Behavioral/Systems/Cognitive

Motion Parallax Is Computed in the Updating of Human Spatial Memory

W. Pieter Medendorp,[1,2,5] **Douglas B. Tweed,**[1,2,4] **and J. Douglas Crawford**[1,2,3]

[1]Canadian Institutes of Health Research Group for Action and Perception, [2]York Centre for Vision Research, and [3]Department of Psychology, York University, Toronto, Ontario, Canada, M3J 1P3, [4]Departments of Physiology and Medicine, University of Toronto, Toronto, Ontario, Canada, M5S 1A8, and [5]Nijmegen Institute for Cognition and Information and FC Donders Centre for Cognitive Neuroimaging, University of Nijmegen, NL 6500 HE, Nijmegen, The Netherlands

As we move through space, stationary objects around us show motion parallax: their directions relative to us change at different rates, depending on their distance. Does the brain incorporate parallax when it updates its stored representations of space? We had subjects fixate a distant target and then we flashed lights, at different distances, onto the retinal periphery. Subjects translated sideways while keeping their gaze on the distant target, and then they looked to the remembered location of the flash. Their responses corrected almost perfectly for parallax: they turned their eyes farther for nearer targets, in the predicted nonlinear patterns. Computer simulations suggest a neural mechanism in which feedback about self-motion updates remembered locations of objects within an internal map of three-dimensional visual space.

Key words: motion parallax; remapping; spatial perception; translations; three-dimensional; human; spatial memory; spatial updating

Introduction

The representation by the brain of visual space is controversial (Von Helmholtz, 1876; Howard, 1982; Andersen et al., 1985; Duhamel et al., 1992). Objects not currently in view are stored in memory, and increasing evidence indicates that their locations are stored in a coordinate system fixed with respect to the eyeball, a retinal frame (Heide et al., 1995; Henriques et al., 1998; Colby and Goldberg, 1999) even for near targets (Medendorp and Crawford, 2002). However, this means that every time we move our eyes, we must update our memories of these objects to keep them in register with their true spatial locations. Such updating has now been demonstrated at the level of single cells in monkey extrastriate visual areas (Nakamura and Colby, 2002), posterior parietal cortex (Duhamel et al., 1992; Batista et al., 1999; Colby and Goldberg, 1999), frontal cortex (Goldberg and Bruce, 1990), and superior colliculus (Walker et al., 1995), as well as in the human posterior parietal cortex (Medendorp et al., 2003).

In some cases, we know from behavioral studies that the updated locations are accurate and available for motor planning: people can direct their eyes to remembered targets, compensating for any intervening rotary motion of the eyes (Hallet and Lightstone, 1976; Schlag et al., 1990; Blouin et al., 1998; Herter and Guitton, 1998; Smith and Crawford, 2001; Medendorp et al., 2002a). However, rotary motion is an atypically simple, special case (Smith and Crawford, 2001; Medendorp et al., 2002a). In real life, our eyes also translate through space, for instance, when we walk or even when the head rotates about the vertebral column (Medendorp et al., 1998). When the eye translates, motion parallax makes the images of near objects move faster across the retina than the images of far objects (Howard and Rogers, 1985; Regan, 1991), as shown in Figure 1*A*. Surprisingly, it is unknown whether spatial updating works for translational motion. Translational updating in a retinal frame would be computationally complex, because each object would have to be handled differently depending on its distance. In other words, this requires the brain to possess a dynamic three-dimensional (3-D) representation of visual space.

Here, we assessed translational updating in human subjects by having them look to remembered objects at different distances after self-generated eye translations, brought about by either head translations or head rotations. We present a model that allows us to interpret these results in terms of their computational and physiological significance for the brain.

Materials and Methods

General. All experiments were approved by the York University Human Participants Review Subcommittee. Seven subjects, 21–46 years of age, gave informed consent to participate in the experiments. All participants had normal or corrected-to-normal visual acuity, and all were free of any known sensory, perceptual, or motor disorders.

Measurement of head and eye motion. Our techniques for measuring and analyzing eye and head kinematics are described by Medendorp et al. (2000, 2002b). Here, we summarize them and specify the response parameters we used to analyze the data.

Location and orientation of the head as well as the locations of the ears and eyes in space and the locations of the stimulus light-emitting diodes (LEDs) were recorded using an OPTOTRAK 3020 system (Northern

Digital, Waterloo, Ontario, Canada) This system provided online information about the 3-D position of infrared-emitting diodes (ireds) with an accuracy of <0.2 mm.

During the experiment, the subject wore a lightweight helmet (<0.2 kg), to which six ireds were attached. Before the experiments, we measured the locations of the eyes and ears with respect to the ireds on the helmet: we placed ireds on the auditory meatus and on each closed eyelid, and recorded these together with the ireds on the helmet. With this information, we were able, during the subsequent experiments, to compute the positions of the ears and eyes in space on the basis of the helmet ireds alone. The center of rotation of each eye was assumed to be 1.3 cm behind its cornea.

Data were sampled at 100 Hz and stored on hard disk for off-line analysis. The coordinates of the ireds were transformed to a right-handed space-fixed coordinate system. The x–y plane was aligned with each subject's horizontal plane. The positive x-axis pointed forward, the positive y-axis was directed to the left (i.e., along the shoulder line), as seen from the subject, and the z-axis pointed upward. The origin of the coordinate system coincided with the center of the interaural axis when the subject was looking straight ahead. The orientation and location of the head were also determined with respect to this reference posture of the head. Measurements of head orientation and location were accurate to within 0.2° and 0.2 mm.

Binocular horizontal and vertical eye-in-space orientations (i.e., gaze) were measured using search coils (Skalar, Delft, The Netherlands) in three mutually orthogonal magnetic fields generated by field coils 2 m across. The field coils were calibrated before the experiment. After demodulation, the three voltages from each coil were sampled at 100 Hz and converted off-line into angular measures (Henriques et al., 1998). During the experiment, the eye-coils were calibrated by having subjects face straight ahead and fixate the stimulus LEDs four times each. By combining the locations of the stimulus and the reconstructed locations of both eyes (using the helmet calibration data), we were able to compute the direction of the stimulus LEDs with respect to each subject's eyes. In this way, both eye-coil signals could be matched to the corresponding vertical and horizontal LED locations. To describe the eye movements, we used a binocular coordinate system that distinguishes eye movements in direction (version) from eye movements in depth (vergence) (see below). Calibration errors were 0.32° (SD = 0.10°) for direction and 0.37° (SD = 0.12°) for depth, respectively, averaged across subjects. For the latter, this resulted in an uncertainty in the depth estimates of <5 cm (SD = 1 cm) within the target range tested (<1 m). For both eye and head, leftward and downward rotations were taken as positive.

