Title
Accumulation of Visual Memory for Natural Scenes: A Medium-Term Memory?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Melinger, Alissa
im Walde, Sabine Schulte

Publication Date
2005

Peer reviewed

Accumulation of Visual Memory for Natural Scenes: A Medium-Term Memory?

David Melcher (dmelcher@brookes.ac.uk)
Department of Psychology, Oxford Brookes University
Gipsy Lane, Oxford OX3 0BP, United Kingdom

Abstract

We studied visual memory for objects in natural scenes. Participants viewed photographs or pictures for periods of 1 to 20 seconds, and were then asked questions about the color, location, and identity of objects that they had seen, as well as given recognition tests. Performance improved as a function of display duration for all question types and for objects of both central and peripheral interest. On some trials, previously viewed stimuli, which had been shown 4 to 6 trials earlier with no subsequent memory test, were repeated to see if performance continued to improve across separate presentations. There was no loss of information across retests, such that memory for a display shown for 10 and then 5 seconds was equal to performance after a single trial of 15 seconds. Overall, memory performance exceeded the capacity and duration limits of short-term or working memory, supporting the idea of a medium-term visual scene memory.

Introduction

The ability to remember the contents of visual displays is critical for a variety of real-life tasks. Traditionally, this process has been characterized as "storage" of a few items in visual short-term memory (VSTM) followed by either information loss or transfer into long term memory (LTM). Visual short-term memory is generally considered to contain only about four items (Cowan, 2001; Vogel & Machizawa, 2004), although the exact capacity remains a matter of debate (Brockmole et al., 2002; Rensink, 2000).

Unfortunately, such a limited short-term memory system would make learning about the world extremely difficult. For example, imagine that the typical view of a cluttered office desk contains at least three to four objects. It would follow that if someone looks at that desk, then it would cause the person to forget everything they had just seen unless it had already been stored in long-term memory. The dichotomy between short and long-term memory implies that people either spend a great deal of their waking life rehearsing information to keep it from fading away, or that they must be constantly filling their long-term memory with information that may later turn out to be useless.

Real-life memory, as opposed to memory for relatively meaningless stimuli (letters, numbers, colored shapes), may not fit neatly into the storage model of memory. Objects in naturalistic scenes, for example, are linked with the scene context in memory rather than stored as a discrete set of objects (Melcher, 2001; Melcher & Kowler, 2001). Moreover, visual memory for objects and other details in photographs exceeds the three to four item memory limit after just a few seconds of viewing (Melcher, 2001; Melcher & Kowler, 2001; Tatler et al., 2003). The build-up and persistence of memory does not, however, depend on transfer to long-term memory, since memory for scenes did not persist across testing on separate days.

The finding that memory for objects in a room exceeds both the capacity and duration of traditional short-term visual memory, while not necessarily finding a permanent place in long-term memory, suggests the existence of a "medium-term" working memory for visual scenes (Melcher, 2001; Pierrot-Deseilligny et al., 2002). There is considerable support for the idea that useful visual information accumulates across separate glances. For example, Tatler et al (2003) found that performance in answering questions about the visual details of a photograph increased over a 10 second interval. Henderson and colleagues have reported accurate recognition memory for the visual features of objects shown several minutes earlier (Henderson et al., 2003). In addition, there is evidence that traditional measures of short-term memory underestimate the visual processing of a larger number of items in the brain (VanRullen & Koch, 2003)

The goal of the present study was to examine the build-up and persistence of working memory for different visual attributes in a complex image. Natural scenes contain a myriad of visual information, including color, texture, shape, absolute and relative object location, object identity, overall spatial layout and scene gist. Recent studies suggest that different visual features may have different rates of memory accumulation and decay (Melcher & Morrone, 2003; Tatler et al., 2003). Such differences in memory for particular visual features might, in fact, help to explain the conflicting reports in the visual memory literature.

