Journal of Experimental Psychology:
Human Perception and Performance
2017, Vol. 43, No. 7, 1343–1358

0096-1523/17/$12.00 http://dx.doi.org/10.1037/xhp0000392

Embodied Memory Allows Accurate and Stable Perception of Hidden Objects Despite Orientation Change

Jing Samantha Pan
Sun Yat-sen University

Ned Bingham
Cornell University

Geoffrey P. Bingham
Indiana University

Rotating a scene in a frontoparallel plane (rolling) yields a change in orientation of constituent images. When using only information provided by static images to perceive a scene after orientation change, identification performance typically decreases (Rock & Heimer, 1957). However, rolling generates optic flow information that relates the discrete, static images (before and after the change) and forms an embodied memory that aids recognition. The embodied memory hypothesis predicts that upon detecting a continuous spatial transformation of image structure, or in other words, seeing the continuous rolling process and objects undergoing rolling observers should accurately perceive objects during and after motion. Thus, in this case, orientation change should not affect performance. We tested this hypothesis in three experiments and found that (a) using combined optic flow and image structure, participants identified locations of previously perceived but currently occluded targets with great accuracy and stability (Experiment 1); (b) using combined optic flow and image structure information, participants identified hidden targets equally well with or without 30° orientation changes (Experiment 2); and (c) when the rolling was unseen, identification of hidden targets after orientation change became worse (Experiment 3). Furthermore, when rolling was unseen, although target identification was better when participants were told about the orientation change than when they were not told, performance was still worse than when there was no orientation change. Therefore, combined optic flow and image structure information, not mere knowledge about the rolling, enables accurate and stable perception despite orientation change.

Public Significance Statement

Previously in the *Journal of Experimental Psychology: Human Perception and Performance*, Pan, Bingham, and Bingham (2013) introduced the notion of embodied memory as a new way to understand how motion-generated optical structure and static optical structure (images) interact to provide visual information about the 3D structure of surfaces that is both effective and stable. These authors now show the essential relevance of this analysis to the perception of spatial relations in the context of postural changes in head orientation that are intrinsic to human viewing. Performance with and without such orientation changes was found to be equivalent when both motion-generated and static optical structure were available, but not otherwise. This implies that the embodied memory theory is essential to understand how human visual perception normally works.

Keywords: optic flow, image structure, orientation change, embodied memory

Optic flow provides effective information about the 3D spatial relations among surfaces and an observer. The problem is that optic flow ceases to remain available when the motion that generates it stops. Embodied memory provides a solution to this problem (Pan, Bingham, & Bingham, 2013). The image structure from surfaces surrounding an observer remains available when motion stops and is intrinsically related to previously available optic flow because the flow is of that image structure. The image

---

Jing Samantha Pan, Department of Psychology, Sun Yat-sen University; Ned Bingham, School of Electrical and Computer Engineering, Cornell University; Geoffrey P. Bingham, Department of Psychological and Brain Sciences, Indiana University.

This study is supported by the National Natural Science Foundation of China (General Program, Grant 31571116) and by Sun Yat-sen University (Grant 16wkpy29). The funders have no role in study design, data collection, or preparation of this abstract. We thank Hongyuan Wu for assisting data collection and making the experimental demo video.

Correspondence concerning this article should be addressed to Geoffrey P. Bingham, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47404. E-mail: gbingham@indiana.edu

structure that has been hypothesized to provide stability to the optic flow information by acting as embodied memory.

Visual perception is active. Observers actively look around, turning their heads and eyes to sample and detect optical information specifying surrounding environmental structures. The head motions often result in changes of retinal image orientations. In this work, we investigate perception of hidden object locations with the presence of changes in image orientations.

In image-based approaches to the study of visual perception, the visual system uses retinal image structure to form representations of objects and their locations. Image structure has been used in the study of vision since the work of Johannes Kepler in the 16th century (Hergenhahn, 2008). Image structure refers to static optical structures or patterns that are detectable by the eye and may contain information about surrounding spatial relations (e.g., in Marr & Hildreth, 1980, and in Koenderink, 1984). Examples of image structure include colors, shades, edges, contours, or shapes, which, in combination, may be used as static cues for visual recognition. In this image-based approach, motions of the visual targets or the observers have been treated as hindrance to perception because they perturb retinal images. For example, visual object recognition studies attempted to explain how objects are recognized based on the distinct retinal images they project before and after some rotation in depth, and found that the more complex an object and the more it rotates in depths, the harder it is to recognize (Riesenhuber & Poggio, 2000; Tarr & Bülthoff, 1998; and Tarr & Vuong, 2002). Another line of research reveals that objects with rotated retinal images were harder to recognize than those with upright images when retinal images were subjected to orientation change in the frontoparallel plane as a result of either a rotating scene or a rotating observer (e.g., Rock & Heimer, 1957; see Jolicoeur, 1990, for a review). Taken together, both visual object recognition studies and orientation change studies suggest that large-scaled motions lead to large differences between retinal images before and after the motions (whether rotation in depth or rolling in the frontoparallel plane), and hence make image-based visual recognition less effective.

However, from an ecological perspective, motion does not impede perception; instead, it provides important information for perceiving objects, surfaces, and events. In an environment populated by opaque surfaces, light is reflected from, and therefore deterministically structured by, surfaces surrounding the observer. The structured light available to an observer is described as the optic array. Motions of the observer and/or surfaces in the environment yield continuous and lawful changes in the optic array, known as *optic flow* (Gibson, 1979/1986; see also Nakayama & Loomis, 1974, for a quantitative formulation of optic flow). Optic flow is generated by motion in the environment and thus corresponds to and specifies these motions. When an individual locomotes through the environment, the motion generates a global pattern of optic flow that enables the perception of heading (e.g., Warren & Hannon, 1990), control of steering (e.g., Li & Warren, 2000, 2002; Wann & Land, 2000; Wilkie & Wann, 2006), or control of braking (e.g., Anderson & Bingham, 2010, 2011; Fajen, 2005a, 2005b, 2008; Yilmaz & Warren, 1995). Furthermore, when observers or environmental objects are going though large-scale motion, optic flow also provides immediate and powerful information about the 3D structure of the surroundings, including metric shapes and the layout of surfaces in cluttered terrain (Domoni & Caudek, 2003; Todd, 1995). This is called structure from motion (SFM; Koenderink & van Doorn, 1991; Tittle & Braunstein, 1993; Todd & Bressan, 1990). However, optic flow varies in quality with the relative speeds of motion. The strength of optic flow is directly proportional to the relative speed of object movements and weakens as motion speed decreases until, as motion stops, it becomes unavailable. Thus, although motion-generated optic flow is strong in specifying 3D structure, it is lacking of stability. Image structure, on the other hand, is weaker in its ability to specify 3D structures, but it is persistent. Image-based vision relies on cues (such as image size, texture gradients, or height in the visual field) to deduce spatial relations based on experience. It reduces the 3D dynamical environment to flat and static snapshots from which observers extract useful cues, resorting to their working memory or mental rotation to undo the rotation or orientation change, to identify objects (Jolicoeur, 1990). Thus, identifying complex objects before and after some spatial transformation is challenging if the continuous motion is unavailable and image structure is the only available information. Despite its weakness in specifying depth relations, however, image structure is stable. It remains as long as the objects remain visible.

Pan, Bingham, and Bingham (2013) hypothesized that embodied memory incorporates the two sources of optical information, and explains how stable and accurate perception might be formed with a combination of optic flow and image structure, allowing them to compensate for one another's weaknesses. Specifically, in normal viewing environments, when opaque surfaces move, both optic flow and image structure information become available. They are intrinsically related in respect to the layout of surfaces from which they are projected, because optic flow carries one structured image into the next structured image (one image "flows" to the next). In other words, optic flow and image structure specify the same events but play different roles in the perception process. In the process of perception, the relation between optic flow and image structure, in part, could be cast as a calibration of image based information about 3D structure by the more powerful optic flow information. (Just as knowledge of results provides feedback from visually guided actions that calibrates the perceptual information used to guide those actions to allow them to be accurate, optic flow specifies spatial layout and can be used to calibrate image structure.) Optic flow specifies the changes in 3D spatial structure that relates sequential images. Once the optic flow has ceased, the calibrated image structure remains, and thus helps preserve the information provided by the optic flow. This combination promotes the effectiveness of perception because information provided by transient optic flow is offloaded to external stable image structure information, rather than being kept in the head. Thus, the calibrated image structure becomes embodied memory for spatial layout that has been specified by optic flow. It is embodied because the image structure is projected from the substantial surfaces of the environment in which the (substantial or embodied) observer is situated and, therefore, to which the observer is related.

Empirically, embodied memory has been shown to yield accurate and stable perception of hidden objects (Pan et al., 2013), metric shape (Lee & Bingham, 2010), and blurry events (Pan & Bingham, 2013). In all of these studies, the experimenters manipulated the availability and/or quality of optic flow and image structure and tested the stability and accuracy of perception-based performance. In the study of hidden object perception (Pan et al., 2013), participants first observed and learned locations of multiple target objects. Then, targets were progressively occluded until they went completely out of view. After some delay, participants identified locations of the hidden objects. The results showed that participants successfully identified locations of hidden objects in conditions in which optic flow was ongoing, or when optic flow and image structure coexisted. When there was only optic flow information, identification performance was accurate (100% identification) when the visual scene was moving, but as soon as the motion stopped, performance dropped to, on average, less than two of 15 items. Performance was not temporally stable. In contrast, with both optic flow and image structure and when image structure (previously calibrated by optic flow) was continuously available, large numbers of targets (more than eight of 15) were identified with response delays of up to 25 s. Performance, in this case, was both accurate and stable.

