Published in final edited form as:

Vis cogn. 2009 September 1; 17(8): . doi:10.1080/13506280802469510.

The Effect of Spatial and Nonspatial Contextual Information on Visual Object Memory

Hsin-Mei Sun and Robert D. Gordon

Department of Psychology, North Dakota State University

Abstract

Recent research has found visual object memory can be stored as part of a larger scene representation rather than independently of scene context. The present study examined how spatial and nonspatial contextual information modulate visual object memory. Two experiments tested participants' visual memory by using a change detection task in which a target object's orientation was either the same as it appeared during initial viewing or changed. In addition, we examined the effect of spatial and nonspatial contextual manipulations on change detection performance. The results revealed that visual object representations can be maintained reliably after viewing arrays of objects. Moreover, change detection performance was significantly higher when either spatial or nonspatial contextual information remained the same in the test image. We concluded that while processing complex visual stimuli such as object arrays, visual object memory can be stored as part of a comprehensive scene representation, and both spatial and nonspatial contextual changes modulate visual memory retrieval and comparison.

Do we form detailed object representations in visual memory during the viewing of scenes? Studies of change blindness have suggested that visual scene representations are incomplete since coherent visual representations disintegrate as soon as attention is withdrawn from an object (Rensink, O'Regan, & Clark, 1997; Simons, 1996). For example, Simons (1996) argued that only spatial layout, and not detailed object identities, can be maintained in visual memory after viewing scenes. Furthermore, Rensink et al. (1997) used a flicker paradigm and found that changes to the scene (e.g., movement of a railing behind people) can not be detected easily, apparently demonstrating that visual memory is not detailed.

However, recent research has revealed that the occurrence of change blindness can be due to various factors, such as failures of encoding, retrieval and comparison (Hollingworth, 2006b; Hollingworth, Williams, & Henderson, 2001; Mitroff, Simons, & Levin, 2004). Moreover, recent research has found robust visual memory for objects and scenes (Hollingworth, 2005; Hollingworth & Henderson, 2002). For example, Hollingworth (2005) had participants view various images of scenes and tested their memory for the visual form of individual objects in the scenes. The test image was either presented immediately after scene viewing or delayed for one trial, until the end of the study session, or 24 hr. There was little or no decline in change detection performance from immediate test to a delay of one

trial or until the end of the entire study session. Furthermore, after viewing more than 40 scenes and hundreds of individual objects, change detection performance after 24 hr delay remained well above chance. According to Castelano and Henderson (2005), these robust object representations can be formed either intentionally or incidentally during scene viewing.

Given that visual object representations can be maintained reliably after viewing scenes, an interesting question is whether visual object memory is stored as part of a larger scene representation or independently of scene context. To probe this question, we can take advantage of the encoding specificity principle of episodic memory retrieval which suggests that the effectiveness of retrieval cues depends on how the to-be-retrieved item was encoded (Tulving & Thomson, 1973). For example, if visual object representations are episodically constructed into a larger scene representation with other items, visual memory performance should be affected by the representation of other items. In other words, memory performance should be higher when the encoding context is reinstated at test. However, if visual object representations are represented independently, the memory for each item should not be affected by others.

Several studies have revealed that multiple items are episodically organized in visual short-term memory (VSTM) rather than represented alone (Jiang, Olson, & Chun, 2000; Vidal, Gauchou, Tallon-Baudry, & O'Regan, 2005; Zimmer, 1998). For example, Zimmer (1998) used a delayed matching task to investigate whether changing the spatial configuration of elements (e.g., triangular or linear arrays) from study to test affects VSTM for geometric figures and words. He found that participants' recognition performance was better if the spatial configuration at test was identical to the one at study. The results implied that spatial configuration is co-represented with individual elements so that spatial congruence between the study and test stimuli can affect test performance.

Similarly, Jiang et al. (2000) manipulated contextual information in a VSTM task and found that change detection performance of object color, shape, and location was impaired when the spatial configuration of array elements changed from study to test. Thus, Jiang et al. proposed that VSTM stores relational information between individual items, and the relational processing is mediated by the global spatial configuration. Specifically, Jiang et al. argued that the representation of the configuration is purely spatial, coding object locations but not necessarily information about object identities.

