Anticipating the three-dimensional consequences of eye movements

Mark Wexler*

Laboratoire de Physiologie de la Perception et de l’Action, Centre National de la Recherche Scientifique, Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France

Communicated by Richard M. Held, Massachusetts Institute of Technology, Cambridge, MA, December 13, 2004 (received for review February 16, 2004)

Rapid eye movements called saccades give rise to sudden, enormous changes in optic information arriving at the eye; how the world nonetheless appears stable is known as the problem of spatial constancy. One consequence of saccades is that the directions of all visible points shift uniformly; directional or 2D constancy, the fact that we do not perceive this change, has received extensive study for over a century. The problems raised by 3D consequences of saccades, on the other hand, have been neglected. When the eye rotates in space, the 3D orientation of all stationary surfaces undergoes an equal-and-opposite rotation with respect to the eye. When presented with an optic simulation of a saccade but with the eyes still, observers readily perceive this depth rotation of surfaces; when simultaneously performing the corresponding saccade, the 3D orientations of surfaces are perceived as stable, a phenomenon I propose calling 3D spatial constancy. In experiments presented here, observers viewed ambiguous 3D rotations immediately before, during, or after a saccade. The results show that before the eyes begin to move the brain anticipates the 3D consequences of saccades, preferring to perceive the rotation opposite to the impending eye movement. Further, the anticipation is absent when observers fixate while experiencing optically simulated saccades, and therefore must be evoked by extraretinal signals. Such anticipation could provide a mechanism for 3D spatial constancy and transsaccadic integration of depth information.

depth perception | vision | saccades | spatial constancy

Directional or 2D spatial constancy holds insofar as the uniform, 2D shifts of the retinal image accompanying each eye movement (see Fig. 1a) do not lead to perceptions either of change in direction of individual points or of 2D motion. This type of spatial constancy was noted by Descartes in his Traité de l’Homme and has been studied systematically since the 19th century (refs. 1–3 and see refs. 4–8 for reviews). Distortions in spatial vision in the temporal vicinity of saccades are considered as signs of processes that give rise to constancy. For example, the threshold for perceiving motion rises just before a saccade (9, 10), which may be why motion is usually not perceived after saccade-induced retinal shifts, and may allow for optimal transsaccadic integration of information (11). Furthermore, points flashed around the time of a saccade are systematically mislocalized in a way that suggests slow build-up of compensation for retinal shifts (12–19). The dynamic properties of neurons that remap their receptive fields in the anticipation of saccades may be closely connected with these distortions and contribute to spatial constancy (7). Such neurons have been found in posterior parietal cortex (20–22), superior colliculus (23), and the frontal eye field (24) of monkeys; recently, neuroimaging has demonstrated similar spatial updating in human parietal cortex (25).

An aspect of the spatial constancy problem that has received little or no attention is the 3D stability of the world during eye movements. Vision serves not only for detecting the directions of points, but, at least as importantly, for extracting the 3D layout of the environment, and in particular the orientations of surfaces (26). The 3D consequences of saccades have been neglected probably because in humans eye rotations result in almost no motion parallax (as opposed to head movements, for instance), and therefore, to a good approximation, generate no new 3D information in the optic array (27, 28).

Nonetheless, eye movements do have 3D consequences. When the eye rotates in space, the 3D orientations of all stationary surfaces undergo an equal-and-opposite rotation in the reference frame of the eye (see Fig. 1b). To demonstrate that these relative depth rotations† are a potential problem for the brain, it is instructive to carry out simulated saccades, in which observers are shown the optical consequences of a saccade, but while keeping the gaze fixed. Stimuli for simulated saccades can easily be created by taking a photograph, rotating the camera about its optical center while keeping the scene unchanged, and taking a second photograph.

An example of a simulated saccade is shown in Fig. 1c, in which the camera was rotated by ∼10° to the right, a typical amplitude for a human saccade, while the scene remained stationary. When the two images are shown in rapid succession to observers whose gaze is fixed (see Movie 1, which is published as supporting information on the PNAS web site), they perceive two things. First, they see a lateral, uniform shift of all objects to the left, the standard 2D shift. Second, they also perceive a depth rotation of the large surfaces in the scene, equal and opposite to the rotation of the camera. On the other hand, if the corresponding saccade is actually performed at the same time (e.g., by executing the saccade that corresponds to the camera movement in Fig. 1c), neither the leftward shift nor the depth rotation is normally perceived. The perception of depth rotation of surfaces in simulated saccades shows that the brain faces a problem of constancy of 3D surface orientation during eye movements; the absence of this perception during real saccades shows that the brain usually solves the problem. I propose to call this phenomenon 3D spatial constancy. The simulated-saccade demonstration strongly suggests that extraretinal signals are involved in the brain’s solution of the 3D spatial constancy problem.‡

An important idea in the psychological and physiological literature on spatial constancy is the link between constancy and sensorimotor anticipation (2, 7, 20, 29, 30): predicting the sensory reafferences of one’s actions could help to distinguish

© 2005 by The National Academy of Sciences of the USA

Abbreviations: FP, fixation point; SfM, structure from motion; SO, stimulus onset; PSML, perisaccadic mislocalization.