Lee and Bingham (2010) investigated whether large perspective changes (>45°) would enable seated observers to perceive metric shape and use the information to guide accurate feedforward reaches-to-grasp with optic flow and image structure information. Multiple cylinders with various aspect ratios were placed on a turntable. Participants in this experiment viewed the objects as the turntable was rotated and then performed reaches-to-grasp. Participants' action responses (e.g., how wide a hand opened when about to grasp) were used as a measure of shape perception. The authors found that large perspective changes (>45°) enabled observers to perceive metric shape and guide accurate feedforward reaches-to-grasp. Without the large perspective changes, reaches-to-grasp were inaccurate. Participants were able to perform accurate reaches-to-grasp both immediately after viewing the rotating objects, and after 15-s delays during which image structure, but not optic flow, was available. Thus, optic flow information was necessary to enable accurate shape perception (and, thus, accurate reaches-to-grasp), and this accuracy of performance remained with persistent image structure after optic flow had ceased.

Pan and colleagues applied this embodied memory approach to event perception with low vision (Pan & Bingham, 2013). They first showed participants individual blurry pictures taken from 20 frame videos depicting daily activities, and then showed them the sequence of blurry frames of the videos with motion masks between the frames. Then, participants were shown the blurry videos (without the motion masks). Finally, they showed participants the individual blurry pictures again. In each phase, participants reported what activities they saw from the blurred stimuli. The participants failed to identify events using static images, even in sequence, but succeeded when these images were played in videos that yielded motion. Thereafter, events were perceivable using static blurry images, which had previously failed to allow perception. Weakened image structure in low vision was calibrated by unimpaired optic flow. The static images preserved the spatiotemporal information in optic flow. This synergy allowed low-vision observers to perceive, remember, and potentially interact with events.

In the current study, we extended the scope of the embodied memory hypothesis and tested its application in perceiving hidden objects with orientation change. Orientation change is a particular case of a more general problem, object constancy. It has been studied using identification tasks, in which observers view a static stimulus (a letter, an object, or a scene) and then identify it in another rolled static view. Rock and Heimer (1957) found that when upright observers identified familiar objects from images that were turned by 90° in the frontoparallel plane, the rate of correct identification was more than 50% lower than when upright observers identified these objects from upright images. When naming objects, response time increased when the objects' orientation change increased from 0° to 180° (Jolicoeur, 1985). Using a slightly different task, Simons and Wang (Simons & Wang, 1998; Wang & Simons, 1999) showed that change detection performance was affected by the relative orientation between an observer and the object display. Specifically, change detection dropped when the array of objects rotated (about 50°) and the observer remained still compared with when the observer and the objects both remained stationary.

The existing studies on the effects of orientation change only provided image structure information before and after motion. With image structure alone, observers have to encode and preserve the original stimuli in the working memory and mentally rotate the postrotation stimuli to undo the orientation change and match them to the remembered prerotation stimuli (Jolicoeur, 1990). However, orientation change involves a continuous motion that generates strong optic flow and in natural viewing conditions, both image structure projected from the moving objects and optic flow from the motion would be available. According to the embodied memory hypothesis, optic flow would specify the orientation change (its direction and amount) and simultaneously calibrate the image structures projected by the moving objects. Calibrated image structures would, in turn, preserve the orientation change, making postmotion spatial relations available to an observer so long as the image structure is available. Thus, if an observer continuously experiences both optic flow and image structure when a visual scene rolls, (s)he should be able to identify visual targets equally well with or without the orientation change.

We tested this prediction in three experiments in which participants identified the locations of multiple objects after they became occluded and after the visual scene rigidly rolled in the frontoparallel plane. We modified the paradigm used in Pan et al. (2013) and inspired by the original Kaplan displays (Gibson, 1979/1986, p. 189, Kaplan, 1969). In the Kaplan kinematogram displays, a randomly textured square was perceived to move in front of a background composed of identical random textures. When viewing any discrete frame from the display (that is, a static picture of the square and the background), an observer could only see a single textured surface (that is, the square was invisible). However, when the continuous motion was presented, an observer could immediately see both surfaces, one in front of the other, separated in depth. In the current study, the stimulus display involved two planar frontoparallel surfaces, covered with random texture and separated in depth, one in front of the other. Like in the Kaplan display, the spatial relation of the two identically textured surfaces could only be detected when the surfaces were moving. Targets on the back surface could be seen through cutout windows in the front surface. All windows were bordered with visible black contours, which projected image structure information. The rear surface rigidly translated relative to the unmoving front surface to produce progressive occlusion taking the targets out of view beyond the windows in the front surface. After the targets were occluded and only the black contour windows were visible, the two surfaces

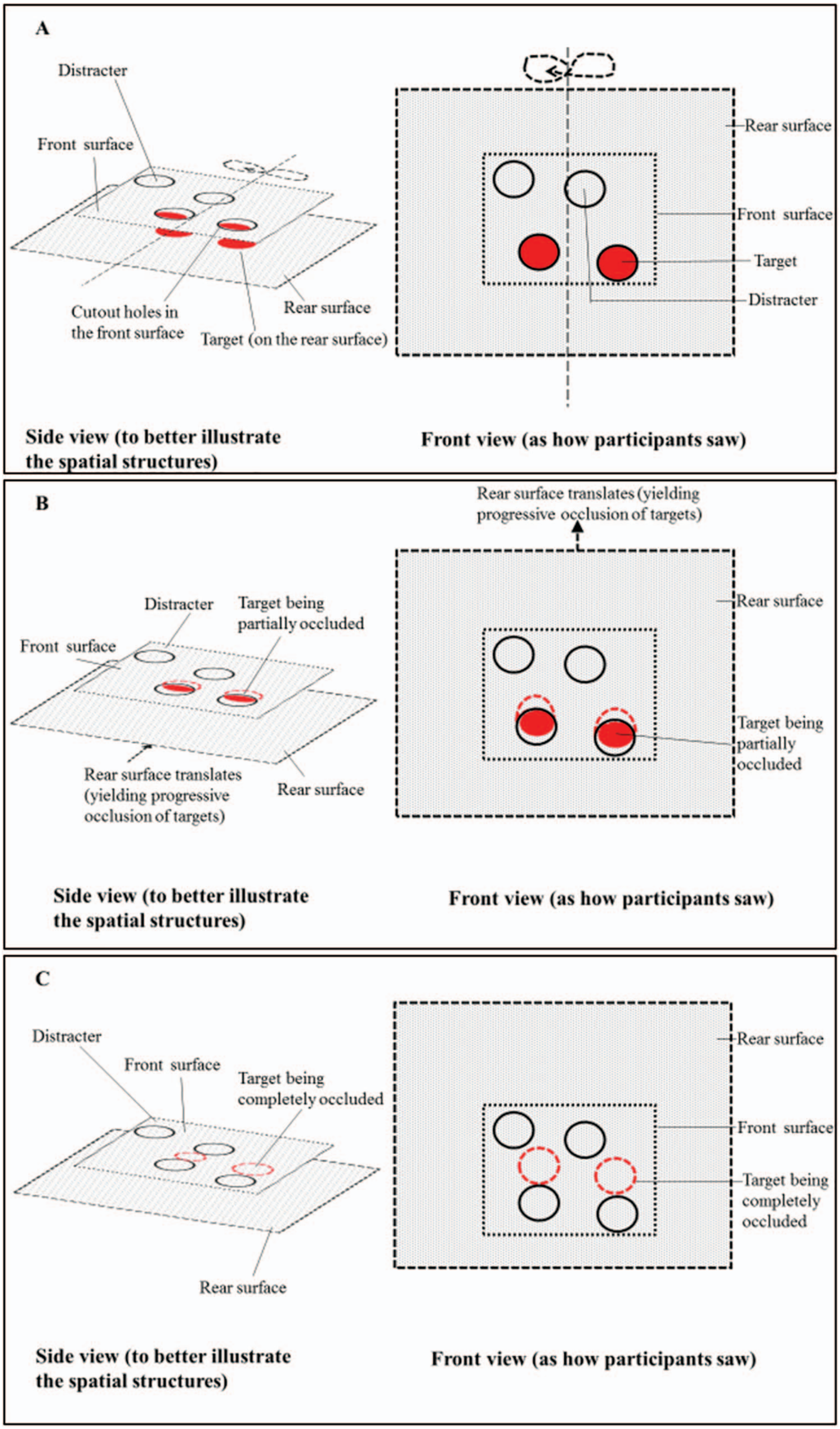

Figure 1 (opposite)

rolled rigidly in a frontoparallel plane and around the line of sight, causing a change of orientation of all visible windows and occluded targets (see the Method section for more detailed descriptions). We tested recall of the locations of occluded targets in conditions without orientation change (Experiment 1, baseline) with continuously available optic flow and image structure information about the orientation change (Experiment 2, continuous change in orientation), and with image structure information (but not optic flow) before and after rolling (Experiment 3, discrete change in orientation). We used a spatial memory task to test participants' perception and retention of object locations. The response task was performed after optic flow had ceased, and entailed the identification of previously perceived, but now occluded, targets.

Note that in a previous experiment using this general paradigm, Pan et al. (2013) showed large changes in performance levels between presentation of both optic flow and image structure compared with presentation of either alone (that is, when only optic flow was presented with no image structure or when only image structure was presented with no optic flow). In the former case, more than eight of 15 targets were identified, and in the latter two cases, less than two of 15 targets were identified, on average. The latter level of performance was characteristic of traditional visual memory. Conditions were also tested in which both optic flow and image structure were presented but neither were available during an extended delay (25 s) between the original display and the response. This yielded a drop in performance level when compared with a condition in which image structure was visible throughout the delay interval. This drop (from ≈8.5 of 15 to ≈6.5 of 15) was not as large as the drop found between continuously available joint versus single forms of information. The reason is that the image structure was made available again at the end of the delay. The initial pairing of image structure with optic flow provided information about the location of hidden targets within the image. However, the subsequent interruption in the availability of the image structure yielded characteristic decay in memory performance (blank condition). The surprising result was an absence of such decay when the image structure remained continuously available during the delay (no blank condition). The current experiments are in line with these latter manipulations, and the scale of the effects was expected to be comparable. The interruption of image structure is a perturbation of embodied memory in which removal of image structure eliminates embodied memory.