Three personal computers controlled the experiment. One acquired the search-coil measurements, one controlled the visual stimulus, and the third collected the ired data.

Stimuli. All of our experiments took place in complete darkness, with the target LEDs as the only visual cues. We used this impoverished environment to test the ability of the brain to update its visual representations on the basis of self-motion clues alone. Nine space-fixed targets (green LEDs) were placed in front of the subject in the transverse plane. The targets were at different distances, 20–100 cm from the subject. Because of the large variety of initial eye locations and/or subsequent eye translations during the test (described below), these targets had various distances and eccentricities when expressed relative to the subject's eyes. We ensured that the target flash always stimulated both retinas and always lay within an oculomotor range of ±40°. Furthermore, a central fixation light (blue LED) was centered between the two eyes at 2 m distance when the subject was in central position. Four other targets (red LEDs), also 2 m from the subject, showed the subject where to point his or her nose
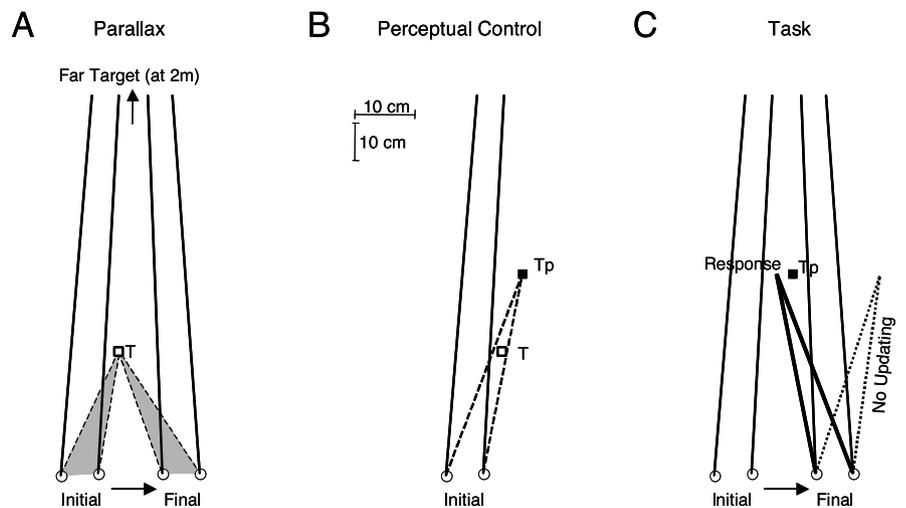

Figure 1. Is motion parallax incorporated in the updating of spatial memory? *A*, When an observer translates sideways, a near object moves through a larger visual angle than a far object because of parallax. *B*, Our subjects foveated a far central target 2 m away. A near target flashed at location *T*, although subjects typically mislocalized it at point *T*$_p$. *C*, The subject translated the head rightward, still fixating the far target, then looked to the remembered location of the near, flashed target. If the subject failed to compute parallax, his or her final gaze would be directed as indicated by the dotted lines. However, in fact, subjects looked to the remembered location (indicated by thick lines), suggesting a nearly perfect updating of the internal representation of space during the head translation.

("nose directing" targets) during the experiment (see below). These targets lay on either side of the central fixation light at 7.5 and 15 cm (for the pure head translation paradigm) or at 35 and 50 cm (for the head rotation paradigm).

Experimental paradigm. We studied the case in which neural computation of parallax, or a failure to compute it, would be most obvious and physiologically relevant: subjects had to update the locations of near targets while translating through space and looking at far targets. The reverse situation (looking at near targets and updating far ones) would involve the same computations and equal parallax but seemed less relevant to real life, in which large head movements more often accompany visual fixations of far targets than near ones, and tracking near objects is usually more relevant for planning motor action.

Updating was tested for two types of eye translation: that caused by lateral head translations and that caused by horizontal head rotations. In the first experiment, subjects actively translated their heads along the interaural y-axis of the coordinate system by bending sideways at the lower back and neck. In the second experiment, they actively rotated the head about its natural axis (near the back of the head), causing the eyes (near the front of the head) to translate along an arc.

For pure translations, in each trial, the subject began by fixating the illuminated far central target (F) while the head was in its central position (Fig. 2). The subject then translated the head sideways (by moving the trunk) to one of four positions, indicated by a nose LED (flashed for 1 sec), so that the nose was pointing at this LED. While the subject maintained this head position (e.g., 10 cm right from the central reference position), ~3 sec after the start of the trial, one of the near targets flashed for 200 msec. Then, within a 2 sec time period, the subject had to translate the head in the opposite direction from the central head position (e.g., 10 cm left from the reference position), while still fixating the far central light. At the end of this 2 sec period, the central light was extinguished and an audio tone cued the subject to look to the remembered location of the flashed target, keeping the head still. The horizontal head rotation paradigm was performed in a similar manner, the only difference being the positions of the nose LEDs (10° and 15° eccentric). In a control test, subjects performed the same task without intervening eye translations to estimate their perceived locations of the targets. For all test conditions, each trial lasted 7 sec. Target and head position combinations were selected randomly. In the motion conditions, for each of the four head positions, all targets were tested three times each, for a total of 108 trials.
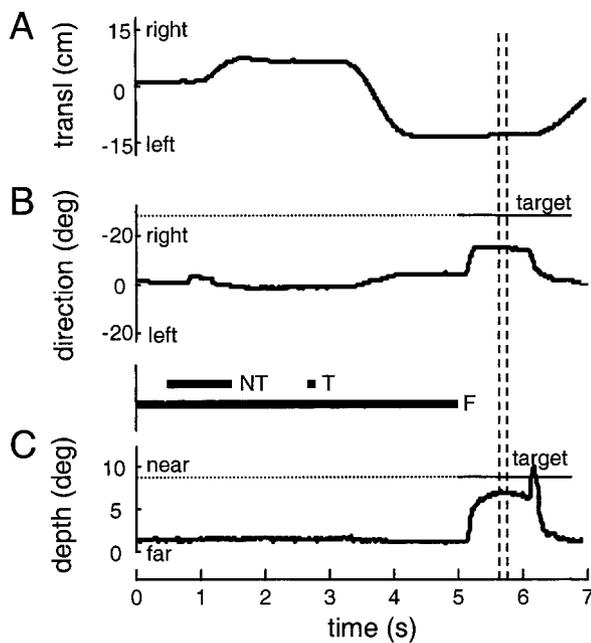
Figure 2. A typical trial illustrating the translational updating paradigm. *A*, The translational motion of the eye in space (in centimeters). *B*, The pointing direction of the eyes (version component in degrees). *C*, The movement of the eyes in depth (vergence in degrees). Thin horizontal lines indicate the direction and depth of the target relative to the eyes after the motion. Filled bars mark the durations of the fixation target (F), the nose target (NT), and the flashed target (T). Vertical dashed lines mark the period over which the saccadic updating response was taken. See Materials and Methods for additional explanation.

