect a target presented surrounded by multiple, competing distractor stimuli. The spatial configuration of the search items was defined as a context, and particular display layouts were presented repeatedly throughout the entire experimental session. The result showed that participants could detect targets in the invariant configurations faster than targets in configurations newly generated in each blocks to serve as a control. Furthermore, participants were not aware of the repetition of configurations and did not have any explicit memory of the context, as demonstrated by incidental recognition tests. These results were interpreted as the evidence that implicit memory of visual context can modulate attentional deployment by top-down guidance. This facilitatory effect was called *contextual cueing*.

On the other hand, Jiang and Chun (2001) demonstrated that attention also modulates implicit learning of visual context in visual search. In their study, visual search display consisted of eight red items and eight green items. Participants were instructed to attend to only one of colors (e.g., red) and ignore another color (e.g, green). The configuration of attended items or of ignored items was made invariant across different experimental blocks. The result showed that the repetition of attended color configuration could facilitate the performances in the visual search task, while that of ignored color resulted in no contextual cueing effect.

In the present study, we investigated contextual cueing further based on following two questions. First question concerns the generality of the contextual modulation of visual attention. Previous studies on contextual cueing used only visual search task, so that it is unclear whether contextual information affects other attentional processing such as attentional tracking. Thus, we used a multiple object tracking (MOT) paradigm (Pylyshyn & Storm, 1988). In this paradigm an observer was asked to track a number of identical items which were independently and unpredictably moving among identical distractors. It is assumed that MOT task demands the resource of attention and visual memory. If contextual information can facilitate performances in MOT task, it would generalize the contextual cueing effect. Second question is whether the interaction between visual attention and implicit memory of visual context exists in the dynamic display. Clearly, the visual world is not static, and the contextual guidance of visual attention would be more useful in visual processing if it were not restricted to the static display. We tested this question using the display which changed randomly and unpredictably over time. Although Chun and Jiang (1999) reported that the dynamic context defined by complex motion trajectories facilitated performance in the visual search task, it is important to clarify the converse effect. That is, we test the influence of attentional set on learning contextual information in dynamic display.

Experiment 1

Method

Sixteen undergraduate students (8 males and 8 females) participated in Experiment 1. They had normal or corrected-to-normal vision, and were naïve to the purpose of the experiment.

All stimuli were presented on a CRT monitor controlled by a NEC computer. Figure 1 shows the sequence of events during a trial. The display consisted of 10 white circle items (0.83° in diameter), on a black background (14.7° × 14.7°). At the start of a trial 5 of 10 items were designated as the target items which were to be tracked by flashing them on, and off, five times. After the targets had been designated each item started to move with a velocity vector that was changed randomly every few hundred milliseconds. The items also changed their velocity vectors when the distance from the edge of the display, or from other items, was less than 0.2°. A smooth motion continued for 7,500 ms. The maximum velocity of the items was adjusted in a practice session (50 trials) for each observer so that, on average, four items could be tracked at the beginning of the main experimental session. At the end of the movement a mouse cursor appeared at the center of the display. An observer then had to indicate five targets by mouse clicks. No feedback was given.

The investigation used a 3 × 5 design. Each of three repetition conditions was used as an experimental factor: an “all old” condition; an “old target” condition; or an “all new” condition. In the “all old” condition, all of the items (targets and distractors) had motion patterns which were made invariant by repeating them throughout the entire experimental session. In the “old target” condition, the pattern of the target items was made invariant but the pattern of the distractor items was made from a new generation for each block. In the “all new” condition all of the motion patterns were new generations for each block. So that the initial and the final layout of the items could not be memorized, the trajectories for the first second and the last second of the display period were recalculated for each repetition. Another experimental factor involved an “epoch” or group of trials. Each epoch was comprised of three blocks, each of which included Five “all old” condition trials, five “old target” condition trials and five “all new” condition trials.

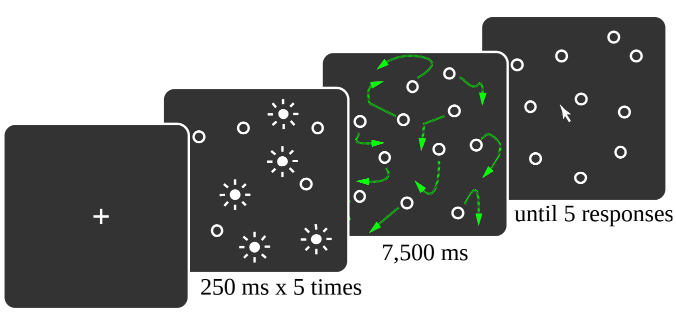

Figure 1. The sequence of events during a trial.