†The term depth rotation refers to a rotation about an axis in the image plane, perpendicular to the direction of gaze. Because according to Listing’s law, which is approximately true, the eye also rotates about an axis this plane, relative rotations between stationary surfaces and the eye will be (mostly) depth rotations.

‡This demonstration can be refined in several ways. First, instead of showing just two frames, the camera can be made to follow a typical saccadic trajectory. Second, the observer should be placed so that the angular displacement of the images equals the camera rotation. Third, the images should cover as large a part of the visual field as possible. Finally, one should eliminate static borders and any other stationary objects in the visual field that might lead to a perception of relative rotation. With or without these refinements, observers report strong perceptions of surface rotations in simulated saccades.

E-mail: wexler@ccr.jussieu.fr.
them from external events. As mentioned above, there is psychophysical (12) and neurophysiological (20) evidence of the anticipation of the 2D consequences of saccades. Does the brain also anticipate the 3D consequences?

In the experiments reported here, anticipation of the 3D consequences of saccades was probed by presenting subjects with ambiguous displays of 3D rotations. These displays were presented immediately before, during, or after saccades. The ambiguity was such that subjects could perceive depth rotation either in the same direction as the eye rotation or in the opposite direction. In control conditions, subjects received the same optic information, either in an allocentric or retinotopic reference frame, but without performing the eye movement. The two perceptual solutions in each ambiguous display were predicted to occur equally often in the control conditions. In the eye movement conditions, on the other hand, anticipation of the 3D consequences of a saccade should show up as a bias toward the solution rotating opposite to the eye, before the onset of the eye movement.

**Methods**

The ambiguous stimulus was a two-frame structure-from-motion (SfM) sequence, in which 3D shape and motion can be perceived from 2D optic flow alone (31, 32). Each proximal stimulus can be the result of infinitely many distal configurations, and this ambiguity is partly broken by the *a priori* hypotheses of rigidity [minimal relative motion (33)] and stationarity [minimal absolute motion (34)]. However, as in the case of the Necker cube or the Mach book, there remain residual ambiguities that lead to reversals of 3D depth and motion.

An example of such an ambiguity is shown in Fig. 2; the corresponding two-frame animation, similar to the actual stimuli, can be found in Movie 2, which is published as supporting information on the PNAS web site. Each of the two frames is composed of a set of dots with a uniform or nearly uniform density in the image plane, therefore offering only trivial depth cues. The displacement of the dots between the first and second frames gives rise to the perception of 3D structure and motion, as can be seen in the animated version. This perception is
ambiguous because a simultaneous inversion of the direction of motion and the direction of surface inclination (known as tilt§) results in approximately the same optic flow (see Fig. 2). Thus, each stimulus is compatible with two solutions, rotating in opposite directions and having tilts that differ by 180°.¶

The stimuli were chosen to be compatible with a plane slanted in depth, undergoing a small depth rotation about a vertical axis. The subjects’ task was to report the perceived surface tilt, from which the perceived direction of rotation can be inferred. For instance, if an observer who is presented with the stimulus in Fig. 2a and b perceives a surface with a tilt to the right, we can conclude that he perceives a 3D rotation such as in Fig. 2c, without inquiring about motion directly. Similarly, a surface tilted to the left goes with the rotation depicted in Fig. 2d. The opposite combination of 3D structure and motion would violate the rigidity hypothesis, known to be closely followed in depth vision (33). The closer the reported tilt to one of the two possible tilts for a given stimulus, the more we can be certain of the perceived direction of rotation. The advantage of this indirect way of measuring the perception of 3D motion by asking about surface orientation is to make the task cognitively opaque and therefore safe from any cognitive bias.