In the control condition, all targets were tested once for each head position, yielding 36 control trials.

Data analysis. Data were analyzed using Matlab software (Mathworks, Natick, MA). We analyzed only horizontal motion (horizontal eye and head rotations and translations and target locations in the transverse plane), because planar motions suffice for our purpose: looking for parallax in spatial updating. In the pure translation experiment, head translation along the other axes averaged just 1.1 cm (SD = 1.5 cm) for the *x*-axis and 1.9 cm (SD = 1.5 cm) for the *z*-axis. We described the eye movements in a binocular coordinate system that distinguishes eye movements in direction (version) from eye movements in depth (vergence). Version angle (the conjugate part) was computed from left (L) and right (R) eye-in-head orientation data as (L + R)/2. Vergence angle was calculated as L − R, indicating the angle between the gaze directions of the two eyes, and is a measure for the binocular fixation distance of the subject. We determined where the subject was looking in depth (D) using $D = \sqrt{[(b \times b) + (I \times b \times \sin[R]) + (I \times I/4)]}$ and $b = I \times \cos(L)/\sin(v)$, where I indicates interocular distance, v is vergence angle, L is the gaze angle of the left eye, and R is the gaze angle of the right eye.

In their response, subjects made a saccade to fixate at the remembered location of the flashed target (see above). The endpoint of the saccade was detected based on visual inspection of both the direction and depth component. Because the change in depth was generally much slower than the change in direction, the endpoint (averaged across 10 samples) was taken when both the fast version and slow vergence components of the saccade were completed (generally within 1 sec after the start of the saccade as indicated by the vertical dashed lines in Fig. 2). Then we characterized each response in terms of both binocular fixation direction (in degrees) and binocular fixation depth (either in degrees or centimeters).

The geometric variables required to compute optimal spatial updating were determined using the instantaneous locations of both eyes with respect to the target (from OPTOTRAK data; Northern Digital). In other words, these variables provided the direction and depth of the target relative to the center of the subject's eyes (cyclopean eye).
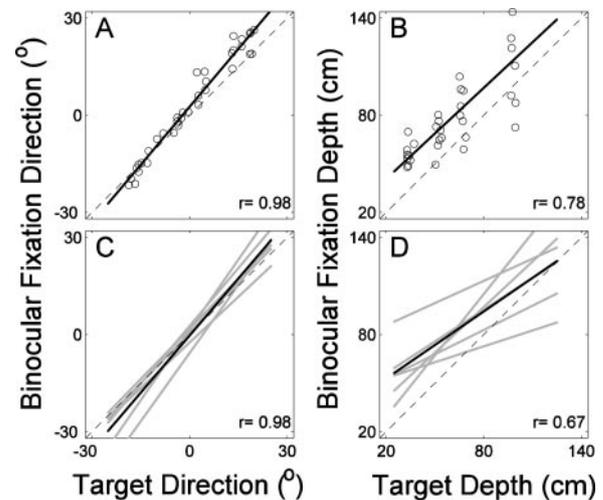

Figure 3. Characteristics of target localization. Binocular fixation direction versus target direction (*A, C*) and binocular fixation depth versus target depth (*B, D*) are shown. *A, B*, Raw data of one subject together with the fitted regression line. *C, D*, The regression lines of all subjects (gray lines), together with the average across all subjects (black lines).

Results

Initial mislocalization of target position

In a control study, we first measured our subjects' perception and eye control in the absence of translation, as depicted in Figure 1*B*. The subject fixated a distant target 2 m away and a nearby light at location *T* (at 20 cm, marked by an open square) was flashed for 200 msec onto the retinal periphery. The subject remained stationary and, after 2 sec, looked to the remembered location of the flash. A filled square marks the spot T_p, where this subject localized the flash, indicating that this subject substantially overestimated the distance of the near target. This mislocalization of target position was observed in all of our subjects, as demonstrated in Figure 3.

Figure 3*A,B* shows the control data of the same subject, plotting measured binocular fixation direction versus target direction (Fig. 3*A*) and binocular fixation depth versus target depth (Fig. 3*B*). As shown, the directions of the target were better judged than their depths. We used linear regression to relate subjects' fixation direction and depth to the actual direction and depth of the vanished target. For direction, the intercept and slope were 2.5° (SD = 0.5°) and 1.19 (SD = 0.04), respectively; for depth, they were 22 cm (SD = 9 cm) and 0.94 (SD = 0.14), respectively. This analysis revealed significant correlations for both the direction ($r = 0.98$) and depth ($r = 0.78$) components in this subject ($p < 0.001$).