After the last block of the tracking task the observer was asked to perform an incidental explicit recognition test in which five “all old” patterns and five newly generated patterns were presented in a random order. The observer was required to answer whether or not a pattern had been presented in the preceding tracking task.

Results and Discussion

Figure 2 shows the mean number of correctly tracked items as a function of the epoch for each repetition condition. The data were entered into a two-way within-participants analysis of variance (ANOVA) with condition (all new, old target or all old) and epoch (1 – 5) as the main variables.

Both main effects were significant: condition, $F(2, 32) = 16.63$, $p < .0001$; and epoch, $F(4, 64) = 2.84$, $p < .05$. The two-way interaction was marginally significant: condition × epoch, $F(8, 128) = 1.95$, $p < .058$. To compare the three conditions more closely, we conducted three separate two-way ANOVA comparing “all new” vs. “all old”, “all new” vs. “old target” and “all old” vs. “old target”.

“All new” versus “old target” condition. Main effect of condition was significant, $F(1, 16) = 11.36$, $p < .001$, reflecting the finding that the performances in “old target” condition were greater than those in “all new” condition. No other effect or interaction was significant.

“All new” versus “all old” condition. Both main effect were significant: The tracking performances in “all old” condition were greater than those in “all new” condition, $F(1, 16) = 11.36$, $p < .005$; and the performance increased as epoch increased, $F(4, 64) = 3.07$, $p < .05$. There was also

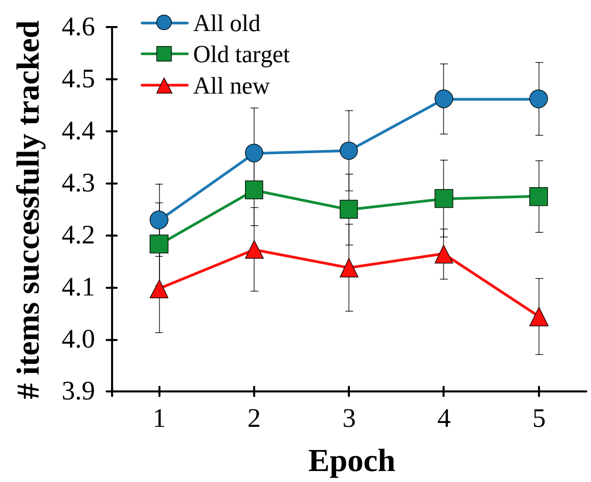

Figure 2. Mean tracking performance data as a function of epoch for Experiment 1. The error bars indicate the standard error of mean.

a significant condition × epoch interaction, $F(4, 64) = 3.94$, $p < .01$.

"All old" versus "old target" condition. Both main effect were significant: The tracking performances in "all old" condition were greater than those in "all new" condition, $F(1, 16) = 8.25$, $p < .05$; and the performance increased as epoch increased, $F(4, 64) = 2.90$, $p < .05$. The interaction was not significant.

So that we could rule out the possibility that the facilitatory effects on the multiple object tracking were based on explicit memory we analyzed the incidental recognition test. We found that although 7 of 16 observers reported that they were aware of the repetition manipulation, the accuracy of explicit recognition was at chance level (overall accuracy = 49.4%, hit rate = 42.3%, false alarm = 43.5%).

There were three main results in Experiment 1. First, the repetition of the motion pattern facilitated the tracking performance in "target old" and "all old" condition. This indicates that the dynamic contextual cueing effect could be observer using MOT task, as well as visual search task (Chun & Jiang, 1999). Second, this learning of the dynamic motion pattern was implicit. Third, the tracking performance was greater in "all old" condition than in "old target" condition. It is unclear that, however, the improvement of the tracking performance was due to the additional distractor information or the invariant global motion pattern. To clarify this issue, we carried out the second experiment in which the motion pattern of the distractor set was invariant across blocks whereas the pattern of the target set was variable. If the invariant motion pattern of the distractor set itself improves the tracking performance, we should find comparable facilitatory effect between Experiment 1 and 2.

Experiment 2

Method

Eighteen naïve undergraduate and graduate students served as participants in Experiment 2. They had normal or corrected-to-normal vision.