A second, and related, issue is the role of scene-related salience in the guidance of attention and subsequent memory for objects in a scene. In studies of change detection, alterations to the scene are noticed more quickly when they occur to an item of "central interest", rather than one that is peripheral to the gist or general meaning of the picture (Rensink, 2000). This raises the question of which types of items benefit from extended viewing of a scene. One possibility might be that salient items would show more improvement because they are viewed earlier and more often, increasing the chance that their details would be in memory, while peripheral items would tend to be ignored On the other hand, peripheral items might show the greatest benefit from longer presentations, compensating for a lack of earlier attention, while central items would lose their initial dominance in memory.

Our earlier studies showing memory accumulation across separate views of the same display did not differentiate between central and peripheral interest items (Melcher, 2001). By comparing memory performance as a function of time and for repeated stimuli, we can examine the way that memory for different attributes of items of central and peripheral interest accumulates and persists in memory.

Experiment 1

Methods

Participants

20 participants participated in the first experiment, and 10 participants in the second experiment. All participants gave informed consent and were paid a small fee for their time.

Stimuli

The stimuli were displayed on a 21" SONY high-resolution monitor viewed from 60 cm and controlled by MATLAB and the Psychophysics toolbox (Brainard, 1997). Each image was approximately 20 degrees of visual angle in height and width. Images included drawings of realistic scenes, reproductions of paintings and photographs of both indoor and outdoor natural scenes. About one-half of the images contained people.

The memory test contained a series of written questions on the screen about specific objects in the scene and a list of three choices. Examples include: (1) What color is the tablecloth? peach, white, or blue (2) Where is the teacup? bottom right, bottom left or center (3) What food is in the middle plate of the three-tiered plate stand? cake, sandwiches or sausage rolls?

The questions were written by the author and three research assistants. Questions were divided into those regarding "central" or "peripheral" interest in the scene (Rensink, 2000). The group collaborated in choosing whether to define a certain question as regarding a central item or not, and in case of disagreement the question was discarded and a new one written.

Three-alternative forced choice tests were also given for an object from the scene, with all items sharing a conceptual category but differing visually (see Figure 1).

Procedure

The first study extended the memory accumulation methodology used in previous studies (Melcher, 2001; Melcher & Kowler, 2001) to photographs and drawings shown for longer periods of time (up to 20 seconds). On each trial, the stimulus display was presented for a time period of 5, 10 or 20 seconds, run in separate blocks. After each trial, participants were either given a memory test or instructed to continue on to the next trial. Participant responses were given by keyboard press and recorded for later analysis. The order of conditions was randomized across observers.

Critically, some images were not followed by a memory test after the first presentation but only after being shown a second time. This manipulation served to measure whether the memory accumulation across repeats of the same stimuli found previously for object recall (Melcher, 2001) extended also to questions about specific object attributes. Re-test trials were shown 4 to 6 trials after the initial display of that stimulus.

(1) (2) (3)

Figure 1. Three alternative forced choice recognition test

Results

The first experiment tested the accumulation of memory for visual detail. Overall, performance improved for longer views ($F[1,647] = 8.30$, $p < 0.001$) and was enhanced for stimuli that were shown more than once before the questions were asked ($F[1,756] = 6.76$, $p < 0.02$) Figure 2 shows the average percent correct performance for all types of questions as a function of total viewing time. The solid squares show performance after 5, 10 or 20 seconds on the non-repeat trials. Open circles show re-test trials, in which the memory test was given only after the second time that the picture had been shown. The leftmost open circle shows performance after seeing a display twice for 5 seconds each time on repeat trials, for a total viewing time of 10 seconds. The rightmost open circle shows percent correct response after seeing a 10 second and a 5 second display for a total viewing time of 15 seconds. On half of those trials, the 10 second duration trial preceded the 5 second trial, while the other half of trials contained the opposite pattern.