More recently, studies using real world objects have further demonstrated that objects in scenes are episodically bound to scene context in memory (Hollingworth, 2006a, 2007). For example, Hollingworth (2006a) asked participants to view images of natural scenes and tested their memory for the visual form of a target object presented either within the original scene context or in isolation. The results showed that object memory performance was more accurate when the target objects were presented within the original scene context than when they were presented in isolation. Thus, while viewing scenes, object memory is stored as part of a more comprehensive scene representation since context deletion impairs the retrieval of target properties from memory.

Clearly, items in visual memory can be constructed into a larger representation with other items so that individual items in visual memory can influence each other. The extent to which different aspects of this representation influence memory retrieval, however, remains an open question. According to Jiang et al. (2000), the reinstatement of abstract spatial layout is the key factor supporting memory retrieval regardless of the object identities occupying those locations. However, Hollingworth (2007, experiment 8) asked participants to view images of object arrays and tested their memory for the orientation of a target object, while manipulating background object-position binding either by retaining or trading contextual objects' locations in the test image. Thus, the abstract spatial configuration could be preserved when the binding of contextual objects to locations was disrupted. The results showed that memory performance was reduced when background object-position binding was changed. Therefore, Hollingworth (2007) suggested that spatial contextual representations of scenes maintain not only the abstract spatial configuration but also the object information bound to each location. Maintenance of the abstract spatial layout is not sufficient to assist memory retrieval. Instead, the binding of individual objects to each location within that spatial configuration determines subsequent memory performance.

However, none of these studies permit isolation of nonspatial properties (e.g., object identity) of the context from spatial properties (e.g., object location). For example, Hollingworth (2007) disrupted background object-position binding by changing each object's location within the original spatial configuration. The objects within the scene, however, were identical in the study and test images. Therefore, it is still unclear whether object identity *per se* plays an influential role in subsequent memory retrieval. The goal of the present study was to test this possibility.

In the current study, we examined the influence of spatial and nonspatial contextual information on object visual memory with a change detection task. On each experimental trial, participants viewed a study image of an object array, followed by a test image. In the test image, participants had to identify the orientation of a target object which was either the same as it appeared during initial viewing or changed (mirror reversed). We examined the influence of spatial contextual information either by retaining the surrounding objects' original positions or by placing them in previously unoccupied locations. Furthermore, we examined the influence of nonspatial contextual information either by maintaining or removing the surrounding objects' identities. If visual memory uses both spatial and nonspatial contextual information to benefit memory retrieval and comparison, we should observe independent contextual contributions from spatial and nonspatial information to the detection of orientation change.

Experiment 1

The purpose of Experiment 1 was to investigate the individual influences of spatial and nonspatial contextual information on visual object memory while viewing arrays of objects. The spatial contextual information was specified as the background objects' locations, whereas nonspatial contextual information was specified as the background objects' identities. We were interested in whether participants' change detection performance would be affected by manipulations of these two elements.

Method

Participants—Twenty-five undergraduates (12 males) from North Dakota State University performed an orientation change detection task for course credit. All participants had normal or correct-to-normal vision and were naïve to the purpose of the study.

Apparatus—The stimuli were presented on a 17-inch CRT monitor with a 100-Hz refresh rate. Participants were tested individually in a room with normal interior lighting. The viewing distance was held constant at 90 cm. Stimulus presentation and response collection were controlled by Presentation software (Neurobehavioral Systems, Inc.). Response accuracy for the experimental trials was collected through mouse-clicks.

Stimuli—Eighty objects were selected from the Tarrlab object databank (Michael J. Tarr, Brown University, <http://www.tarrlab.org/>). Each image contained an array of 6 objects. The screen was divided into an invisible 3×4 matrix, and the 6 objects were randomly assigned to six of the 12 possible locations. The entire array subtended $15.2^\circ \times 20.7^\circ$. The background was white. For the test images, the target object was cued by a red square, and its orientation was either the same as it appeared in the initial array or mirror reversed. The other objects surrounding the target object (which hereafter we refer to as “contextual objects”) either stayed the same as in the study image or changed location or identity. In other words, the contextual objects either were presented in the same location or randomly assigned to previously empty locations, and either maintained their identities or had their identities obscured by blurring. Figure 1 shows sample stimuli and manipulations for different conditions.