Figure 3*C,D* shows the regression lines for all six subjects (gray lines) superimposed on the mean regression line (black line) across subjects for both direction and depth. In all subjects (Fig. 3*C*), the directions of the eye and the flashed target were closely correlated ($r > 0.96$; $p < 0.001$); averaged across subjects, the relationship was well fitted by a line of slope 1.17 (SD = 0.21) and intercept −0.3° (SD = 3.2°), indicating that target direction was fairly well perceived. Target distance was judged somewhat less accurately. In all subjects (Fig. 3*D*), however, binocular fixation depth and target depth were significantly correlated ($0.46 < r < 0.81$; $p < 0.01$); the average relationship had slope 0.69 (SD = 0.35) and intercept 39 cm (SD = 24 cm). As did subjects in previous studies (Komoda and Ono, 1974; Gogel, 1977; Philbeck and Loomis, 1997), our subjects underestimated the depths of distant targets and overestimated the depths of near targets, per-
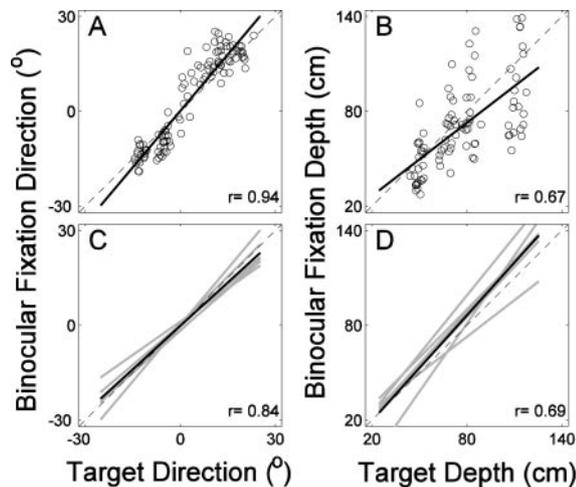
Figure 4. Spatial updating performance. Direction (*A, C*) and depth (*B, D*) of binocular fixation versus the values that would be predicted if the subject perfectly updated the perceived location of the target are shown. *A, B*, One subject's performance, together with the fitted regression line. *C, D*, Regression lines of all subjects (gray lines), together with the average across all subjects (black lines).

haps because of the impoverished visual stimuli (flashed lights in the dark). Because the brain can only be expected to update information that it has actually received, we fitted each subject's eye positions as a mathematical function of flash location so that in the following studies, we could estimate where they initially localized the flash. As we show in the next section, the data support the idea that these eye positions reflect subjects' perceptions of target location rather than their motor inaccuracy in directing their eyes.

Updating during head translation

What happens when the head translates? Figure 1*C* shows a bird's eye view of a typical trial. Again, the subject fixated a distant target and the same near target was flashed for 200 msec. T_p marks the spot where this subject localized the flash (according to the control study; Fig. 1*B*). Next, within 2 sec, the subject translated rightward while maintaining fixation on the distant target. That target was then extinguished, and the subject looked to the remembered location of the flash. Thick lines show the final gaze point. Although it is far from the actual site of the flash, it is close to the point at which the eyes converged in the control task, when the head did not move (T_p). In other words, the subject mislocalized the target, but then almost perfectly updated this stored target location.

The same happened in all trials at all distances tested. Figure 4*A,B* plots the direction (*A*) and distance (*B*) of fixation versus the values that would be predicted if the subject perfectly updated the perceived location of the target. For direction, the correlation was 0.94 and the slope was very close to the ideal value of 1. For distance, the correlation was lower, at 0.67, but the slope was again close to 1.

Results for all subjects are plotted in Figure 4*C,D* (gray lines), together with the average across all subjects (black). For all subjects, we found significant correlations for both direction ($0.71 < r < 0.94$; $p < 0.001$) and depth ($0.59 < r < 0.85$; $p < 0.001$). On average, the correlation was better for direction ($r = 0.84$) than for depth ($r = 0.69$), but in both cases, the slopes were close to the ideal value of 1. Averaged across subjects, the slope was 0.91 (SD = 0.16) for direction and 1.11 (SD = 0.25) for depth. This means that except for some variability, the updating
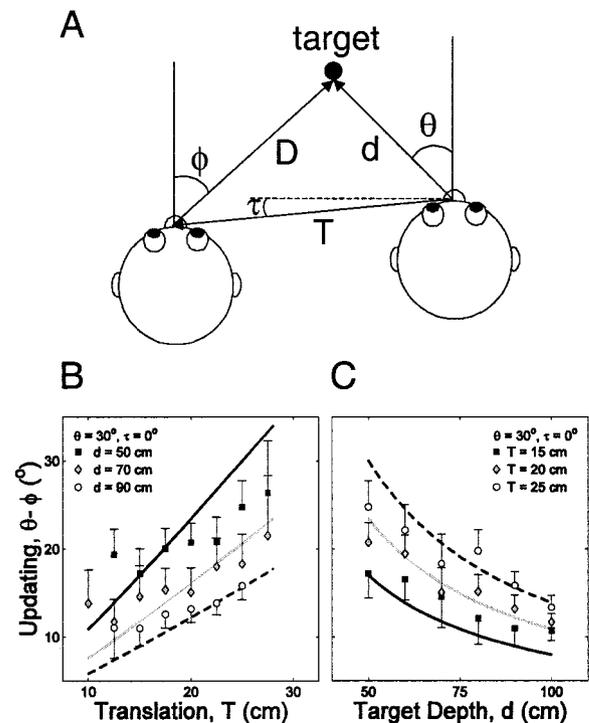

Figure 5. Translational-depth geometry. *A*, Geometry of translational updating. Target angle (*φ*) and target distance (*D*) after translation depend on four parameters: the initial distance of the target from the eyes (*d*), its direction relative to the eyes (*θ*), the translation of the eyes (*T*), and the direction of eye translation (*τ*). The relationships are given by Equations 1 and 2. *B, C*, The average amount ± SE of updating (*θ − φ*) of six subjects (data binned) for updating targets at initial direction (*θ*) of 30°. They show the same nonlinear patterns as perfect updating (lines).

algorithm accurately simulated motion parallax. All subjects initially misjudged the locus of the target, presumably because distance cues were so unnaturally impoverished in our experiment, but thereafter, they accurately simulated the motion, relative to themselves, of the perceived locus.