Experiment 1: Baseline—No Orientation Change

In Experiment 1, we tested participants' recollection of the locations of targets after they were occluded. This replicated the fourth experiment of Pan et al. (2013) and established a baseline for further comparisons.

Method

Participants. Twenty-four normally sighted observers, of ages between 18 and 32 years, completed this experiment. Among them, 14 were students at Indiana University and eight were students from Sun Yat-sen University. The study was approved by the institutional review boards at both universities. Each participant received a small amount of monetary reward after finishing the experiment.

Apparatus. Participants sat in front of a computer monitor (display width = 43 cm; height = 27 cm) with a viewing distance of 50 cm. The refresh rate of the monitor was 60 Hz, with spatial resolution being 1680 × 1050.

Display. We used a simulated 3D display on a computer screen (see Figure 1). The display size was 43 cm (width) × 27 cm (height) and viewing distance was 50 cm. The display consisted of two rectangular surfaces, a smaller surface in front of a larger surface, both parallel to the computer screen (or perpendicular to the line of sight), with a depth separation between the surfaces. The surfaces were randomly textured in exactly the same way, that is, identical density of texture elements. The rear surface extended well beyond the edges of the computer screen and its edges never appeared on screen during experimental trials. The front surface was smaller (27° × 27° visual angle) so that it occluded only the central portion of the rear surface, leaving approximately 8° of the rear surface on the two sides and 2° on the top and bottom that could always be seen beyond edges of the front surface. The front surface contained cutout circles with black borders, through which part of the rear surface could be seen. The black borders constituted the image structure information in this task. Through some of these circular windows, pink circles could be seen lying on the rear surface. The pink circles were targets. Through some other circular windows, only the random textures on the rear surface could be seen. Windows showing only random textures were distracters. The windows, targets, and distracters were circles with diameters of 11 mm (roughly 1.3° visual angle). The number of targets varied as an experimental manipulation, and the number of distracters remained at 15. Although the two surfaces were separated in depth, this depth structure was only perceptible with continuous motion that generated differential optic flow along the depth edges. When the display was static, the depth separation could not be seen and the two surfaces appeared to be one that contained random textures, pink dots (targets) and black circles (window borders).

Figure 1. An illustration of the basic experimental display (not drawn to scale) used in all experiments. A front view (which is what participants actually saw) is shown on the right, and a side view (which is for clarifying the 3D structure) is shown on the left. (A) SFM phase: Two identically textured surfaces were separated in depth. The front surface contains cutout holes, through which pink disks (targets) or random textures (distracters) on the rear surface are seen. The two surfaces rigidly rotate and precess in depth (or make figure-8 movement around a vertical axis in the frontoparallel plane). This motion reveals the 3D structure. (B) Translation phase: The rear surface then translates and hence targets on it are progressively occluded by the front surface. (C) Eventually, targets are completely occluded by the front surface. Through the cutout holes in the front surfaces, only random textures on the rear surface are seen. SFM = structure from motion. See the online article for the color version of this figure.

Procedures. Participants read and signed consent forms and then completed three to 10 practice trials (with the experimenter present) to become familiar with the task.

Each experimental trial consisted of four phases: SFM, translation, delay, and response. An experimental trial started with the two planar surfaces rotating and precessing in depth, around an upright frontoparallel axis (which was not visible). This SFM generated differential optic flow for the two surfaces at different depth layers, and hence revealed the depth relation between the surfaces. Participants watched the surfaces rotate for 8 s and studied the locations of the targets. After the motion stopped, the two surfaces once again were frontoparallel. Immediately after that, the rear surface translated rigidly in one of the four directions—up, down, left or right—while the front surface remained stationary. This rigid translation of the rear surface could be seen both through the windows in the front surface and beyond the four edges of the front surface. As the rear surface translated, the pink circles on it passed beyond the cutout windows in the front surface and thus became occluded behind the front surface. The rigid translation resulted in progressive occlusion and yielded optic flow that specified the depth relation of the two surfaces as well as the locations of the pink targets behind the front surface. The rear surface translated for 11 mm, which was equivalent to the diameter of the targets and cutout windows, and this translation took 3 s. After the translation ended, targets were completely occluded by the front surface. The hidden locations of the targets would only be perceived in terms of the distance and direction that the entire rear surface had moved relative to the detectable image structures (i.e., the visible black borders of the cutout windows). After the translation was completed, either a long delay of 25 s or a short delay of 5 s was introduced, during which we manipulated the availability of image structure information (i.e., the visibility of both the random texture and window borders). Participants then used the mouse to click on the locations of the targets, which were now hidden behind the front surface (see Figure 2 for an illustration of the experimental procedures).

The structure of these displays was similar to that of the Kaplan kinematogram displays (Kaplan, 1969; see also Gibson, 1979/1986, pp. 189–191). The presence of two surfaces could only be seen with optic flow during the SFM and translation phases. Once motion stopped, only a single static display of random textures and black circles could be seen (see Figure 1 and Figure 2[1]).

When making responses, we encouraged participants to click accurately instead of randomly by introducing a point system: Starting with 200 points, if they identified a target correctly (that is, a "hit"), they gained a point; if they identified a target incorrectly (that is, clicking on a pixel on the display that was not a target, which would be counted as a "false alarm"), they lost a point; if they did not attempt to identify, there would be no point change. At the end of the experiment, participants received bonus payment (in addition to the standard participation payment) in proportion to their final points.[2] This was designed to prevent guessing and to promote accurate performance. The method was effective. In three experiments, there were very few false alarms. (Overall, the median of false alarms was 0. The Results sections under each experiment provide the respective false alarms rates.) This means that in more than half of the trials in each experiment, participants did not err. The extremely small number of false alarms in these experiments suggested that participants were careful and conservative when making responses—they did not guess.[3] Therefore, we simply analyzed the number of targets correctly identified (that is, hits) as a measure of performance.

In the current experiment, we manipulated the number of targets available (10, 15, or 20), the length of the delay phase duration (5 s or 25 s), and the persistence of image structure information during delay: On trials with persistent image structure, random textures and black circles were continuously visible during delay; on trials with interrupted image structure, a black screen was seen during delay. We referred these to the "no blank" and "blank" conditions, respectively, and we referred this factor as "blank" in the following analyses. We tested how these factors affected the identification of hidden targets. Table 1 and Table 2 provide optic flow and image structure information in each condition for all experiments. The number of distracters was fixed at 15 for all experiments. In Experiment 1, the combination of number of targets, delay duration, and blank, all randomized within a testing block, yielded 12 unique combinations of input variables. In one testing block, a participant completed two trials for each combination. Each participant completed two blocks, or 48 trials, in this experiment.

Results

In this study, we tested the identification of hidden targets when the numbers of available targets, the delay duration, and the persistence of image structure information (blank) varied. The number of targets that were correctly identified was significantly affected by the main effects of number of targets, $F(2, 46) = 75.89$, $p < .001$, $\eta^2 = 0.27$, and the persistence of image structure information (blank), $F(1, 23) = 36.47$, $p < .001$, $\eta^2 = 0.06$. Target identification was not affected by the length of delay, $F(1, 23) = 1.36$, $p = .26$, but it was affected by the interaction between delay and blank, $F(1, 23) = 8.93$, $p < .001$, $\eta^2 = 0.01$.

With up to 20 targets, the more targets there were, the more targets that participants identified (see Figure 3). The rate of increment was 0.28, $t(1150) = 13.71$, $p < .001$, that is, hits increased by 1 as the number of available targets increased by 3 or 4. Participants identified fewer targets in trials in which image structure was not persistent, that is, in which participants saw blank screens with no detectable image structure information during the delay. The mean number of hits for no-blank

---

[1] The experiments (Experiments 1, 2, and 3) can be downloaded and run from http://www.indiana.edu/~palab/Resources/Demos/Orientation_Change_Demo.zip. Instructions are included in the demo folder. Note that the demos are for Mac OS only. Motion speed may vary depending on the computer. Alternatively, a video of all three experiments can be obtained from http://www.indiana.edu/~palab/Resources/Demos/demo5.mp4.

[2] As bonus payment, participants at Indiana University received $0.01 per point gained, and participants at Sun Yat-sen University received ¥0.10 (≈ $0.015) per point gained.

[3] Participants did not guess, although sometimes they misremembered. For instance, Experiment 3 contained 1,728 trials, and in 41 trials, there were high false alarms and no hits. In these trials, it was likely that participants misremembered the direction in which the rear surface had moved, and hence where the targets were hidden relative to the visible window borders. Participants then systematically clicked on the wrong locations throughout. This kind of memory error was not equivalent to random guessing.