Procedure and design—Figure 2 shows the sequence of events in each trial. A trial started with a central cross as a fixation point for 1000 ms. Then the study image was displayed for 4000 ms, followed by a 900-ms blank interval. After the interval, a test image appeared until the participant responded. This was followed immediately by feedback concerning the accuracy of the response.

The experiment used a 2 (target’s orientation same, changed) $\times 2$ (contextual objects’ location same, different) $\times 2$ (contextual objects’ identity same, blurred) factorial design. Participants first completed a practice session of 16 trials, 2 in each of the 8 conditions. Following the 16 practice trials, participants completed 160 experimental trials, 20 in each of the 8 conditions. The 160 experimental trials were presented randomly for each participant, in four blocks of 40 trials each.

Results

For each participant in each condition, the mean hit rate when the target changed and the mean false alarm rate when it did not were used to calculate a nonparametric index of sensitivity to change, A' . A' scores vary from .5 to 1.0, where .5 represents chance level and 1.0 represents perfect sensitivity. Table 1 shows the mean A' values in the experimental conditions.

The results of a two-way (contextual objects' location $\times$ contextual objects' identity) ANOVA performed on the mean A' performance revealed a significant main effect of the contextual objects' location. Participants were better able to detect changes to the target object's orientation when the contextual objects' locations remained the same in the test image, $F(1, 24) = 9.13$, $MSE = 0.004$, $p < .01$. There was also a significant main effect of maintaining the contextual objects' identities, with better performance when the contextual objects' identities remained the same in the test image, $F(1, 24) = 4.57$, $MSE = 0.004$, $p < .05$. There was no significant interaction between contextual objects' location and identity, $F(1, 24) < 1$.

Discussion

Experiment 1 showed that the overall change detection performance was well above chance because A' scores ranged from .78 to .84 across all four conditions, demonstrating that detailed object representations can be formed after viewing arrays of objects. In addition, the current results showed that not only the global spatial configuration but also the contextual objects' identities had a significant impact on participants' visual memory performance. Participants' performance was significantly better when either the contextual objects' locations or identities maintained the same in the test image, indicating that the preservation of spatial or nonspatial contextual information facilitated visual memory retrieval and comparison.

In Experiment 1, contextual object identities were changed by blurring the objects, rendering them unidentifiable. In doing so, however, we also changed gross perceptual properties of the array itself. For example, while the study images contained contextual objects that were perceptually dissimilar from each other, the test images contained perceptually similar blurred objects. It is possible that the change in these properties of the array, rather than the change in contextual object properties per se, influenced visual object memory in Experiment 1. In Experiment 2, rather than blurring the contextual objects, those objects were replaced with new objects in the test image.

Experiment 2

Experiment 2 replicated Experiment 1 but altered the nonspatial contextual information by replacing the original contextual objects with new, previously unseen objects. In addition to allowing us to examine the robustness of previous findings, the experiment provided a further test of whether the influence of nonspatial contextual information can be found when gross array properties (e.g., contextual object similarity) are similar in study and test images. If nonspatial contextual information is encoded while viewing object arrays, then similar facilitation effects of detecting the change of target's orientation should appear when the contextual objects remain the same in the test image.

Method

Participants—Twenty-five undergraduates (18 males) from North Dakota State University participated in the experiment for course credit. All participants had normal or correct-to-

normal vision, had not participated in the previous experiment, and were naïve to the purpose of the study.

Apparatus, stimuli, procedure, and design—Experiment 2 was identical to Experiment 1, except that eighty additional objects were selected from an online database (<http://www.photoobjects.net/>) and used to replace the contextual objects in the nonspatial contextual information changed condition. Figure 1 shows sample stimuli and manipulations for different conditions.

Results

Mean A' performance in each of the experimental conditions is displayed in Table 2. A two-way (contextual objects' location $\times$ contextual objects' identity) ANOVA performed on the mean A' performance revealed significant main effects for contextual objects' location, $F(1, 24) = 5.54$, $MSE = 0.003$, $p < .05$, and for contextual objects' identity, $F(1, 24) = 5.68$, $MSE = 0.006$, $p < .05$. The interaction between contextual objects' location and identity was not significant, $F(1, 24) < 1$.