One could argue that the errors in the control test reflected motor inaccuracies in the eye movements rather than perceptional errors. In other words, the subject's saccadic goal could have been at a different location than the terminus of the actual saccade. To test this possibility, we assumed that it was true, and that the same motor inaccuracy would also occur in the updating condition, and we computed the eye positions that would be expected in the latter task. These predictions failed to match the data. In particular, with these assumptions, the direction of final gaze showed low gains and poor correlations (mean $r = 0.49$) when plotted against the predicted directions. In contrast, the assumption that depth errors reflected perceptional errors yielded unity gains and high correlations (mean $r = 0.84$). Thus, the latter assumption yielded the better description of the data (note, however, that the combination of perceptual and motor errors might do even better).

What is the computational task facing the updating system? Obviously, the required directional updating depends on target depth. A distant target keeps almost the same direction relative to the eyes during translations, so almost no updating is required. For closer targets, the required updating depends nonlinearly on target distance, target eccentricity, and eye translation. Figure 5*A* depicts the variables involved in this geometry. Final target angle (*φ*) (the direction of the target relative to the eyes after the trans-
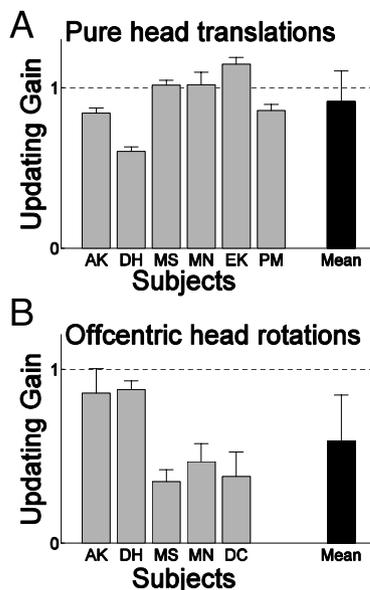
Figure 6. Updating gains of all subjects. Translational updating works better for eye translations attributable to head translations (*A*) than for eye translations resulting from head rotations (*B*). A gain of 1 indicates perfect updating (shown by a horizontal dashed line). Error bars indicate SDs.

lation, or in other words, the final orientation needed to point the eye at the target site) and final target distance (*D*) depend nonlinearly on four parameters: the initial distance of the target from the eyes (*d*), its direction relative to the eyes (*θ*), the translation of the eyes (*T*), and the direction of eye translation (*τ*). These relationships are described by the following two functions:

$$D = \sqrt{(d^2 + T^2 - 2dT\sin(\theta - \tau))} \qquad (1)$$

$$\phi = \arcsin[(T\cos\tau - d\sin\theta)/D]. \qquad (2)$$

The lines in Figure 5*B,C* show the ideal relationship between updating responses (described as *θ* − *φ*) and eye translation (*T*) for targets at different initial distances (*d*) and an initial direction (*θ*) of 30°. As shown, the required updating response increases with increasing eye translation (Fig. 5*B*) and decreases with increasing target distance (Fig. 5*C*).

How well do subjects match these ideal relationships? Averaged data from six subjects (Fig. 5, ■, ○, ◇) for a target that was initially perceived to be 30° left show that the amount of updating varied as predicted with head translation (Fig. 5*B*) and perceived target distance (Fig. 5*C*). In other words, human updating showed the same nonlinear patterns as perfect updating. This key result shows that the updater combines translational and depth information remarkably well when it computes object motion during eye and head translation.

To summarize these findings, we fitted to each subject's data a variant of Equation 2 of the form *φ* = arcsin [*G* (*T*cos*τ* − *d*sin*θ*)/ *D*], with *G* as a free parameter to quantify interindividual differences. *G* can be regarded as an updating gain, because it characterizes how well the updater handles translational and target-depth information. If *G* = 1, the subject perfectly combines depth and translation to update the percept of target location. The fits showed high correlations: averaged across subjects, *r* = 0.90. Updating gains, shown in Figure 6*A*, averaged 0.91 (SD = 0.20), a value not significantly different from the ideal of 1 (*p* = 0.31; *t* test), indicating that the brain takes translational-depth geometry into account when updating spatial information.

Eye translation caused by head rotation

As stated previously, the eyes can translate as a result of either head translations or head rotations. When the head translates, otoliths sense the motion. However, when the head rotates horizontally around its natural rotation axis, there is little, if any, otolith stimulation. In that case, the updater would have to estimate eye translation from other sources of information, including the amount of head rotation and the distance of the eye from the axis of head rotation (Medendorp et al., 1998).

To determine whether updating compensates for eye translation resulting from head rotation, we tested five subjects in our head rotation paradigm. We performed the same analysis as above (i.e., again correcting for the perceptual errors and fitting Eq. 2). Results are shown in Figure 6*B*: the fits to Equation 2 yielded high correlations, with *r* = 0.97, averaged across subjects, but the updating gains averaged just 0.57 (SD = 0.25), which was significantly higher than 0 (*t* test; *p* < 0.01) but lower than the ideal value of 1 (*t* test; *p* < 0.05). This indicates that spatial memory computes parallax less exactly when the eye translation is caused by head rotation. One reason could be that the potential errors are so much smaller than for translational head motion. When we analyzed the response errors in this task, they were smaller (2.7° across subjects) than in the pure translation task (4.9° across subjects). In this sense, then, the system was actually more accurate in the head-rotation task.

Discussion

It is known that the brain correctly updates the locations of objects in spatial memory when the eyes rotate, for both smooth pursuit and saccadic eye and head movements (Hallet and Lightstone, 1976; Schlag et al., 1990; Blouin et al., 1998; Herter and Guitton, 1998; Smith and Crawford, 2001). Our previous study (Medendorp et al., 2002a) showed that during eye and head rotations, the brain compensates almost perfectly for the motions of objects in spatial memory (i.e., gain of >0.98). The present results show that spatial updating also works during translational motion, which is a much more demanding task computationally. As in motion parallax (Howard and Rogers, 1985; Regan, 1991), the required updating varies from object to object, depending nonlinearly on depth and direction. Our results match the predicted nonlinear patterns very well (i.e., gain ∼0.94). Our subjects translated in darkness and looked between vanished targets, so they could not see the parallax but had to compute it internally. We conclude that the brain simulates the geometry of motion parallax when it updates its three-dimensional representation of visual space.