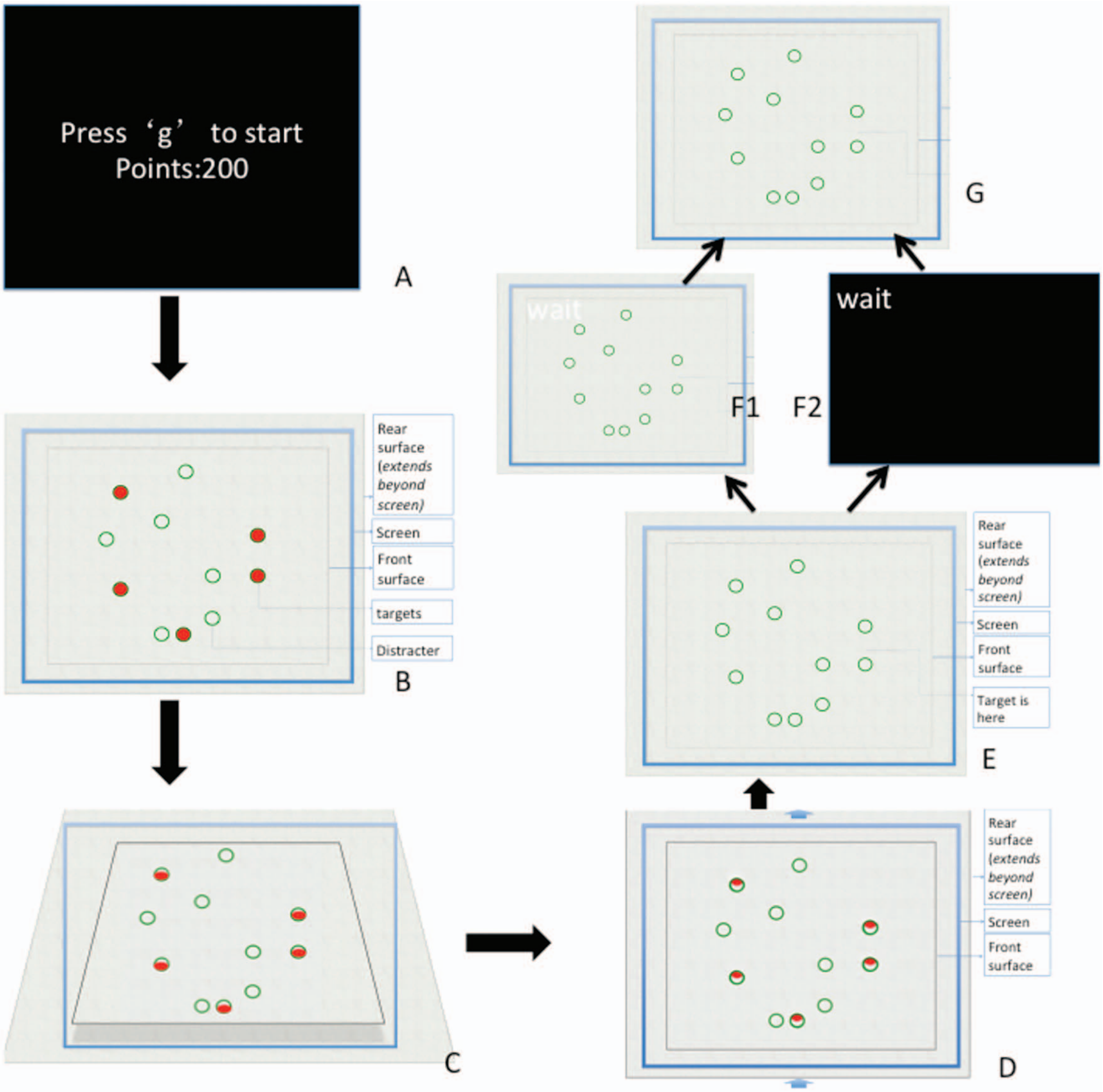

Figure 2. Illustrations of experimental stimuli and procedures (not drawn to scale) in Experiment 1. A trial began with a black screen showing the current points and an instruction to begin (A). The rear surface (containing targets and distracters) and the front surface (containing cutout windows) were separated in depth. Through the windows in the front surfaces, targets (red disks) and distracters (random textures) were seen (B). Then, the rear surface and the front surface rotated rigidly around an upright frontoparallel axis that also precessed (SFM) and revealed the 3D structure of the display (C). Then, the rear surface, along with the targets it contained, translated in one of four directions—up, down, left, or right (D) until targets moved out of the windows and were completely occluded by the front surface (E). Participants waited for 5 or 25 s while the monitor displayed either the ending scene (F1) or turned blank (F2). Finally, participants clicked on locations of the now hidden targets given the ending scene of the display (G). Go to http://www.indiana.edu/~palab/Resources/Demos/Orientation_Change_Demo.zip to download and see the experiments; or http://www.indiana.edu/~palab/Resources/Demos/demo5.mp4 to see a video of the experiments. See the online article for the color version of this figure.

trials was 8.90 (*SD* = 2.98), and that for blank trials was 7.95 (*SD* = 3.14).

In addition to the significant main effect of blank, the interaction between blank and delay was significant. As shown in Figure 4, long delays affected targets identification in trials with blank (or no persistent image structure information), but not in trials with no blank (or persistent image structure information). Specifically, in trials with blank, on average of 8.23 targets (*SD* = 3.05) were identified with short delays, and 7.68 targets (*SD* = 3.20) were identified with long delays. The difference was significant (p = .035, using the Benjamini and Hochberg (BH) correction for false discovery; Benjamini & Hochberg, 1995). In trials with no blank,

Table 1
Availability of Optic Flow Information in the Blank and No-Blank Conditions of the Three Experiments in This Study

		Availability of optic flow information during each phase			
Experiment	Condition	SFM	Translation	Roll	Delay
Experiment 1: 24 subjects	No blank	Available	Available	Not applicable	Not available
	Blank	Available	Available	Not applicable	Not available
Experiment 2: 12 subjects	No blank	Available	Available	Available	Not available
	Blank	Available	Available	Available	Not available
Experiment 3: 12 subjects	No blank	Available	Available	Available	Not available
	Blank	Available	Available	Not available	Not available

Note. SFM = structure from motion.

on average of 8.82 targets (*SD* = 2.98) were identified with short delays, and 8.98 targets (*SD* = 2.98) were identified with long delays. The difference was not significant ($p = .52$, using the BH post hoc correction for false discovery).

The false alarm rates (wrongly clicking on a pixel on the screen as if it was a target) were low in this experiment. In all no-blank trials, participants made 5,676 clicks, among which 551 were clicks on nontarget pixels of the display. In all blank trials, participants made 5,152 clicks, among which 570 were clicks on nontarget pixels of the display. Although number of wrong clicks counted for about 10% of total clicks, the false alarm rates in both cases had a median of 0. This means that in more than half of the trials, with or without blank, there was no false alarm; but in some trials, there were many false alarms. This was consistent with our conjecture that participants might misremember the direction of translation and therefore wrongly click many times in one trial. None of the factors significantly affected false alarm rates in this experiment.

Experiment 2: Continuous Change of Orientation

An observer's head often rolls around the line of sight in the act of looking. Phenomenologically, one's perception is not affected by such a head roll. In this experiment, we simulated a head roll by rolling the display in the frontoparallel plane (while keeping the observer's head unmoved) and examined hidden target perception when observers were allowed or not allowed to see the rolling motion. Specifically, in this experiment, participants viewed a dynamic display consisting of rigid rotation/precession in depth of two surfaces followed by rigid translation of the rear surface, and then rigid rolling of the two surfaces in the frontoparallel plane. Whereas the rigid rotation/precession in depth and rigid translation contained differential flow from the front and rear layers, frontoparallel roll did not provide any information for relative depths. Nonetheless, the frontoparallel rotation simulated a common visual experience, which was the change of orientation as a result of change in head orientation of an observer.

Method

Participants. Twelve participants who completed Experiment 1 also did this experiment. They were rewarded a small amount of money to compensate for their time and effort. The experiment was approved by the institutional review boards at Sun Yat-sen University and Indiana University.

Display and procedures. The apparatus, display, and procedures were the same as those in Experiment 1, except that a rolling phase was added immediately after the rigid translation phase and before the delay period began (see Figure 5). Participants first watched the SFM phase and learned the locations of targets. Then they saw the rigid translation of the rear surface taking the targets out of view in a specific direction beyond the visible window borders. The SFM and translation lasted for 8 s and 3 s, respectively (the same as in Experiment 1). Next, the two surfaces rolled rigidly together in the frontoparallel plane in the clockwise or counterclockwise direction for 30° around the line-of-sight axis. The window borders and hidden targets were also carried by the rolling of the surfaces. This frontoparallel transformation simulated rolling of the head to yield a change in orientation of the visual image. The frontoparallel rolling took 3 s, after which there

Table 2
Availability of Image Structure Information in the Blank and No-Blank Conditions of the Three Experiments in This Study

		Availability of image structure information during each phase			
Experiment	Condition	SFM	Translation	Roll	Delay
Experiment 1: 24 subjects	No blank	Available	Available	Not applicable	Available
	Blank	Available	Available	Not applicable	Not available
Experiment 2: 12 subjects	No blank	Available	Available	Available	Available
	Blank	Available	Available	Available	Not available
Experiment 3: 12 subjects	No blank	Available	Available	Available	Available
	Blank	Available	Available	Not available	Not available

Note. SFM = structure from motion.

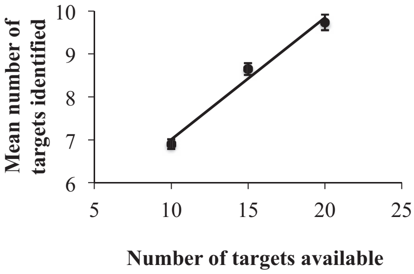

Figure 3. In Experiment 1, participants identified more targets when more targets were available to them. Error bars represent 1 *SE*.

was either a long (22 s) or a short delay (2 s) before the participant used the mouse to click on the hidden targets. The 22-s and 2-s delay durations were chosen to compensate for the 3 s of rolling duration and to keep the time elapsed between differential optic flow that specifies depth structure (translation phase) and participants' clicking (response phase) the same as that in Experiment 1.