Discussion

The results of Experiment 2 were consistent with Experiment 1. The overall A' scores were high across the four conditions, demonstrating that observers retain an accurate memory for the visual properties of the objects they view. However, participants' change detection performance was significantly decreased when contextual objects were either moved to previously unoccupied locations or replaced with new objects. Therefore, Experiment 2 replicated and strengthened the findings of Experiment 1, revealing that both spatial and nonspatial contextual properties of an object array can be encoded into visual memory and used to modulate the subsequent retrieval of target information.

General Discussion

The two experiments reported in the present study were designed to examine the influence of spatial and nonspatial contextual information on visual object memory after viewing arrays of objects. In Experiment 1, we tested participants' visual memory with a change detection task. After viewing the study image, participants had to discriminate whether the orientation of a target object had changed in the test image. In the test image, we manipulated spatial and nonspatial contextual information by changing the locations and identities of the contextual objects that surrounded the target object. Specifically, in Experiment 1 we blurred the contextual objects to obscure the availability of the nonspatial contextual information. The results showed that participants successfully detected orientation changes on a significant proportion of trials. Moreover, participants' change detection performance was significantly better when the contextual objects retained either their positions or identities. In Experiment 2, we replicated the first experiment except that we altered the nonspatial context by replacing the original contextual objects with new, previously unseen objects. Compared to the blurring manipulation in Experiment 1, Experiment 2 used a more stringent manipulation to isolate the influence of nonspatial contextual information on the change detection performance. The results obtained from

Experiment 2 were consistent with the findings of Experiment 1, suggesting that both spatial and nonspatial contextual information are beneficial for supporting visual memory retrieval and comparison.

When we see an object in a scene or among other objects, how is it represented in visual memory? Our results showed that change detection performance was well above chance, demonstrating that detailed object representations can be formed after viewing complex visual stimuli such as arrays of objects. The results were consistent with previous studies showing that robust object representations can be maintained after scene viewing (e.g., Hollingworth, 2005; Hollingworth & Henderson, 2002). Furthermore, our results were in accordance with previous studies that have demonstrated that object representations are wedded to a more extensive scene representation (e.g., Hollingworth, 2006a, 2007). Therefore, an inconsistency in contextual information between study and test images produces modest, but reliable, impairment in visual memory for a target object.

Importantly, we also found individual influences of spatial and nonspatial contextual information on visual memory performance. When the global spatial configuration changed in the test image, visual memory performance was significantly reduced regardless of whether the contextual objects' identities were the same or different. This is consistent with the spatial configuration hypothesis (Jiang et al., 2000), which suggested that spatial contextual information is essential for the organization of visual memory. In fact, spatial contextual information is important for various cognitive activities, such as scene recognition and visual search (Chun, 2000; Chun & Jiang, 1998; Oliva & Torralba, 2006). For example, Chun and Jiang (1998) developed the contextual cueing paradigm and found that visual attention can be quickly guided to a target when the spatial configuration of the search array has been repeated. Furthermore, as long as the abstract spatial configuration is maintained, contextual cueing effect can be observed even when distractors' identities change (Chun & Jiang, 1998). Similar to Jiang et al. (2000), Chun and Jiang suggested that contextual cueing may not need detailed identity information.

While spatial configuration information appears to play an important role in supporting visual memory, the extent to which purely nonspatial contextual information plays a role has been less clear. The results of the present study address this fundamental gap in the literature. Surprisingly, when we changed nonspatial aspects of the context either by blurring contextual objects' identities (Experiment 1) or by replacing contextual objects with novel objects (Experiment 2), visual memory performance was also impaired regardless of whether or not the contextual objects' location (which forms the global spatial configuration) was the same. Our results are similar to prior studies of object-position binding which demonstrated that spatial configuration coding is not the only key to link items together in our visual memory. Object information bound to locations within that configuration is important as well (Hollingworth, 2007; Hollingworth and Henderson, 2002).

Our results extend this previous work, however, by showing that, regardless of the spatial configuration, contextual object identities provide support for visual memory retrieval and comparison. When the contextual objects' identities were retained in the test image,

participants had better memory performance. While the binding of objects to particular locations is important for generating an episodic scene representation (Hollingworth, 2007), the nonspatial information originated from the contextual objects also contributes independently to memory retrieval.