Even when the targets were visible, they were small lights in an otherwise dark room, so the visual cues to depth were sparse. One consequence of these impoverished stimuli was that subjects consistently misperceived the initial location of the target. However, they correctly updated that perceived location; they computed how an object in that location would have moved relative to the head, using information about head motion. The sources of this information are unknown, but three possibilities are the vestibulum, efference copy, and proprioceptors in the neck muscles (Medendorp et al., 2000).

Relationship to previous studies

The circuitry that drives the vestibulo-ocular reflex (VOR) is known to compute the motion of the fixation target relative to the head, including motion parallax, so that it can keep the gaze line trained on the target when the head moves (Blakemore and Donaghy, 1980; Viirre et al., 1986; Angelaki, 1998; Paige et al., 1998;

Medendorp et al., 2002b). Here, we have shown that the brain can update the locations of objects that are not targets of the VOR: in our experiment, gaze remains on the far target throughout the head motion and is redirected to the computed location of the nearer object only afterward. This result demonstrates updating, not within the VOR, but in a more versatile type of spatial memory that can be used to plan and guide later actions by other motor systems (in this case, the saccadic system). It may or may not be the same core circuitry that computes the updated locations for both the VOR and spatial memory.

Israël and Berthoz (1989) showed that people can saccade accurately to a remembered space-fixed target after being passively translated in darkness, but their study did not consider parallax associated with different target depths and translational paths. Our results agree with theirs but reveal new computational capabilities in spatial memory.

Our study is also relevant to locomotion and navigation, which require similar updating of spatial information (Berthoz and Viaud-Delmon, 1999). Several studies in humans (Israël et al., 1993, 1997; Amorim et al., 1997; Philbeck and Loomis, 1997) have shown that subjects slightly underestimate self-motion when they have no visual cues. These findings agree with our observation that people slightly underestimate the size of the eye movement needed to compensate for their head translation (their average gain was slightly smaller than 1.0). Of course it remains to be seen how accurately parallax is computed in a variety of natural activities such as walking and running and so on, but our study establishes the basic fact that spatial memory is capable of using nonvisual data on egomotion to compute motion parallax.

Modeling and neural implications
Psychophysical studies suggest that humans use an eye-fixed coordinate system for representing and updating targets in both near and far space (Henriques et al., 1998; Medendorp and Crawford, 2002). Moreover, many brain regions, such as extrastriate visual areas (Nakamura and Colby, 2002), posterior parietal cortex (Duhamel et al., 1992; Batista et al., 1999; Colby and Goldberg, 1999; Medendorp et al., 2003), frontal cortex (Goldberg and Bruce, 1990), and superior colliculus (Walker et al., 1995) have been shown to store object locations in eye-fixed coordinates. In theory, updating could be performed in any coordinate frame, as long as the correct updating signals and operations are used. However, we chose to model parallax-sensitive updating in an eye-fixed frame so that specific experimental predictions could be generated for the physiological structures discussed above.

Figure 7A shows a schematic of the model. A motion generator selects one target from a retinotopic map of many objects and generates a coordinated, three-dimensional eye-head gaze shift toward that target (Tweed, 1997); it can also translate the head. Furthermore, the motion generator can be driven by voluntary commands, as in the present study. In generating the motion, it provides feedback signals about eye and head motion (based on physiological signals found in the brainstem), which are transformed into eye coordinates and fed into a network that updates all of the objects in the retinotopic map. Our simulations of up-
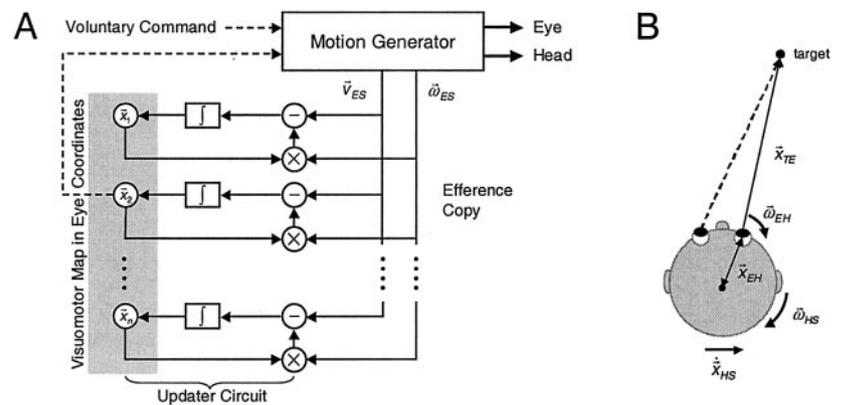

Figure 7. Retinotopic remapping. *A,* A target is selected from an eye-centered target map $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)$, and its retinal error is passed to a motion generator that can rotate the eyes and head as well as translate the head. The motion generator can also be driven using voluntary commands. In generating the motion, it provides feedback signals about eye and head motion, as given by $\vec{\omega}_{ES}$ (the rotary velocity of the eyes in space, $\vec{\omega}_{ES} = \vec{\omega}_{HS} + \vec{\omega}_{EH}$) and $\vec{v}_{ES}$ (the translational velocity of the eyes, $\vec{v}_{ES} = \vec{\omega}_{HS} \times \vec{x}_{EH} \times \vec{x}_{HS}$, with *x* indicating the vector cross-product). Both signals, computed in eye coordinates, are used to update the retinotopic target map. Updating circuits use local information about target depth and direction (as stored in the target map) to determine how that same representation is remapped during the motion. *B,* The derivation of the updating equation (Eq. 3). If the eyes are rotating, the target moves in opposite direction relative to the eyes, according to $\dot{\vec{x}}_{TE} = \vec{x}_{TE} \times \vec{\omega}_{ES}$. Likewise, when the eyes are translating, the target moves with opposite velocity relative to the eyes, as given by $\dot{\vec{x}}_{TE} = -\vec{v}_{ES} = \vec{x}_{EH} \times \vec{\omega}_{HS} - \dot{\vec{x}}_{HS}$. Combining the two equations yields Equation 3.

dating during torsional head rotations (Medendorp et al., 2002a) have suggested that updating is not a simple vectorial shift of stored locations in the brain (Goldberg and Bruce, 1990) but rather a nonlinear process that takes into account both the direction of eye rotation and the location of the target on the map. In the present study involving translations of the eye through space, because of either head rotation or translation, updating becomes even more complex.