Procedure. The time course of a saccade trial is summarized in Fig. 3. Subjects first fixated a central fixation point (FP), initially accompanied by an arrow indicating saccade direction. The signal to perform the eye movement was the simultaneous appearance of a peripheral saccade target (7.5° to the left or right), the dimming of the central FP, and a brief auditory tone. At the same time, the first frame of the SfM sequence appeared around the central FP. After a variable delay [calculated so that stimulus onset (SO) would be ideally at −100, 0, or +100 ms with respect to predicted saccade onset, the latter calculated from the previous four trials], the first stimulus frame was replaced by the second, which was displayed for 50 ms. In data analysis, SO with respect to actual saccadic onset was used as an independent variable.

§Slant and tilt are a common way of parametrizing the orientation of a plane. Slant is the magnitude of the plane’s inclination from the frontoparallel; tilt is the projected direction of that inclination. For example, the surface in Fig. 2c has tilt 0°, whereas the one in Fig. 2d has tilt 180°, and the two have equal slant.

¶There are two solutions in the space of possible tilts. If the solution space is extended to include arbitrary tilts and axes of rotation.

Ocular trajectories were tested on-line (see below). Trials that did not meet the criteria were aborted and later repeated. On trials that passed, after the disappearance of all stimuli subjects used a joystick to adjust the 3D tilt of a visual probe (centered on the saccade target) to match the average perceived tilt of the SfM stimulus.

The fixation condition was identical to the saccade condition (the same stimuli with respect to the display monitor), except that subjects were required to continue fixating the central FP throughout. SO was timed as if the reaction time had been 250 ms. In the simulated saccade condition, subjects fixated a central FP while receiving the same retinal stimulus as in a previous saccade trial (i.e., stimuli directions rotated about the eye in the opposite direction as the saccade). In both conditions, any trial with a saccade was repeated.

Visual Stimuli. The visual stimulus consisted of two image frames. The first frame was a texture of randomly placed dots with uniform density in the image plane. The second frame was generated by (i) parallel-projecting the first frame onto an inclined plane with slant 45° and tilt chosen from 15°, 45°, . . . , 345°; (ii) rotating this object by 7.5° about a vertical axis through its center; and (iii) parallel-projecting it onto the screen. The dots had a mean density of five dots per deg2 and were displayed on the screen as white circles (radius of 0.03°), and only those dots falling within a circle of radius 2.5° were shown. The two-frame animation in Movie 2 closely

![Fig. 2. The ambiguous stimuli used in the experiments. (a) Two frames of a SfM sequence, similar to the actual stimuli used. The frames can be seen as an animation in Movie 2. (b) The corresponding optic flow (with length of arrows exaggerated). This flow is ambiguous and can have either c or d as interpretation, each of which is a combination of 3D orientation and motion. The two solutions have opposite tilts and directions of rotation. A similar ambiguity holds for arbitrary tilts and axes of rotation.](#)
reflects an actual stimulus with tilt 0° (or 180°). The only difference, other than the timing, is the size of the dots, which has been increased to make them easier to see.

The FPs were displayed as red dots of radius 0.1°, and the arrow that previewed the saccade direction had a length of 2°. The probe used for the tilt response was the perspective projection of a flat object with irregular, star-shaped edges (radius of 2.5° face-on, width of 0.5°). The probe always had a slant of 45°, and its tilt was controlled by a joystick.

Stimuli were displayed on a computer monitor, with pixel size of 1.7 arcmin and a refresh rate of 100 Hz. The experiment was performed in darkness, with the monitor and its edges covered by an attenuating filter, so that nothing other than the stimuli could be seen (including the edges of the monitor).

**Eye Movement Recording.** Eye movements were monitored with a Skalar Iris IR limbus eye tracker (Skalar Medical, Delft, The Netherlands). The horizontal position of the left eye (the right eye being occluded) was digitally sampled at 100 Hz × 12 bits. Head movement was restrained by a chin rest, the eye tracker was calibrated before each block, and the initial FP on each trial was used for drift correction.

Saccade criteria were as follows: eye speed had to attain 100 deg/s, and saccadic onset and offset were defined as the samples at which speed went above or fell below 20 deg/s. Additionally, saccadic onset had to occur between 50 and 600 ms after the go signal, and its amplitude had to be between 0.5 and 1.5 times the 7.5° jump in the FP.

**Subjects.** Eight subjects (six naïve) participated in the saccade condition of the main experiment. Subjects with no experience with SfM tasks performed a preliminary simplified training block. Subjects then performed four blocks in the normal condition, each with 72 trials. Four of the subjects later performed a single block (72 trials per SO). The horizontal position of the left eye (the right eye being occluded) was digitally sampled at 100 Hz × 12 bits. Head movement was restrained by a chin rest, the eye tracker was calibrated before each block, and the initial FP on each trial was used for drift correction.