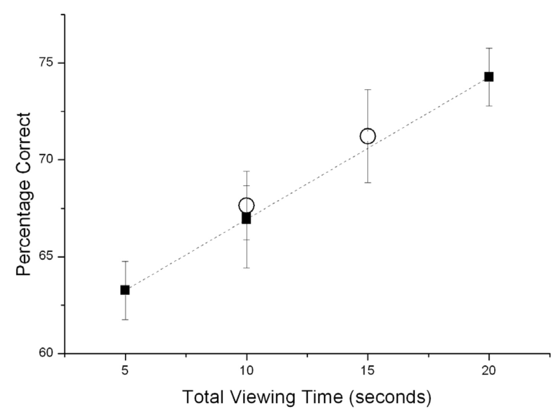

Figure 2: Average percent correct performance on the memory test in Experiment 1 as a function of total viewing time. Filled symbols show performance after a single, continuous presentation , while open symbols show performance after a repeated stimulus.

As shown by the dotted line, performance improved linearly as a function of total viewing time, with no difference between re-test trials and normal trials of the same total viewing time, consistent with our previous study with free recall (Melcher, 2001).

Performance was analyzed separately for each of the different question types, with data divided into trials containing items of central and peripheral interest (see Methods for more details). Overall, performance was superior for questions about items judged as of "central interest" (Figure 3, filled circles), maintaining a consistent advantage of 10%-15% compared to items of less salience (Figure 3, filled squares). The same trend was found for repeat trials (open circles and squares).

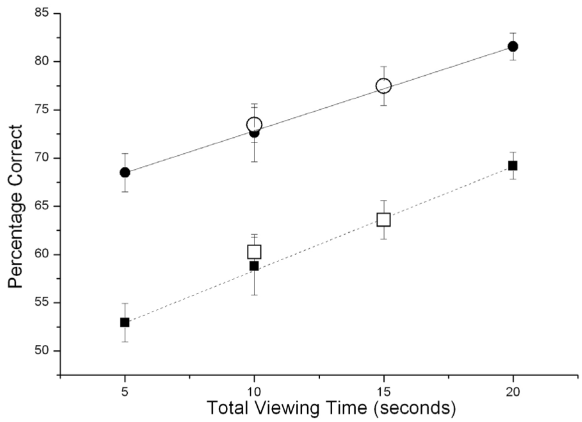

Figure 3: Performance in Experiment 1 divided into questions about items of central interest (circles, solid line) and peripheral interest (squares, dotted line)

Experiment 2

Methods

The second experiment considered both memory accumulation (from 1 to 10 second display durations) and the duration of memory persistence across a delay period of up to one minute. The experiment was similar to the first experiment, with the addition of a delay period between the presentation of the stimulus and the test.

The first two conditions consisted of displays shown once for 1 or 10 seconds. The third block of 20 trials contained a 10 second display of the stimulus, followed by a delay of 10 seconds and then a 1 second re-display of the stimulus. The fourth block extended the delay to 60 seconds. During the delay period, a written paragraph was presented on the screen and participants were instructed to silently read the paragraph throughout the delay period, repeating if necessary. The reading task was designed to occupy visual, verbal and conceptual working memory. The order of conditions was randomized across observers.

Results

There was a general improvement in memory test performance as the display duration increased from 1 second (Figure 4: leftmost bars in each cluster) to 10 seconds (Figure 4, bars second to left) for color, location and identity questions, consistent with the results of the first experiment. Recognition performance (leftmost cluster of bars) also improved as display duration increased. The smallest improvement as a function of total viewing time was found for color.

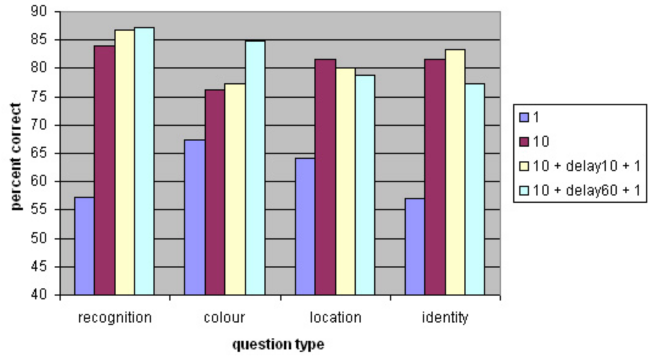

Figure 4: Percentage correct performance as a function of question type and display duration condition (Experiment 2)

There was no effect of the delay period on performance for color, location, identity or shape recognition tasks for the maximum interval tested. Performance on trials containing a ten second stimulus, followed by a delay with reading and then a further one second display was similar to that found after a single ten second trial.