In this experiment, we manipulated the number of targets available (10, 15, or 20 targets), the length of the delay phase duration (short vs. long), and the persistence of image structure information during delay (no blank vs. blank) and tested how these factors affected the identification of hidden targets. The combination of number of targets, delay duration, and blank yielded 12 unique combinations of independent variables. In each testing block, a participant completed two trials for each combination. Each participant completed four blocks or 96 trials in this experiment. (In two blocks, the frontoparallel roll was in the clockwise direction, and in two blocks the roll was in the counterclockwise direction. The roll direction did not affect performance, $p = .07$, because participants always saw the rolling and knew its direction. This factor was hence excluded from the subsequent data analysis.)

Data analysis. First, we looked at how target identification was affected by the number of targets available, delay duration, and blank. More importantly, to study embodied memory for targets with orientation change, we compared performance when there was no frontoparallel roll versus when there was roll (the process of which remained visible to participants). We combined and compared the 12 participants' data from the current experiment and their data from Experiment 1. The dependent measure was hits and all the independent variables were tested within subjects.

Results

We analyzed data from the current experiment with repeated measures ANOVA, and found that, similar to the results in Experiment 1, hidden target identification was significantly affected by the number of targets available, $F(2, 22) = 67.84$, $p < .001$, $\eta^2 = 0.35$, blank, $F(1, 11) = 5.40$, $p < .04$, $\eta^2 = 0.03$, and the blank–delay interaction, $F(1, 11) = 9.07$, $p < .02$, $\eta^2 = 0.01$. Additionally, in this experiment, hidden target identification was also affected by delay duration, $F(1, 11) = 5.22$, $p < .05$, $\eta^2 = 0.02$, and the interaction between blank and number of targets available, $F(2, 22) = 3.67$, $p < .05$, $\eta^2 = 0.01$.

In general, more hidden targets were identified in trials with no blank in which the image structure information was persistently available ($M_{\text{no blank}} = 8.35$, $SD_{\text{no blank}} = 3.32$; $M_{\text{blank}} = 7.78$, $SD_{\text{blank}} = 3.44$), and in trials with short delays ($M_{\text{short delay}} = 8.34$, $SD_{\text{short delay}} = 3.18$; $M_{\text{long delay}} = 7.79$, $SD_{\text{long delay}} = 3.57$). As the number of targets available increased, hits increased at the rate of 0.32, $t(1150) = 9.43$, $p < .001$, or roughly one additional hit for three additional available targets.

Next, we compared the 12 participants' performance in this experiment with their performance in Experiment 1. The difference between the two experiments was that in Experiment 1, there was no rolling of image structures, and in Experiment 2, there was rolling (which was visible to the participants). First, we found that there was no significant difference in performance between these two experiments, that is, without roll or with visible roll, $F(1, 11) = 1.35$, $p = .27$. Second, in both Experiment 1 and Experiment 2, the 12 participants always identified more targets when more were available (see Figure 6). The rates of increase in the two experiments were the same. The slope of a line fit by least square to hits versus number of targets available was 0.28 for Experiment 1, and 0.32 for Experiment 2. The slope, $t(1724) = -1.05$, $p > .29$, and intercept, $t(1724) = 1.54$, $p > .12$, differences between the two lines were not significant. Third, a significant blank–delay interaction was found in both experiments. Specifically, in Experiment 2, in trials with blank, 8.17 targets ($SD = 3.18$) were identified with short delays and 7.50 targets ($SD = 3.47$) were identified with long delays. The difference was significant ($p = .0056$, using the BH post hoc correction for false discovery; Benjamini & Hochberg, 1995). In trials with no blank, 8.54 targets ($SD = 3.06$) were identified with short delays and 8.47 targets ($SD = 3.32$) were identified with long delays. The difference was not significant ($p = .75$, using the BH post hoc correction for false discovery). Thus, in Experiment 2, extended delay resulted in decreased hits only in trials with blank, or when embodied memory was unavailable during delay (see Figure 7). This was similar to the result of Experiment 1.

In these three respects, performance in Experiment 2 showed similar trends as in Experiment 1. This suggested that when embodied memory had been formed by the interaction of optic flow and image structure in the SFM and translation phases, the calibrated image structures continued to preserve locations of previously seen and currently hidden targets, even when there was

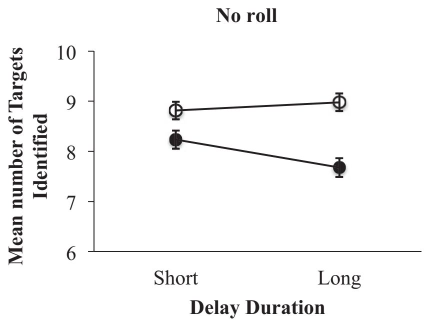

Figure 4. In Experiment 1, hits were affected by the interaction of blank and delay. When the image structure information was persistently available, hits did not drop with extended time delays (no blank: open circles). When the image structure information was unavailable during delays, hits dropped as the delays became longer (blank: filled circles). Error bars represent 1 *SE*.

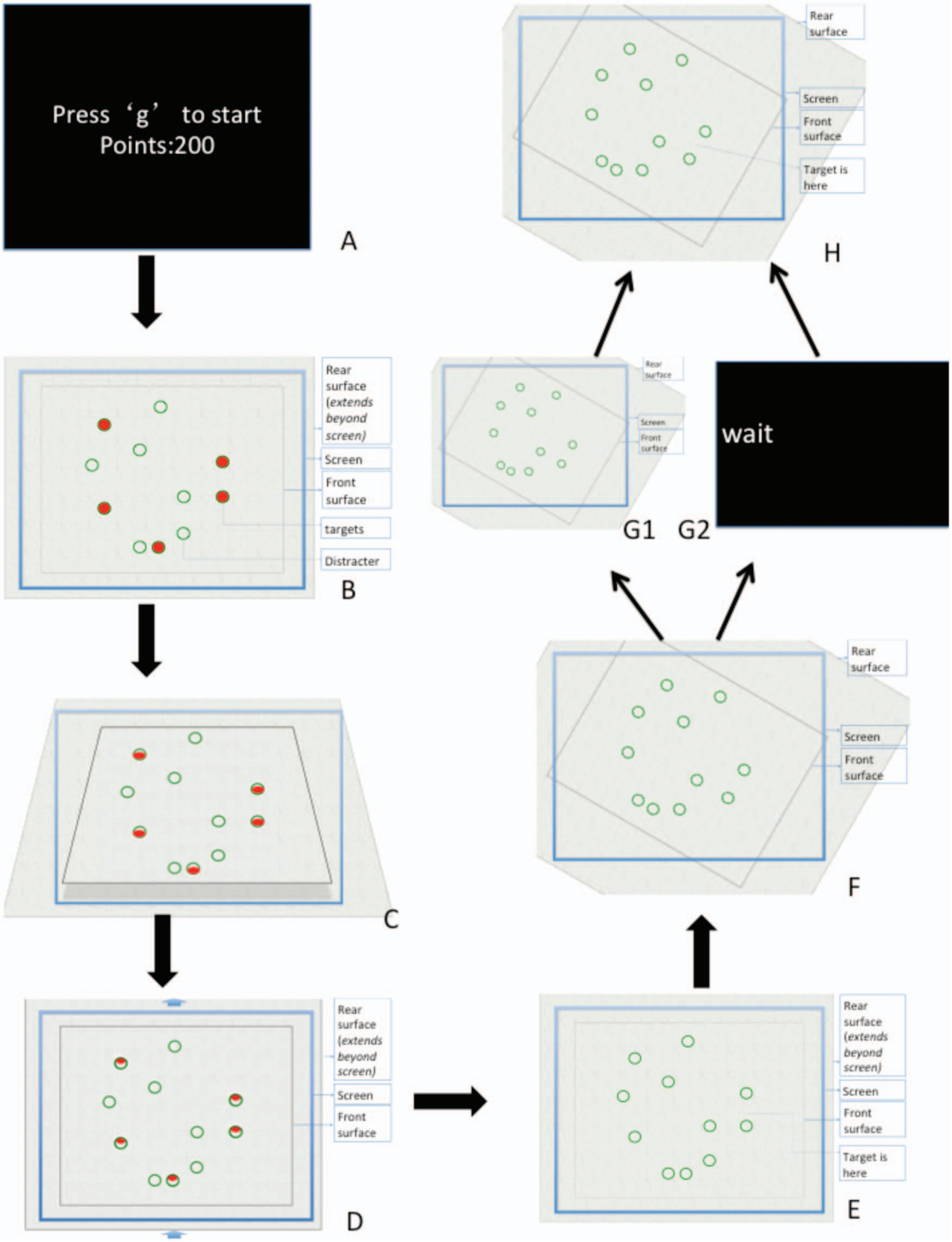

Figure 5. Illustrations of experimental stimuli and procedures (not drawn to scale) in Experiment 2. In Experiment 2, all displays until the progressive occlusion step were identical to those in Experiment 1 (Steps A to E in Figure 2). After targets on the rear surface were occluded (E), the two planes rolled for 30° clockwise or counterclockwise (F). Participants then waited for 5 or 25 s while seeing either the postroll ending scene (G1) or a blank screen (G2). Finally, participants clicked on locations of the now hidden targets (H). See the online article for the color version of this figure.