**Fig. 4.** Results of the main experiment. (a) Mean horizontal eye trajectory (relative to initial position, with leftward saccades multiplied by −1) and standard deviations. Time t = 0 corresponds to saccade onset. (b) Mean bias in 3D rotation perception as a function of SO time (relative to saccade onset, binned by 50 ms). Time scale is the same as in a, so that negative values on the abscissa correspond to SOs before the saccade. The point on the right shows data for the fixation condition. Error bars denote between-subject standard errors. (c) Individual bias data for subjects in the main experiment: before the saccade (means of the −150-, −100-, and −50-ms bins), after the saccade (the 100-, 150-, and 200-ms bins), and in the fixation condition. (d) Tilt error as a function of SO time. Time scale and bins are as in b. (e) Distributions of tilt errors for SOs before (200 ms), during, and after (200 ms) the saccade.
trials) in the fixation condition. Seven subjects (five new) participated in the simulated saccade experiment (four blocks, as in the saccade condition); the subjects who had participated in the saccade condition received their own previous optic flow, whereas the new subjects received optic flow from randomly chosen previous subjects.

**Results**

**Real Saccades.** In a first experiment, subjects were shown the ambiguous 3D stimulus while they were planning, executing, or had just executed a horizontal saccade, either to the left or right. Some subjects also participated in a fixation control condition, in which stimuli were the same relative to the computer monitor, but gaze was kept fixed.

The results are shown in Fig. 4. The time variable for each trial is defined so that \( t = 0 \) corresponds to saccade onset. Fig. 4a shows mean horizontal eye trajectories. The crucial results will be given in terms of bias, defined as the fraction of trials in which subjects reported a tilt compatible with the 3D rotation opposite to that of the eye on that trial (i.e., counterclockwise surface rotation (as seen from the top) for rightward saccades and clockwise rotation for leftward saccades). Fig. 4b shows mean bias in the saccade condition as a function of SO time (relative to saccade onset) and also in the fixation condition. To reduce noise, this analysis was carried out on trials with accurate responses, defined as having tilt error <45°. Chance level for bias is 0.5, which corresponds to the two solutions being chosen equally often. Three points differ significantly from chance (\( P < 0.05 \), Sidak-corrected binomial test), with \( t = -150 \), -100, and -50 ms. For these points, SO occurred on average 190 ms after the go signal. Bias did not differ significantly from chance level for any other SO times, nor for the fixation control condition. Individual subject data are shown in Fig. 4c, with data in the saccade condition grouped into before- and after-saccade bins. There is a significant difference between biases in the before-saccade and fixation conditions (\( t \) test, \( P < 0.001 \)) and between the before- and after-saccade conditions (\( P < 0.01 \)), but not between the after-saccade and fixation conditions.

To check that subjects actually performed the 3D task, tilt error was defined as the absolute value of the angular difference between the tilt response and the closest of the two tilts in the stimulus. Thus defined, tilt error is independent of bias and ranges between 0° and 90° with chance level at 45°. Mean tilt error as a function of SO time is given in Fig. 4d, which shows that performance is better than chance level (45°) before and after the saccade, but drops close to chance during the eye movement (the latter caused by saccadic suppression or failure of transsaccadic integration). Error distributions (Fig. 4e) confirm this analysis of performance. Before and after the saccade the error distribution peaks at 0°, and Kolmogoroff-Smirnov tests show that it is nonuniform (\( P < 0.01 \) in both cases); on the other hand, for trials with SO during the saccade, the same test shows that the distribution does not differ significantly from uniform.

**Simulated Saccades.** The bias demonstrated here could have been caused by either retinal or extraretinal signals associated with the saccade. The fact that it occurs before the saccade points to an extraretinal origin (29, 36, 37). However, a role for retinal flow and smear caused by the saccade cannot be excluded, because, even for trials with SO before the saccade, some retinal signals (e.g., caused by small retinal movements of the saccade target) were present, and the large retinal jump after the stimulus could have resulted in a retroactive or “postdiction” bias (38). To probe the origin of the bias, an additional control experiment was performed with optically simulated saccades (39, 40). Subjects maintained central fixation (as in the fixation condition above), while experiencing approximately the same flow on the retina as in previous saccadic trials. The results (Fig. 5) show that the bias curve is very different in the simulated case than in the case of real saccades: in the case of simulated saccades, although there is a small bias toward antiparallel rotations for stimuli presented before the saccade, it is nowhere significantly different from chance. (The only point that approaches significance is \( t = 0 \), where there is a bias toward parallel rotations.) Thus, the control experiment shows that optic flow in simulated saccades is insufficient to induce antiparallel bias and lends strong support to the hypothesis that the bias is induced by extraretinal signals.