The updating equation (derived from Fig. 7B) is:

$$\dot{\vec{x}}_{TE} = \vec{x}_{TE} \times \vec{\omega}_{ES} + \vec{x}_{EH} \times \vec{\omega}_{HS} - \dot{\vec{x}}_{HS}, \qquad (3)$$

which describes how the translational velocity of a target relative to the eye ($\dot{\vec{x}}_{TE}$) depends on the current location of that target relative to the eye ($\vec{x}_{TE}$), the rotary velocity of the eye in space ($\vec{\omega}_{ES}$), the location of the eyeball in the head ($\vec{x}_{EH}$), the rotary velocity of the head in space ($\vec{\omega}_{HS}$) and the translational velocity of the head in space ($\dot{\vec{x}}_{HS}$) (*x* indicates the vector cross-product). The model uses velocity signals instead of position signals so that it can update continuously while the head is moving, for example, during tasks such as walking. The variables on the right side of Equation 3 could be derived from the semicircular canals, otoliths, efference copy, and stored knowledge about head anatomy (Mergner et al., 2001).

In the networks of the brain, of course, Equation 3 likely takes some unfamiliar "distributed" form, but whatever that form, the relevant variables and their overall interaction must be at least approximately as shown. Thus, the equation leads to robust predictions. First, the updater must receive feedback about the six dimensions of head motion ($\vec{\omega}_{HS}$ and $\dot{\vec{x}}_{HS}$) and the 3-D rotation of each eye ($\vec{\omega}_{ES}$), most likely derived from the brainstem and fed through cortical circuits via the thalamus (Smith and Crawford, 2001; Sommer and Wurtz, 2002). Second, the updater must hold information about the 3-D location of the target itself ($\vec{x}_{TE}$), presumably for several potential targets simultaneously (represented in the model by parallel updater circuits) (Fig. 7A).

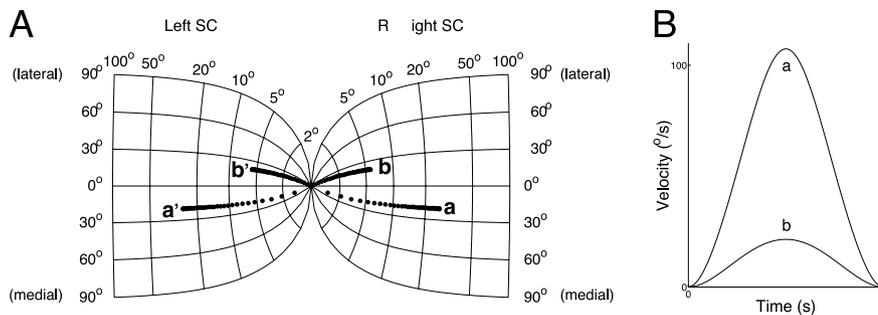Moreover, the model yields quantitative predictions about

A

B

Figure 8. Predictions of the model of remapping in the collicular map during eye translations induced by head translations. *A*, Two space-fixed targets, a (at depth 25 cm) and b (at depth 100 cm), are flashed when the head is in a rightward position (10 cm right from midline). Both stimulate the right side of the retina and are represented, therefore, on the right superior colliculus (SC). However, after the head translates into a leftward position (10 cm, left), the remembered targets are now to the left relative to the retina. In the collicular map, then, their representations must cross the midline, from the right to the left SC. However, for each target, the amount of remapping depends on the target depth: target a, closer to the eye, requires more remapping (a longer trajectory on the map) than target b, which is at farther distance from the eye. *B*, During translations, eye-centered remapping requires target representations to move at different speeds across the SC map, depending on the depth of each target.

activity patterns in neural structures such as the superior colliculus (Munoz et al., 1991; Walker et al., 1995), implying that they should evolve during head motion in a way that depends on object depth. Figure 8*A* shows the prediction of the model for two target representations on that map during a head translation. As the simulation shows, remapping depends on depth: target a, close to the eyes, follows a longer, faster trajectory across the map than target b, which is farther from the eyes. So during translations, eye-centered target representations migrate at different speeds across the colliculus depending on their depths (Fig. 8*B*). A direct physiological demonstration of this would prove that each 3-D target representation contributes to its own updating. Although this has not been shown, consistent with these ideas, there is evidence that the superior colliculus does receive depth information from cortical lateral intraparietal area (Gnadt and Beyer, 1998). Alternatively, collicular remapping could reflect computations in upstream cortical maps, such as in the posterior parietal cortex, which is known to receive information about target depth (Gnadt and Mays, 1995; Gnadt and Beyer, 1998), head motion (Snyder et al., 1998; Andersen et al., 1999), head position (Brotchie et al., 1995), and eye position (Andersen et al., 1985).

Together, our present study of parallax and our previous report on rotary updating (Medendorp et al., 2002a) suggest that the brain computes an evolving representation of remembered 3-D space, at least for those targets we choose to remember or act on (Henriques et al., 1998). As illustrated by our model, this representation must be updated in a nonlinear manner that incorporates the 3-D kinematics of eye and head motion, information about the 3-D location of individual visual targets, and the geometry of motion parallax.

References

Amorim MA, Glasauer S, Corpinot K, Berthoz A (1997) Updating object's orientation and location during nonvisual navigation: a comparison between two processing modes. Percept Psychophys 59:404–418.

Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. Science 230:456–458.

Andersen RA, Shenoy KV, Snyder LH, Bradley DC, Crowell JA (1999) The contributions of vestibular signals to the representations of space in the posterior parietal cortex. Ann NY Acad Sci 871:282–292.

Angelaki DE (1998) Three-dimensional organization of otolith-ocular reflexes in rhesus monkeys. III. Responses to translation. J Neurophysiol 80:680–695.