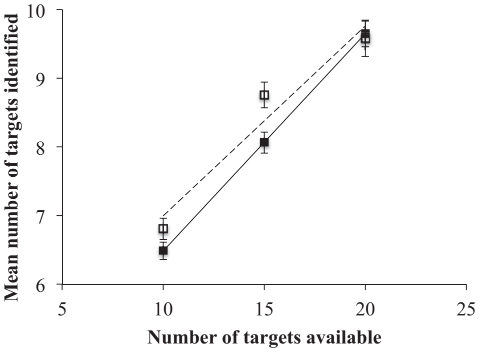

Figure 6. Comparison of performance between trials with no rolling (open squares) and trials with seen rolling (filled squares). Data were from 12 participants who identified hidden targets with no roll (Experiment 1) and trials with continuous visible roll (Experiment 2). In both cases, hits increased as the number of targets increased at similar rates. There was no difference between the slopes of the two trend lines. Error bars represent ±1 *SE*.

an orientation change. As long as the calibrated image structures remained in view, additional orientation change did not add extra difficulty to the task or lower target identification performance. Thus, embodied memory is powerful enough to overcome orientation change.

The false alarm rates were low in this experiment. In all no-blank trials, participants made 5,505 clicks, among which 694 were clicks on nontarget pixels on the display. The median of false alarm rates is 1, with a range of 0 to 17, and an interquartile range (IQR) of 2. In all blank trials, participants made 5,211 clicks, among which 729 were clicks on nontarget pixels on the display. The median of false alarm rates is 1, with a range of 0 to 13, and an IQR of 2. In both cases, although false alarms counted for about 13% of total clicks, their distribution was highly skewed with long right-hand tails. In general, the false alarms were low, but in some trials, participants clicked on many nontarget pixels and no targets. A possible reason was that they incorrectly remembered the translation direction of the rear surface, which would result in many false alarms on one trial. None of the factors significantly affected false alarm rates in this experiment.

Experiment 3: Discrete Change in Orientation

In this experiment, we studied target identification when the frontoparallel rolling phase was not visible. In this case, there was no embodied memory for orientation change, because image structure information was not available during the rolling, so optic flow and image structure were not allowed to interact in this respect. Comparing results from this experiment with those from Experiment 2, we stressed the importance of calibrated image structure information and, hence, embodied memory in overcoming visual image perturbations caused by orientation change.

Method

Participants. The other 12 participants who completed Experiment 1 (but not Experiment 2) completed this experiment. The experiment was approved by the institutional review boards at Sun Yat-sen University and Indiana University.

Display and procedures. Display and viewing conditions were the same as in the previous experiments. Similar to those in Experiment 2, this experiment contained the SFM (which revealed the 3D structures), rigid translation (which yielded progressive occlusion), frontoparallel rolling (which yielded orientation change), delay, and response phases. However, in this experiment, the rolling and delay were treated as one phase and the factor of blank affected both. That is, in half of the trials, the process of frontoparallel rolling was visible, and also in these trials, the black window borders were visible during delay. These were the no-blank trials, in which image structure information was persistently available. In the other half of the trials, rolling occurred when the screen went blank and the screen remained blank during the delay. These were the blank trials, in which image structure information was not persistently available (see Figure 8).

We recorded target identification performance as a function of number of available targets (10, 15, or 20), persistence of image structure information (no blank vs. blank), and delay duration (2 s or 22 s). These manipulations were repeatedly tested in six blocks.

Additionally, for a given participant, in two blocks, the frontoparallel rolling was always to one direction, either clockwise or counterclockwise, for 30°. In these two blocks, the experimenter always told the participants the direction in which the display was going to roll. Prior to the experimental blocks, the verbal instruction to the participants was, “In the next block, you will always see the display rolling by 30 degrees clockwise/counterclockwise or to the right/left, after the targets become occluded. Then you should try to locate the hidden targets and click on them.” The participants also received three to five practice trials to get familiar with the rolling direction and degree. In these two known-rolling-direction blocks, a participant did two repetitions of each combination of number of target (three levels), blank (two levels), and delay (two levels), or 24 trials per block for two blocks.

In the other four experimental blocks, the roll direction was randomized within block (in half of the trials, the display rolled 30° clockwise, and in the other half, the display rolled 30° counterclockwise). In this case, the participants did not know to which

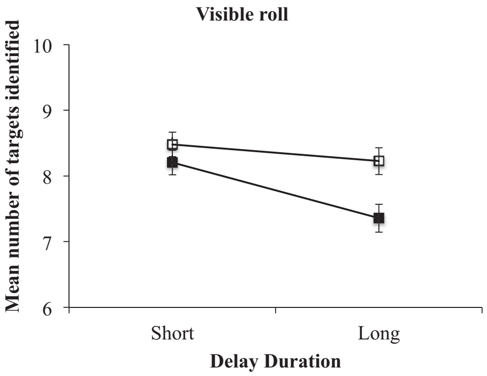

Figure 7. The blank–delay interaction was significant in Experiment 2. Pairwise *t* test (with BH correction) showed that only hits in the long delay and blank condition were significantly lower than those in the other three ($p < .006$ in all cases). No blank: open squares. Blank: filled squares. Error bars represent 1 *SE*.

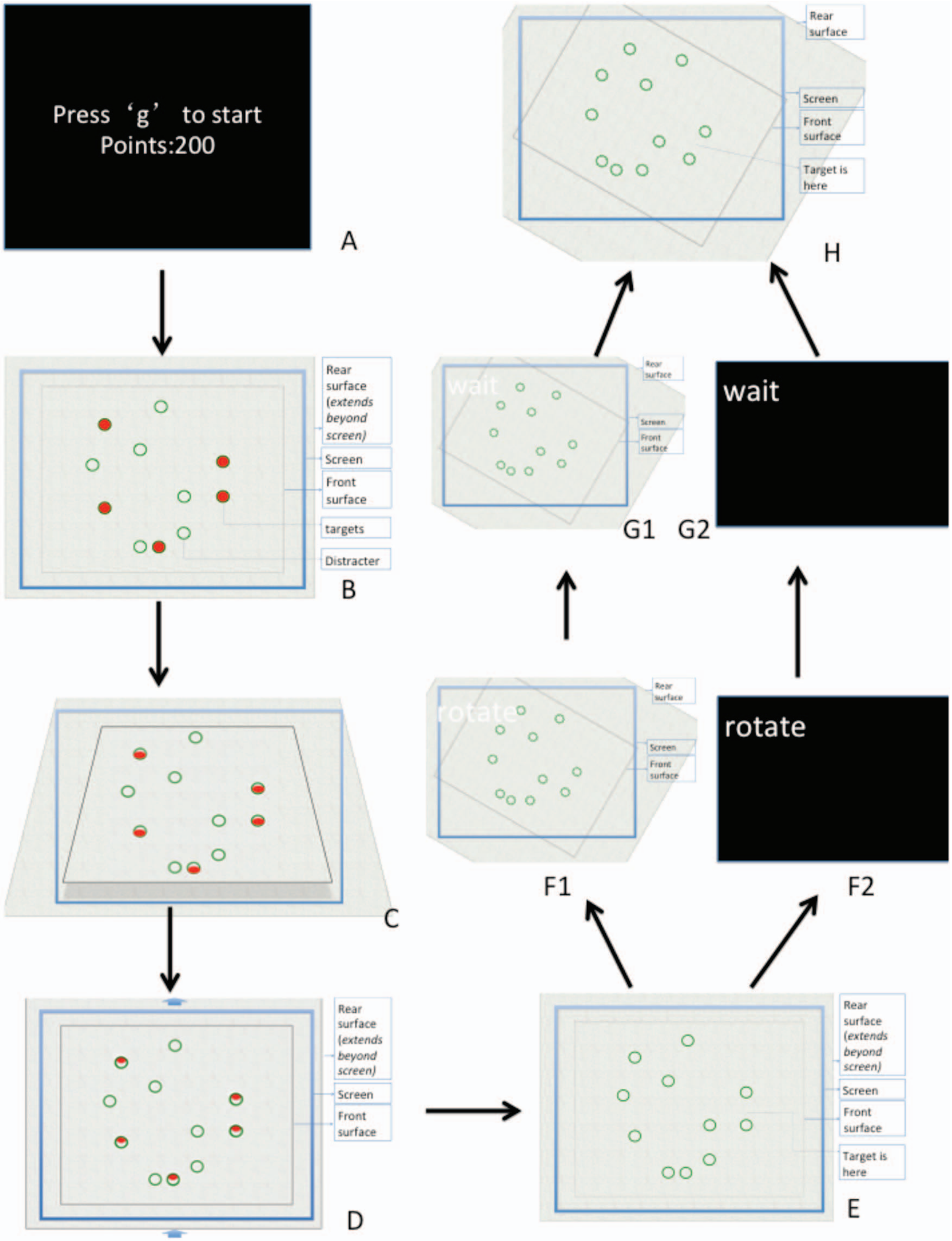

Figure 8. Illustrations of experimental stimuli and procedures (not drawn to scale) in Experiment 3. In Experiment 3, all displays until the progressive occlusion step were identical to those in Experiment 1 (Steps A to E in Figure 2). After targets on the rear surface were occluded (E), participants either saw the two planes rolling in the frontoparallel plane (F1) and continued seeing the ending scene while waiting for 5 or 25 s (G1); or the screen turned blank when rolling and delay took place (F2 and G2). After the delay, participants clicked on locations of the now hidden targets (H). See the online article for the color version of this figure.

direction the roll would be. Prior to the experimental blocks, the verbal instruction to the participants was,

> In the next block, you will see the display rolling by 30 degrees sometimes clockwise or to the right and sometimes counterclockwise or to the left, after the targets become occluded. Then you should try to locate the hidden targets and click on them.

The participants received three to five practice trials to get familiar with the roll. Altogether, in these four unknown roll direction blocks, there were 24 unique combinations of variables formed by rolling direction (two levels), number of target (three levels), blank (two levels), and delay (two levels). A participant did one repetition, or 24 trials per block for four blocks. All 12 participants completed six experimental blocks, the order of which was randomized.