General Discussion

Overall, the results show that memory for the visual detail of natural scenes accumulates over time and across separate glances. There results are in sharp contrast to suggestions that our visual memory representation is relatively sparse (e.g. Rensink, 2000; Todd & Marois, 2004), but add support to the idea that working memory for natural scenes can accumulate information over a period of minutes (Melcher, 2001; Henderson et al., 2003; Tatler et al., 2003).

Overall, memory for items judged as "central" to the gist or meaning of the picture was better than that found for items of peripheral interest. This is consistent with findings that changes to these items are noticed more easily (Rensink, 2000). The rate of accumulation of memory across the 20 second time period, however, was similar for central and peripheral interest items, which did not support the hypothesis that memory representation might be built up initially of central items and later with less important details.

The present results provide confirmation of earlier findings on the accumulation of memory across separate views of a picture (Melcher, 2001; Melcher & Kowler, 2001). As in the earlier studies, the rate of memory accumulation for natural scenes depended on total viewing time across repeated presentations of the same stimulus. Here, a different task (recognition versus recall), different stimulus set, and a larger range of durations was used, but none of these factors affected the basic finding.

It is not easy to accommodate these findings within the traditional stage/storage model of memory, nor is it clear how a single visual-spatial sketchpad could support the simultaneous storage and accumulation of several pictures at once. One possibility is that format of the scene representation takes into account the inherent structure of the environment, rather than dividing information into a limited number of unrelated information chunks.

Previously, physiological studies have given some evidence for a form of visual-spatial memory that persists over periods of minutes (McGaugh, 2000; Pierrot-Deseilligny et al., 2002). Interestingly, recent studies on visual short-term memory, as measured by change detection, have implicated the frontal-parietal attention network (Todd & Marois, 2004; Vogel & Machizawa, 2004). One possibility, based on studies of animal physiology (Murray & Bussey, 1999; Ranganath & D'Esposito, 2001) and human brain imaging (Pierrot-Deseilligny et al., 2002; Ranganath et al., 2004), is that visual scene memory, unlike VSTM, depends critically on areas of temporal cortex involved in recognizing and remembering objects and complex scenes.

The finding that memory builds up from fixation to fixation may help to explain why demonstrations of failures of memory in the laboratory are often so surprising. In normal life, the brain keeps in mind relevant information for cognition and action, building up a useful but ultimately disposable representation of the immediate environment.

Acknowledgments

This project was supported by grants from the British Academy (SG37385) and the Royal Society (2003/R2).

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spaial Vision, 10*, 433-436.

Brockmole, J. R., Wang, R. F., & Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception & Performance, 28*, 315-334.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavior Brain Science, 24*(1), 87-114.

Henderson, J. M., Williams, C. C., Castelhano, M. S., & Falk, R. J. (2003). Eye movements and picture processing during recognition. *Perception & Psychophysics, 65*(5), 725-734.

McGaugh, J. L. (2000). Memory—a century of consolidation. *Science, 287*(5451), 248-251.

Melcher, D. (2001). Persistence of visual memory for scenes. *Nature, 412*(6845), 401.

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research, 41*, 3597-3611.

Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience, 6*, 877-881.

Murray, E. A., & Bussey, T. J. (1999). Perceptual- mnemonic functions of the perirhinal cortex. *Trends inCognitive Science, 3*, 142-151.

Pierrot-Deseilligny, C., Muri, R. M., Rivaud-Pechoux, S., Gaymard, B., & Ploner, C. J. (2002). Cortical control of spatial memory in humans: the visuooculomotor model. *Annals Neurology, 52*, 10-19.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research, 40*, 1469-1487.

Tatler, B. W., Gilchrist, I. D., & Rusted, J. (2003). The time course of abstract visual representation. *Perception, 32*, 579-592.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751-754.

VanRullen, R., & Koch, C. (2003). Competition and selection during visual processing of natural scenes and objects. *Journal of Vision, 3*(1), 75-85.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748-751.