Batista AP, Buneo CA, Snyder LH, Andersen RA (1999) Reach plans in eye-centered coordinates. Science 285:257–260.

Berthoz A, Viaud-Delmon I (1999) Multisensory integration in spatial orientation. Curr Opin Neurobiol 9:708–712.

Blakemore C, Donaghy M (1980) Coordination of head and eyes in the gaze-changing behaviour of cats. J Physiol (Lond) 300:317–335.

Blouin J, Labrousse L, Simoneau M, Vercher JL, Gauthier GM (1998) Updating visual space during passive and voluntary head-in-space movements. Exp Brain Res 122:93–100.

Brotchie PR, Andersen RA, Snyder LH, Goodman SJ (1995) Head position signals used by parietal neurons to encode locations of visual stimuli. Nature 375:232–235.

Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. Annu Rev Neurosci 22:319–349.

Duhamel JR, Colby CL, Goldberg ME (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. Science 255:90–92.

Gnadt JW, Beyer J (1998) Eye movements in depth: what does the monkey's parietal cortex tell the superior colliculus? NeuroReport 26:233–238.

Gnadt JW, Mays LE (1995) Neurons in monkey parietal area LIP are tuned for eye-movement parameters in three-dimensional space. J Neurophysiol 73:280–297.

Gogel WC (1977) An indirect measure of perceived distance from oculomotor cues. Percept Psychophys 21:3–11.

Goldberg ME, Bruce CJ (1990) Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. J Neurophysiol 64:489–508.

Hallet PE, Lightstone AD (1976) Saccadic eye movements to flashed targets. Vision Res 16:107–114.

Heide W, Blankenburg M, Zimmermann E, Kompf D (1995) Cortical control of double-step saccades: implications for spatial orientation. Ann Neurol 38:739–748.

Henriques DYP, Klier EM, Smith MA, Lowy D, Crawford JD (1998) Gaze-centered remapping of remembered visual space in an open-loop pointing task. J Neurosci 18:1583–1594.

Herter TM, Guitton D (1998) Human head-free gaze saccades to targets flashed before gaze-pursuit are spatially accurate. J Neurophysiol 80:2785–2789.

Howard IP (1982) Human visual orientation. New York: Wiley.

Howard IP, Rogers BJ (1985) Binocular vision and stereopsis. New York: Oxford Claridon.

Israël I, Berthoz A (1989) Contribution of the otoliths to the calculation of linear displacement. J Neurophysiol 62:247–263.

Israël I, Chapuis N, Glasauer S, Charade O, Berthoz A (1993) Estimation of passive linear whole body displacement in humans. J Neurophysiol 70:1270–1273.

Israël I, Grasso R, Georges-Francois P, Tsuzuku T, Berthoz A (1997) Spatial memory and path integration studied by self-driven passive linear displacement. I. Basic properties. J Neurophysiol 77:3180–3192.

Komoda M, Ono H (1974) Oculomotor adjustments and size-distance perception. Percept Psychophys 15:353–360.

Medendorp WP, Crawford JD (2002) Visuospatial updating of reaching targets in near and far space. NeuroReport 13:633–636.

Medendorp WP, Melis BJM, Gielen CCAM, Van Gisbergen JAM (1998) Off-centric rotation axes in natural head movements: implications for vestibular reafference and kinematic redundancy. J Neurophysiol 79:2025–2039.

Medendorp WP, Van Gisbergen JA, Van Pelt S, Gielen CC (2000) Context compensation in the vestibuloocular reflex during active head rotations. J Neurophysiol 84:2904–2917.

Medendorp WP, Smith MA, Tweed DB, Crawford JD (2002a) Rotational remapping in human spatial memory during eye and head motion. J Neurosci 22:RC196(1–4).

Medendorp WP, Van Gisbergen JA, Gielen CC (2002b) Human gaze stabilization during active head translations. J Neurophysiol 87:295–304.

Medendorp WP, Goltz HC, Vilis T, Crawford JD (2003) Gaze-centered updating of visual space in human parietal cortex. J Neurosci 23:6209–6214.

Mergner T, Nasios G, Maurer C, Becker W (2001) Visual object localisation in space. Interaction of retinal, eye position, vestibular and neck proprioceptive information. Exp Brain Res 141:33–51.

Munoz DP, Pelisson D, Guitton D (1991) Movement of neural activity on the superior colliculus motor map during gaze shifts. Science 251:1358–1360.

Nakamura K, Colby CL (2002) Updating of the visual representation in monkey striate and extrastriate cortex during saccades. Proc Natl Acad Sci USA 99:4026–4031.

Paige GD, Telford L, Seidman SH, Barnes GR (1998) Human vestibuloocular reflex and its interactions with vision and fixation distance during linear and angular head movement. J Neurophysiol 80:2391–2404.

Philbeck JW, Loomis JM (1997) Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. J Exp Psychol Hum Percept Perform 23:72–85.

Regan D (1991) Binocular vision. In: Vision and visual dysfunction, Vol IX. London: Macmillan.

Schlag J, Schlag-Rey M, Dassonville P (1990) Saccades can be aimed at the spatial location of targets flashed during pursuit. J Neurophysiol 64:575–581.

Smith MA, Crawford JD (2001) Implications of ocular kinematics for the internal updating of visual space. J Neurophysiol 86:2112–2117.

Snyder LH, Grieve KL, Brotchie P, Andersen RA (1998) Separate body- and world-referenced representations of visual space in parietal cortex. Nature 394:887–891.

Sommer MA, Wurtz RH (2002) A pathway in primate brain for internal monitoring of movements. Science 296:1480–1482.

Tweed D (1997) Three-dimensional model of the human eye-head saccadic system. J Neurophysiol 77:654–666.

Viirre E, Tweed D, Milner K, Vilis T (1986) A reexamination of the gain of the vestibuloocular reflex. J Neurophysiol 56:439–450.

Von Helmholtz H (1867) Handbuch der Physiologischen Optik. Hamburg, Germany: Voss.

Walker MF, Fitzgibbon J, Goldberg ME (1995) Neurons of the monkey superior colliculus predict the visual result of impeding saccadic eye movements. J Neurophysiol 73:1988–2003.