Data analysis. First, we studied how target identification was affected by the number of targets available, delay duration, rolling direction, and blank. We paid special attention to the factor of blank. We analyzed target identification performance when observers saw the frontoparallel rolling motion and continued to see the outlines of windows during delay in the no-blank condition. This was the situation in which embodied memory for orientation change was available. (Note that this condition was identical to the no-blank condition in Experiment 2.) We also analyzed target identification performance when observers did not see the frontoparallel rolling motion and did not see any outlines of windows during delay in the blank condition. This was the situation in which embodied memory for orientation change was not available and the completion of the task relied on mental rotation and memory-in-the-head. We contrasted performances in these two cases.

Second, we examined the effect of knowledge by comparing performance between blocks, that is, whether being told about the direction of frontoparallel rolling helped target identification. We compared target identification performance between using knowledge provided by others and using embodied memory based on real-time information, and studied their effects on overcoming orientation change.

Results

An omnibus repeated measures ANOVA showed that, in this experiment, hits were significantly affected by the number of targets, $F(2, 22) = 72.84$, $p < .001$, $\eta^2 = 0.27$, blank, $F(1, 11) = 24.63$, $p < .001$, $\eta^2 = 0.07$, and whether rotation was known or unknown, $F(1, 11) = 5.39$, $p = .040$, $\eta^2 = 0.01$.

In this experiment, hits increased with the number of targets available at the rate of 0.27, $t(1726) = 14.8$, $p < .001$. More targets were identified in trials with no blank ($M_{\text{no blank}} = 7.96$, $SD_{\text{no blank}} = 2.98$; $M_{\text{blank}} = 6.81$, $SD_{\text{blank}} = 3.50$) and in trials with known direction of rolling ($M_{\text{known}} = 7.68$, $SD_{\text{known}} = 3.29$; $M_{\text{unknown}} = 7.24$, $SD_{\text{unknown}} = 3.29$). In the no-blank trials, in which image structure information was always available, whether there was a short or a long delay did not affect hits, $F(1, 45) = 0.03$, $p = .96$. Furthermore, in the no-blank trials, knowing the rolling direction or not did not make a difference, $F(1, 11) = 0.44$, $p = .52$. In the blank trials, however, when no image structure information was available during rolling and during delay, participants performed better when they were told the rolling direction than when they were not, $F(1, 11) = 6.39$, $p < .03$ (see Figure 9). Generally, these results suggested that performance was better with embodied memory for orientation change in the no-blank trials and performances was also better when participants had knowledge about the direction of orientation change. Knowing the direction was only helpful when image structure information was absent during orientation change and delay, because, in this case, embodied memory for orientation change was unavailable and participants had to rely on mental rotation to resolve the orientation change. Hence, being told about the direction of roll helped the mental rotation and improved performance.

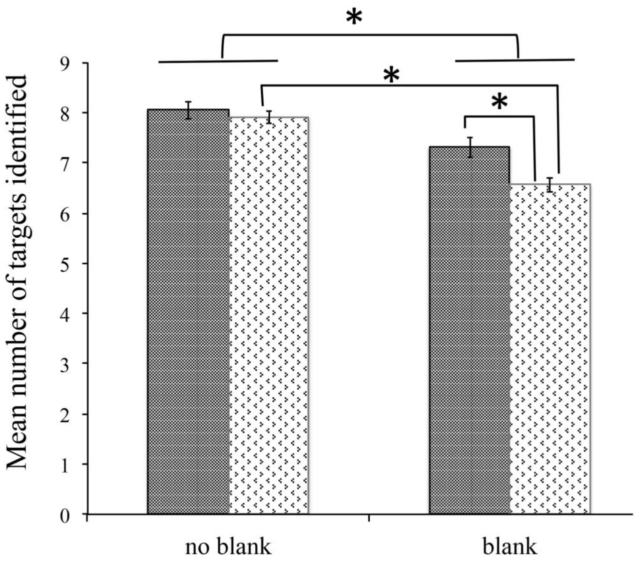

Figure 9. In Experiment 3, target identification was affected by the availability of image structure information (no blank vs. blank). Furthermore, being told about the direction of roll or not only affected performance in the blank trials. Dark bars: known roll direction. Light bars: unknown roll direction. Error bars represent 1 *SE*. * $p < .05$.

The false alarm rates were low in this experiment. In all no-blank trials, participants made 7,678 clicks, among which 807 were clicks on nontarget pixels of the display. In all blank trials, participants made 6,689 clicks, among which 801 were clicks on nontarget pixels of the display. In both cases, the median of false alarm rates is 0, and the IQR was 1. Although misclicks counted for about 11% of total clicks, the low median and low IQR suggested that the distribution of false alarms was skewed and in most trials false alarms were low.

General Discussion

In three experiments, we investigated whether embodied memory functions to facilitate perception with image orientation change. In Experiment 1, we replicated results from Pan et al. (2013) and showed that when optic flow and image structure coexisted, optic flow calibrated image structure information, which formed an embodied memory that preserved a large number of hidden target locations over long time delays. Specifically, with embodied memory, the number of targets identified was above 8, with both short and long delays. This exceeded the capacity, in terms of number of items and time duration, of visual short-term memory (~4 ± 1 items; Luck & Vogel, 1997; Vogel, Woodman,

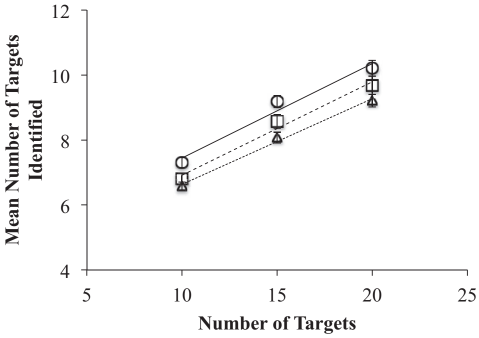

Figure 10. In the no-blank conditions of Experiments 1, 2, and 3, hits always increased with number of targets available at the same rate, so an orientation change did not lower performance, as long as embodied memory was available. Circle: performance in the no-blank condition in Experiment 1; square: performance in the no-blank condition in Experiment 2; triangle: performance in the no-blank condition in Experiment 3. Error bars represent 1 *SE*.

& Luck, 2001). When embodied memory was unavailable, as a result of temporary perturbation of image structure information, performance exhibited a classical decay with time, which is a characteristic of memory-in-the-head (Brown, 1958; Peterson & Peterson, 1959).

In Experiments 2 and 3, an orientation change was simulated as the display rigidly rolled in the frontoparallel plane. This roll did not change the spatial relations between the visible image structures (the window borders) and the hidden target locations, although it did change the locations of both. In Experiment 2, participants always had access to the image structure information during rolling. In half of the trials, participants saw the window borders during delay (no-blank condition), and in the other trials, they did not (blank condition). Experiment 2, therefore, is analogous to Experiment 1, except that there was an additional phase of rolling. In Experiment 3, in half of the trials, participants saw the window borders on screen during the processes of rolling and they continued seeing the borders during delay (no-blank condition). In the other trials, participants did not see the window border during the processes of rolling or of delay (blank condition). Therefore, the no-blank conditions of Experiments 2 and 3 were identical. However, the blank condition in Experiment 2 allowed participants to see the window borders during rolling, and the blank condition in Experiment 3 did not allow participants to see the window borders during rolling, in which case participants had to rely on mental rotation to recover the direction of roll.

When image structure information was calibrated by optic flow and remained persistently available, having a roll did not change target identification performance. There was no difference in hits in the no-blank conditions across the three experiments, $F(2, 45) = 1.73$, $p > .15$ (see Figure 10). As shown in Figure 10, with persistent (and calibrated) image structure information during rolling and delay, hits always increased with the number of targets available at the same rate for all three experiments. Additionally, in all no-blank trials, delay duration did not affect performance, $F(1, 45) = 3.10$, $p > .9$. This result suggested that embodied memory was able to help participants overcome the perturbation of orientation change.

With continuous image structure information available in the no-blank conditions, orientation change did not affect target identification. Next, we studied how orientation change affected performance (or not) when image structure was interrupted in the blank conditions. First, we compared performance in the blank trials of Experiment 1 (with no roll and no image structure information during delay) versus Experiment 2 (with perceptible roll and no image information during delay). There was no significant difference between these, $F(1, 11) = 0.27$, $p = .61$ (see Figure 11, left). Thus, orientation change did not lower target identification performance so long as the rolling was visible, regardless of whether image structure was available or not during delay. Moreover, a comparison between the blank trials in Experiment 1 and the blank trials in Experiment 3 showed that (a) when the rolling

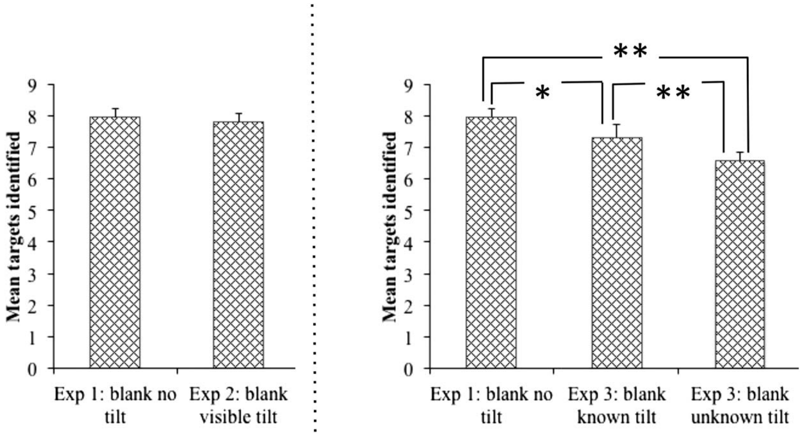

Figure 11. Performance with interrupted image structure information in the three experiments. Left: In the blank trials of Experiments 1 and 2, hits were not affected by orientation change, because the rolling was visible to the participants. Right: Hits in Experiment 1 (no roll) were higher than hits in Experiment 3 (unseen roll), regardless of whether participants were informed of the roll direction. Error bar = 95% CI. * $p < .05$. ** $p < .01$.

was not perceptible, target identification performance was poorer, although (b) being told about the rolling direction led to higher hits than not being told about the rolling direction, $F(2, 22) = 10.2$, $p < .001$. Pairwise *t* tests (with BH correction) showed that hits in the blank trials of Experiment 3 were significantly lower than hits in the blank trials of Experiment 1, both with known and unknown roll directions ($p = .013$ and $p < .001$, respectively); and hits in the blank trials of Experiment 3 with unknown roll direction were lower than hits in the blank trials of Experiment 3 with known roll direction ($p < .005$; see Figure 11, right panel).

