The Operation of the Visual System in Relation to Action

Minireview

Michael F. Land

Neurophysiologists studying the visual representation of the world in the parietal lobe generally find that it is based in a gaze-centred (retinotopic) frame. Students of orientation, however, find that the brain also contains a more panoramic egocentric representation that allows appropriate motor actions to take place independent of the orientation of the eyes and head. This representation can operate temporarily without visual input, but is updated from the vestibular system and from other modalities. In this minireview, I shall consider how these two representations are kept aligned with each other, and how they relate to the organisation of motor actions and to the phenomenal world that we see.

Introduction

For foraging monkeys, or humans working in fields or workshops, the main function of vision is to guide actions while the body is in almost continuous motion. This involves eye, head and body movements, which are concerned not only with stabilizing gaze, but directing it to the sources of information needed for the ever-changing requirements of the task in hand [1].

The world we see — the phenomenal world — is based in the brain, even though it seems to be 'out there'. It has the property that it does not move when we rotate out eyes, head and body, or when we move through it. It is devoid of the saccadic dislocations that characterise the retinal image. Objects in it can be ignored or attended to, depending on their relevance to current actions. None of this looks much like the retinotopic representations found in the early stages of the cortical visual pathway. We will thus need to look elsewhere for a complete account of the phenomenal visual world.

Our actions involve orientation movements of eyes, head, body and limbs to objects whose locations may be within the current field of vision — roughly the frontally directed hemisphere — or may be outside it, but accessible from short-term or long-term memory. It is, for example, unproblematic to point, without turning or with eyes closed, to a door or window that is currently in the hemisphere behind us. In this review I will propose an arrangement that attempts to satisfy the major requirements of a visual system that is capable of guiding action from the moving platform of the body, and which at the same time generates a representation of the world that is more like the one we experience visually. Given the presumption that the way we see the world evolved to make the control of action as straightforward as possible, it is likely that our phenomenal perception of the world is closely related to the mechanisms we use to act upon it.

At present there are two research approaches to the ancient problem of why the phenomenal world remains stationary in spite of the motions of the eyes, head and body. Essentially, the first deals with the seen world as presented to the retina, and the other with the more extensive three-dimensional space around us. Over 20 years ago, students of eye movement control discovered that when the eyes make saccades, the pattern of excitation in the lateral intraparietal area (LIP) is 'remapped' *before* movement occurs, to coincide with the disposition of features in the image *after* each saccade [2,3]. These and other authors have taken this to mean that this remapping of image features provides at least a partial explanation for why saccades do not disrupt the appearance of the visual scene. This remapping is now known to occur in many cortical areas as well as the superior colliculus [4], and is linked to an efference copy of the eye movement signal itself which originates in the colliculus [5].

Over much the same time period, students of human orientation and navigation have examined the ways that the representation of the world, on which our motor actions are based, is 'spatially updated' as we move through it. Even when vision is not available, actions such as pointing can be made with reasonable accuracy after a number of rotational and translational body movements. This, they argue, requires an internal model that is kept aligned with the external world by inputs from the vestibular system, path integration, and, when vision is available, from optic flow and visible landmarks. This representation is referred to as the egocentric representation [6–8], spatial image [9,10], or parietal window [11]; I will use 'egocentric memory' in this review. These authors maintain that it is the updating of this model that keeps the perceived world congruent with the external world. I will assume here that these two quite different traditions of inquiry are both valid, and that each has something to contribute to the question of the stability of the phenomenal world.

I will begin by looking at the problem from the standpoint of the requirements of the motor system, since this, from an evolutionary perspective, is where visual function has its origins. Initially I will only consider rotations in the horizontal plane, but it is quite possible to extend the arguments given here to the changes that result from translational movements and I will return briefly to this towards the end. The problems of fully representing coordinate transformations in three-dimensional space are considerable, and are discussed in [12].

Representations for Action

To point in the direction of an object in the surroundings, the visual signal, in the form of the location of the object on the retina, needs to go through a number of transformations (Figure 1). Ultimately the retinotopic location (α) must be converted into the direction of the arm relative to the trunk axis (δ), but for this to happen the eye-in-head angle (β) and the head-on-body angle (γ) have to be added in. In the case shown in Figure 1, the pointing angle δ is simply $\alpha + \beta + \gamma$. This formulation is not exact, because eye, head, trunk and arm do not revolve around a common axis, but it is adequate for present purposes.