**Discussion**

It has been shown that during the planning of a saccade and \( \approx 150–50 \) ms before its onset, the visual system develops a bias in its perception of 3D rotations, in favor of rotations opposite to the impending rotation of the eye. Because of saccadic suppression, the method used does not allow us to determine whether this effect persists during the saccade. The bias may reappear immediately after the end of the saccade, but here it is weaker and not statistically reliable. The absence of bias in the fixation condition shows that the origin of the effect is not cognitive: simply knowing the saccade direction induces no bias. Nor is the bias caused by subsequent retinal flow and smear: its absence in the simulated saccade condition shows that it is generated by extraretinal eye movement signals. The bias that has been demonstrated is anticipatory, because it is in the same direction as the rotation, in the reference frame of the eye, of all visible surfaces during the upcoming saccade. Thus, before the eye begins to move, the visual system prepares for the reaferent 3D visual consequences of the eye movement.

There are at least two different ways to predict the egocentric surface rotations induced by eye movements, indirectly or directly. An indirect mechanism could use, as an intermediate step, prediction of the 2D shift of directions, for example, a predictive 2D map (30) that dynamically remaps retinal input in anticipation of saccades (e.g., transforming the top image in Fig. 1c into the bottom image, even before the saccade is executed). 3D vision mechanisms then could “read off” the anticipated 3D orientation from the depth cues in the anticipated 2D map. Alternatively, a direct mechanism could act on the high-level neural representations of 3D surface orientation that are known to exist in the brain (41). Although these representations originally are extracted from the retinal image, they may be updated directly in anticipation of a saccade without relying on remapped retinotopic arrays. The advantage of the indirect mechanism is its parsimony in not requiring a specialized mechanism for 3D prediction; the advantage of the direct mechanism is its economy of an intermediate processing step.

A psychophysical effect believed to underlie 2D or directional constancy is perisaccadic mislocalization (PSML) (7, 12, 13, 42). In PSML, the perceived directions of points flashed just before
(≈100–500 ms) a saccade are erroneously perceived as shifted in the same direction as the upcoming eye movement. Could the mechanism that is responsible for PSML, and, supposedly, for 2D constancy (7), also account for the 3D effect reported here? This seems unlikely, because in PSML the presaccadic shift is in the same direction as the eye movement, which would predict a bias toward a 3D rotation also in the same direction in space as the eye, which is precisely the opposite of the effect that has been found. Furthermore, PSML flips direction immediately after the saccade, a flip not observed in the 3D effect. Finally, the time course of the 3D bias also seems to differ from that of a bias in 2D motion perception (43). Thus, if 2D constancy mechanisms induce PSML, as is widely supposed (7, 12), they do not seem to account for the 3D bias.

We seem to be left, therefore, with the hypothesis of a direct mechanism for the prediction of 3D consequences of saccades. A plausible neural mechanism draws inspiration from the anticipatory remapping of 2D neuronal receptive fields in the intraparietal sulcus of monkeys that codes the 3D orientation of surfaces (and in particular, tilt) in a way that seems to be independent of the underlying depth cues (41). Other such maps probably exist elsewhere in the primate brain (44, 45). If some of these 3D representations have anticipatory properties that could support 3D spatial constancy, they could be updated directly by the efference copy of the eye movement command (46), because the egocentric surface rotation induced by the saccade depends on the parameters of the eye movement alone (it is equal and opposite to the eye’s rotation in space). Thus, it could be interesting to explore the dynamic properties of the 3D surface representations in the brain around the time of saccades.

Finally, another use for the anticipation of the 3D consequences of eye movements could be for transsaccadic integration of 3D information. Results on transsaccadic integration have been mixed, with reports of no integration (47–49), as well as evidence for integration of shape and motion across saccades (50–53), with recent results showing integration of 3D information across multiple spatial locations (54). Given the particular slowness of 3D vision (55, 56) compared with the typically brief time between saccades, transsaccadic integration of 3D information would be especially valuable. The difficulty in integrating 3D information across fixations is precisely the egocentric rotation of surfaces with eye movement (Fig. 1b). The anticipatory rotation that has been demonstrated here would keep egocentric surface information up to date. Updating 3D egocentric representations in anticipation of saccades thus could provide a mechanism for both 3D spatial constancy and transsaccadic integration of depth information.

This article has been greatly improved as a result of discussions with R. Held and C. J. Erkelens. I thank J. Droulez and C. Morvan for useful advice and continued support.