As described in the introduction, results of this work echo and extend previous findings in the literature. In particular, the current results suggested that knowing the direction and amount of orientation change (without visually observing it) does not completely compensate for the perturbation or make identification performance as good as that with no orientation change of the visual scene. The drop in target identification was likely because of a challenging cognitive task—mental rotation. This was supported by results from Simons and Wang (Simons & Wang, 1998; Wang & Simons, 1999), who showed that change detection was worse when there was an unseen orientation change. Specifically, in their studies, participants first saw five objects on a tabletop. Then, the tabletop was occluded by a curtain and rotated for 50°. During this time, one object was moved to a different location on the tabletop. Rigidly attached to the tabletop was a rod that extended from the curtain and remained visible while the tabletop was occluded and rotated. The participants did not see the target objects going through rigid rotation, but they knew how much the objects had moved by watching the rod rotate. This setup was analogues to our experimental condition, in which participants did not see the rolling but were told the roll direction and amount. In either case, optic flow was not allowed to interact with image structure information, and therefore orientation change had to be solved by mental rotation, which caused identification and change detection performances to drop, compared with when there was no orientation change. Thus, exposure to combined optic flow and image structure information and their interaction is not replaceable by the mere knowledge of the orientation change. It is the combination of optical information that underlies accurate and stable perception during orientation change.

Finally, we note that a form of rehearsal may be part of the embodied memory dynamic or process relevant to an account of the results in the blank condition compared with the no-blank condition. In this account, the continuous availability of the image structure facilitates or enables rehearsal of the spatial locations of the targets. Using the image structure to rehearse the locations means that the image structure is providing stable information about those target locations in the context of the preceding optic flow. In rehearsing, the perceiver interacts with the image structure attending to actual target locations as previously specified by the optic flow. Rehearsing in this context could also yield a form of perceptual learning, for instance, learning the pattern of the target locations. The blank condition amounts to a perturbation of this interaction. The question is how stable is the perceptual performance in response to such perturbations. Our results show that a perturbation in the availability of image structure destabilizes embodied memory. The performance level declines. The relation between the effective rate of decline and the duration of the perturbation (independent of the delay period) remains to be determined.

Conclusion

In three experiments, we showed that participants identified a large number of progressively occluded targets with good temporal stability when both optic flow and image structure information were available. When an orientation change occurred and the rolling of the object array was visible, participants took advantage of the combined optic flow and image structure information and identified hidden objects just as effectively as when there was no orientation change. Orientation change only produced decrements in identification performance when the rolling was not visible. In this case, optic flow information specifying the orientation change was not available. The absence of optic flow could not be compensated by the knowledge of rolling amount and direction, because even if participants were given this knowledge, their performance still did not match that in trials with no orientation change.

Therefore, accurate and stable perception of targets going through progressive occlusion and orientation change is enabled by a combination of optic flow and image structure information. When the two interact, optic flow provides powerful information about the spatial layout, including the location of objects that become hidden relative to the visible image structures before and after the visual image array rolls. The perceived spatial layout is preserved in image structure that has been calibrated by the preceding optic flow. Offloading the spatial relations specified by the transient optic flow to external stable image structure allows individuals to access and act upon information provided by optic flow without having to hold it all in the head.

References

Anderson, J., & Bingham, G. P. (2010). A solution to the online guidance problem for targeted reaches: Proportional rate control using relative disparity tau. *Experimental Brain Research, 205,* 291–306. http://dx.doi.org/10.1007/s00221-010-2361-9

Anderson, J., & Bingham, G. P. (2011). Locomoting-to-reach: Information variables and control strategies for nested actions. *Experimental Brain Research, 214,* 631–644. http://dx.doi.org/10.1007/s00221-011-2865-y

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B: Methodological, 57,* 289–300.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *The Quarterly Journal of Experimental Psychology, 10,* 12–21. http://dx.doi.org/10.1080/17470215808416249

Domoni, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: A new theory. *Trends in Cognitive Science, 7,* 444–449.

Fajen, B. R. (2005a). Calibration, information, and control strategies for braking to avoid a collision. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 480–501. http://dx.doi.org/10.1037/0096-1523.31.3.480

Fajen, B. R. (2005b). The scaling of information to action in visually guided braking. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1107–1123. http://dx.doi.org/10.1037/0096-1523.31.5.1107

Fajen, B. R. (2008). Learning novel mappings from optic flow to the control of action. *Journal of Vision, 8*(11), 12.1–12.12. http://dx.doi.org/10.1167/8.11.12

Gibson, J. J. (1986). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin. (Original work published 1979)

Hergenhahn, B. R. (2008). *An introduction to the history of psychology* (6th ed.). Belmont, CA: Thomson Wadsworth.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13,* 289–303. http://dx.doi.org/10.3758/BF03202498

Jolicoeur, P. (1990). Identification of disoriented objects: A dual-systems theory. *Mind & Language, 5,* 387–410. http://dx.doi.org/10.1111/j.1468-0017.1990.tb00170.x

Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics, 6,* 193–198. http://dx.doi.org/10.3758/BF03207015

Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics, 50,* 363–370. http://dx.doi.org/10.1007/BF00336961

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America, A, Optics, Image & Science, 8,* 377–385. http://dx.doi.org/10.1364/JOSAA.8.000377

Lee, Y. L., & Bingham, G. P. (2010). Large perspective changes yield perception of metric shape that allows accurate feedforward reaches-to-grasp and it persists after the optic flow has stopped! *Experimental Brain Research, 204,* 559–573. http://dx.doi.org/10.1007/s00221-010-2323-2

Li, L., & Warren, W. H., Jr. (2000). Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vision Research, 40,* 3873–3894. http://dx.doi.org/10.1016/S0042-6989(00)00196-6

Li, L., & Warren, W. H., Jr. (2002). Retinal flow is sufficient for steering during observer rotation. *Psychological Science, 13,* 485–490. http://dx.doi.org/10.1111/1467-9280.00486

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281. http://dx.doi.org/10.1038/36846

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of The Royal Society of London, Series B: Biological Science, 207,* 187–217.

Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception, 3,* 63–80. http://dx.doi.org/10.1068/p030063

Pan, J. S., & Bingham, G. P. (2013). With an eye to low vision: Optic flow enables perception despite image blur. *Optometry and Vision Science, 90,* 1119–1127. http://dx.doi.org/10.1097/OPX.0000000000000027

Pan, J. S., Bingham, N., & Bingham, G. P. (2013). Embodied memory: Effective and stable perception by combining optic flow and image structure. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1638–1651. http://dx.doi.org/10.1037/a0032070

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198. http://dx.doi.org/10.1037/h0049234

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience, 3*(Suppl.), 1199–1204. http://dx.doi.org/10.1038/81479

Rock, I., & Heimer, W. (1957). The effect of retinal and phenomenal orientation on the perception of form. *The American Journal of Psychology, 70,* 493–511. http://dx.doi.org/10.2307/1419441

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9,* 315–320. http://dx.doi.org/10.1111/1467-9280.00062

Tarr, M. J., & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition, 67,* 1–20. http://dx.doi.org/10.1016/S0010-0277(98)00026-2

Tarr, M. J., & Vuong, Q. C. (2002). Visual object recognition. In H. Pashler & S. Yantis (Eds.), *Stevens' handbook of experimental psychology: Volume 1, Sensation and perception* (pp. 283–314). New York, NY: Wiley. http://dx.doi.org/10.1002/0471214426.pas0107

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics, 54,* 157–169. http://dx.doi.org/10.3758/BF03211751

Todd, J. T. (1995). The visual perception of three-dimensional structure from motion. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 201–226). New York, NY: Academic Press. http://dx.doi.org/10.1016/B978-012240530-3/50008-0

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48,* 419–430. http://dx.doi.org/10.3758/BF03211585

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92–114. http://dx.doi.org/10.1037/0096-1523.27.1.92

Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition, 70,* 191–210. http://dx.doi.org/10.1016/S0010-0277(99)00012-8

Wann, J., & Land, M. (2000). Steering with or without the flow: Is the retrieval of heading necessary? *Trends in Cognitive Sciences, 4,* 319–324. http://dx.doi.org/10.1016/S1364-6613(00)01513-8

Warren, W. H., Jr., & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America, A, Optics, Image & Science, 7,* 160–169. http://dx.doi.org/10.1364/JOSAA.7.000160

Wilkie, R. M., & Wann, J. P. (2006). Judgments of path, not heading, guide locomotion. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 88–96. http://dx.doi.org/10.1037/0096-1523.32.1.88

Yilmaz, E. H., & Warren, W. H., Jr. (1995). Visual control of braking: A test of the tau hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 996–1014. http://dx.doi.org/10.1037/0096-1523.21.5.996

Received April 9, 2016
Revision received December 29, 2016
Accepted January 2, 2017 ■