It is important to note that nonspatial contextual information includes various visual features, such as color and shape, in addition to semantic object properties, such as identity. When we altered the nonspatial context by blurring the contextual objects' identities or replacing them with other objects, different visual features of the contextual objects also changed. Our current work does not permit an investigation of the influences of all dimensions of nonspatial context. Future studies should be conducted to investigate how various aspects of nonspatial context modulate visual object memory.

Conclusions

We have the ability to form and maintain detailed object representations during and after viewing scenes. Furthermore, the reinstatement of the encoding context benefits subsequent visual memory performance. To study the question of how contextual information supports visual memory retrieval, most studies focus on the importance of spatial configuration. However, the extent to which purely nonspatial contextual information (e.g., object identity) plays a role has been less clear. By using object arrays, our results suggest that nonspatial contextual information is essential for effective retrieval of visual object memory as well. Furthermore, object identities need not be bound to locations in order to support visual memory retrieval. Thus, when encoding a variety of objects into memory, two types of information are encoded. One is the spatial contextual information which comes from the spatial configuration formed by items in the display. The other is the nonspatial contextual information which comes from properties of objects in the display. The two types of contextual information both benefit visual object memory retrieval and comparison.

Acknowledgments

This project was supported by Grant Number 1P20 RR020151 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). The project was also supported by the National Science Foundation under Grant Number BCS-0443998. We thank two anonymous reviewers for providing helpful comments to improve the manuscript.

References

- Castelhano MS, Henderson JM. Incidental visual memory for objects in scenes. *Visual Cognition*. 2005; 12:1017–1040.
- Chun MM. Contextual cueing of visual attention. *Trends in Cognitive Sciences*. 2000; 4:170–178. [PubMed: 10782102]
- Chun MM, Jiang Y. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*. 1998; 36:28–71. [PubMed: 9679076]
- Hollingworth A. The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2005; 31:396–411.
- Hollingworth A. Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2006a; 32:58–69.

- Hollingworth A. Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*. 2006b; 14:781–807.
- Hollingworth A. Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*. 2007; 33:31–47. [PubMed: 17311477]
- Hollingworth A, Henderson JM. Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*. 2002; 28:113–136.
- Hollingworth A, Williams CC, Henderson JM. To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*. 2001; 8:761–768. [PubMed: 11848597]
- Jiang Y, Olson IR, Chun MM. Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2000; 26:683–702.
- Mitroff SR, Simons DJ, Levin DT. Nothing compares 2 views: Change blindness can occur despite reserved access to the changed information. *Perception and Psychophysics*. 2004; 66:1268–1281. [PubMed: 15813193]
- Oliva A, Torralba A. Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*. 2006; 155:23–36. [PubMed: 17027377]
- Rensink RA, O'Regan JK, Clark JJ. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*. 1997; 8:368–373.
- Simons DJ. In sight, out of mind: When object representations fail. *Psychological Science*. 1996; 7:301–305.
- Tulving E, Thomson DM. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*. 1973; 80:352–373.
- Vidal JR, Gauchou HL, Tallon-Baudry C, O'Regan JK. Relational information in visual short-term memory: the structural gist. *Journal of Vision*. 2005; 5:244–256. [PubMed: 15929649]
- Zimmer HD. Spatial information with pictures and words in visual short-term memory. *Psychological Research*. 1998; 61:277–284. [PubMed: 9870295]

Figure 1.

Sample study and test images. All test images depict trials in which the target object has changed orientation. In the experiment, all stimuli were presented in color. (a) a sample test image used in the same location and same identity condition (b) a sample test image used in the different location and same identity condition (c) a sample test image used in the same location and blurred identity condition (d) a sample test image used in the different location and blurred identity condition (e) a sample test image used in the same location and different identity condition (f) a sample test image used in the different location and different identity condition

Figure 2.

The sequence of events within a trial. In the experiment, all stimuli were presented in color.

Table 1

Mean A' data in Experiment 1 as a function of nonspatial (contextual objects' identity) and spatial (contextual objects' location) contextual information.

Contextual Objects' Identity	Contextual Objects' Location	
	Same	Different
Same	0.84	0.80
Blurred	0.81	0.78

Table 2

Mean A' data in Experiment 2 as a function of nonspatial (contextual objects' identity) and spatial (contextual objects' location) contextual information.

Contextual Objects' Identity	Contextual Objects' Location	
	Same	Different
Same	0.82	0.78
Different	0.77	0.76