The materials and procedure were same as Experiment 1, expect that "old distractor" condition was conducted instead of "old target" condition in Experiment 1. In "old distractor" condition, the pattern of the distractor items was made invariant but the pattern of the target items was made from a new generation for each block.

Results and Discussion

Figure 3 shows the mean tracking data. Two factors were used in a repeated measures ANOVA: condition (all new, old distractor or all old) and epoch (1 – 5). This revealed main effects of condition, $F(2, 36) = 4.68$, $p < .05$, and epoch, $F(4, 72) = 4.39$, $p < .01$. There was also a significant condition × epoch interaction, $F(8, 144) = 3.05$, $p < .01$.

"All new" versus "all old" condition. There were significant main effects of condition, $F(1, 17) = 15.05$, $p < .01$,

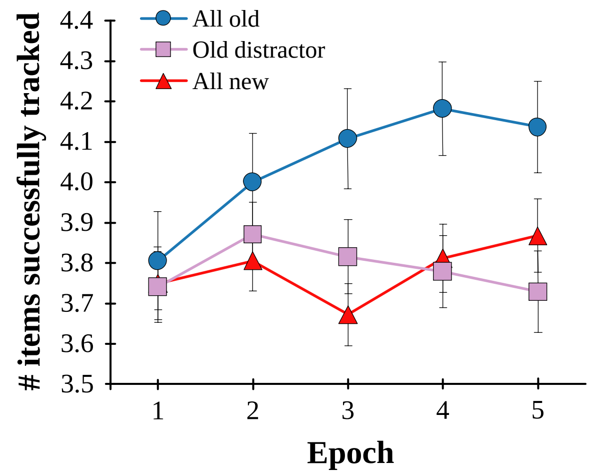

Figure 3. Mean tracking performance data as a function of epoch for Experiment 1. The error bars indicate the standard error of mean.

and epoch, $F(4, 68) = 5.02$, $p < .01$. Condition × epoch interaction was also significant, $F(4, 68) = 4.19$, $p < .01$.

"All new" versus "old distracter" condition. No main effects or interaction were significant.

"All old" versus "old distractor" condition. There were significant main effects of condition, $F(1, 17) = 9.31$, $p < .01$, and epoch, $F(4, 68) = 4.69$, $p < .01$. Condition × epoch interaction was also significant, $F(4, 68) = 4.52$, $p < .01$.

4 of 18 participants reported that they noticed that certain motion patterns were being repeated. But, none of them stated that they tried to explicitly memorize the patterns. The incidental recognition test showed that Hit rate (37.5%) didn't differ from false alarm rate (44.4%), $t < 1$.

We found the facilitatory effect by the invariant motion pattern in "all old" condition, as in Experiment 1. But, tracking performance in "old distractor" condition was not improved, indicating that implicit learning of invariant motion pattern was not affected by the pattern of the distracter set.

General Discussion

In two experiments, we found that the dynamic, complex visual motion patterns in MOT are memorized implicitly and facilitate tracking performance. These results provide converging evidence for contextual cueing effect in dynamic display (Chun & Jiang, 1999). This study extends prior work for implicit learning of context in visual search task (Chun & Jiang, 1998) to learning in multiple object tracking, generalizing the contextual effect on attentional processing.

Furthermore, we found the additive effect of distractor-set information on implicit learning for the motion pattern of the target set, whereas the invariant pattern of only distractor set did not improve tracking performance. These results suggest that implicit learning observed in this task was sensitive to the invariance of the global motion pattern. This is consistent with the notion that the organization of visual short-term memory is based on global spatial configuration (Jiang, Olson, & Chun, 2000).

However, our results are inconsistent with the result of Jiang and Chun (2001), who reported that that the repetition of attended color configurations could facilitate the performances in the visual search task, while those of ignored color resulted in no contextual cueing effect. The discrepancy between our study and that of Jiang and Chun may be due to the difference of the dynamics of display. Recently, Olson and Chun (2002) showed that contextual cueing effect is biased toward spatially grouped information. The dynamic display used in this study might cause mixture of local grouping component. This might result in implicit learning of distractor information that was to be ignored in MOT.

In conclusion, the present results suggest the importance of implicit learning of context information in a dynamic scene. The human visual system encodes complex dynamic covariation between objects in a visual scene and this encoding serves to reduce uncertainty.

Acknowledgement

This study was supported in part by Grant-in-Aid for JPSP Fellows to HO.

References

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28-71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10*, 360-365.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 683-702.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology, 54A*, 1105-1124.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9*, 273-302.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179-197.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368-373.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*, 261-267.