The angle β can be obtained either from an efference copy of the signal to the eye muscles or from eye muscle proprioceptors, and similarly the neck angle γ can come from
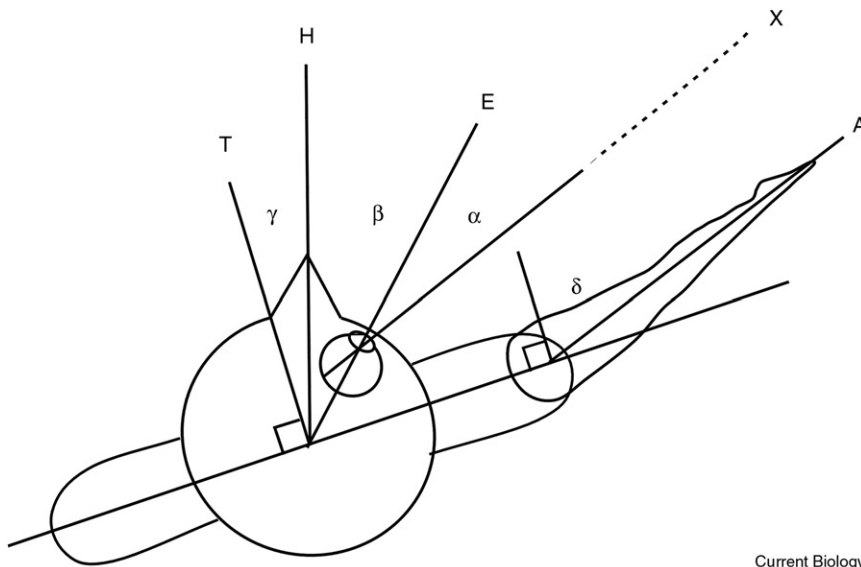
School of Life Sciences, University of Sussex, Brighton, BN1 9QG, UK.
E-mail: m.f.land@sussex.ac.uk

Figure 1. The information required when pointing to a target at X.

efference copy or from proprioception. At some stage in the calculation of δ these additions must be made, and there is general agreement that the posterior parietal cortex is crucial for these transformations. Directional information provided by the parietal cortex is then passed to the premotor and motor cortices for the formulation of limb movements. In their seminal review, Colby and Goldberg [13] presented extensive evidence that the parietal cortex contains multiple representations of space that are linked to spheres of action that involve the eye, head or hand. Attention seems to be crucial in these transformations: parietal cells respond not simply to the presence of a stimulus, but only if the stimulus is relevant to the current task.

What is not entirely clear is the coordinate frame in which these transformations take place, and in which the arm direction is represented. They might take place through a series of transformations representing sequentially each of the additions implied by Figure 1, to arrive finally at the arm angle δ. In fact, much of the evidence from single-unit studies in the parietal and premotor areas indicates that the opposite occurs, with the arm direction being converted into gaze-centred coordinates by an inverse transformation, so that $δ_R$, now in retinotopic coordinates, is equal to $δ - β - γ$ [14–16]. There are some advantages to this apparently paradoxical way of doing things. If $α$ and $δ_R$ are represented in the same frame then they can be compared with each other, and the angle between them — the 'pointing error' — can be obtained directly from the difference. Furthermore, this error will be insensitive to eye movements, because the target image and arm direction will shift together in the retinal image. Also, if the hand is visible, its location can be coded in the same frame. From this, it seems that a retinotopic (gaze-centred) representation can provide an ideal theatre in which the visual control of the actions of the arm and hand can be played out.

Of course to move the arm to a new direction the opposite transformation must be reapplied: a new value of δ must be constructed from its components, as in Figure 1. Presumably both reverse and forward transformations take place at least in part in the regions of the parietal cortex identified with reaching: in monkeys this is the parietal reach region (medial

intraparietal area; MIP). Some recent studies of MIP have suggested that this does not involve a straightforward translation between reference frames, and that neurons can employ spatial encodings with a range of idiosyncratic representations [17]. The premotor and motor cortices are also involved in the formulation of appropriate movements of the limbs, and some information is available about the relative roles of individual cells and cell populations during reaching behaviour [18,19].

How these transformations are brought about is still far from clear. Whilst the representations in most parietal areas have a retinotopic frame, the gain of many parietal neurons with retinotopic fields is modulated by the positions of the eyes, head or limbs, in a way that could provide the basis for coordinate transformations [20,21], (reviewed in [22]). As Crawford *et al.* [16] put it: gain fields can do this "by tweaking the individual contributions of units so that the overall population vector rotates". For example, in this way the addition of an eye position signal can turn the retinotopic location of an object into a craniotopic location.

In what follows, I will assume that some kind of head-centred (craniotopic) mapping of visual space does indeed exist. Clearly a gain-field modified retinotopic representation can perform this function. The reason for choosing to deal with a head-centred (rather than a gaze-centred or body-centred) representation will become apparent later, but has to do with the fact that the vestibular system measures motion of the head.

Vision and Memory in Pointing and Reaching
In this section, I consider the way the visible and invisible worlds map onto the motor system that enables pointing. I assume that there exists in motor areas of the cortex a region that formulates the limb movements required for making pointing or reaching movements relative to the trunk, and that for each pointing direction there are neurons or neuron populations with fixed loci in the brain [23]. In Figure 2, their distribution is represented by the circle M, and it is assumed that they connect directly or indirectly with the muscles of the arm.

Such direction-tuned cells certainly exist in the dorsal pre-motor cortex, where their activity and the tightness of their directional tuning is related to the probability of their engagement in a particular task [18]. In both the premotor and motor cortices, direction is likely to be coded in population responses rather than single units [24]. Thus, although the details of the mapping are likely to be complex, a representation of direction in motor space (M) does exist. In Figure 2A, this mapping receives an input from the (head-centred) visual representation V, and at some point the angle γ is added (Figure 1), possibly in the ventral premotor cortex [25], to take account of the varying angle between head and

Figure 2. Target representations for action.

Representations of the location of a target (●) onto a head-based visual representation (V) and a motor map responsible for pointing or reaching (M). In (A,B) the target is within the field of vision. In (C,D) the target is outside this field in egocentric spatial memory (dashed line E).

A

B

C

D
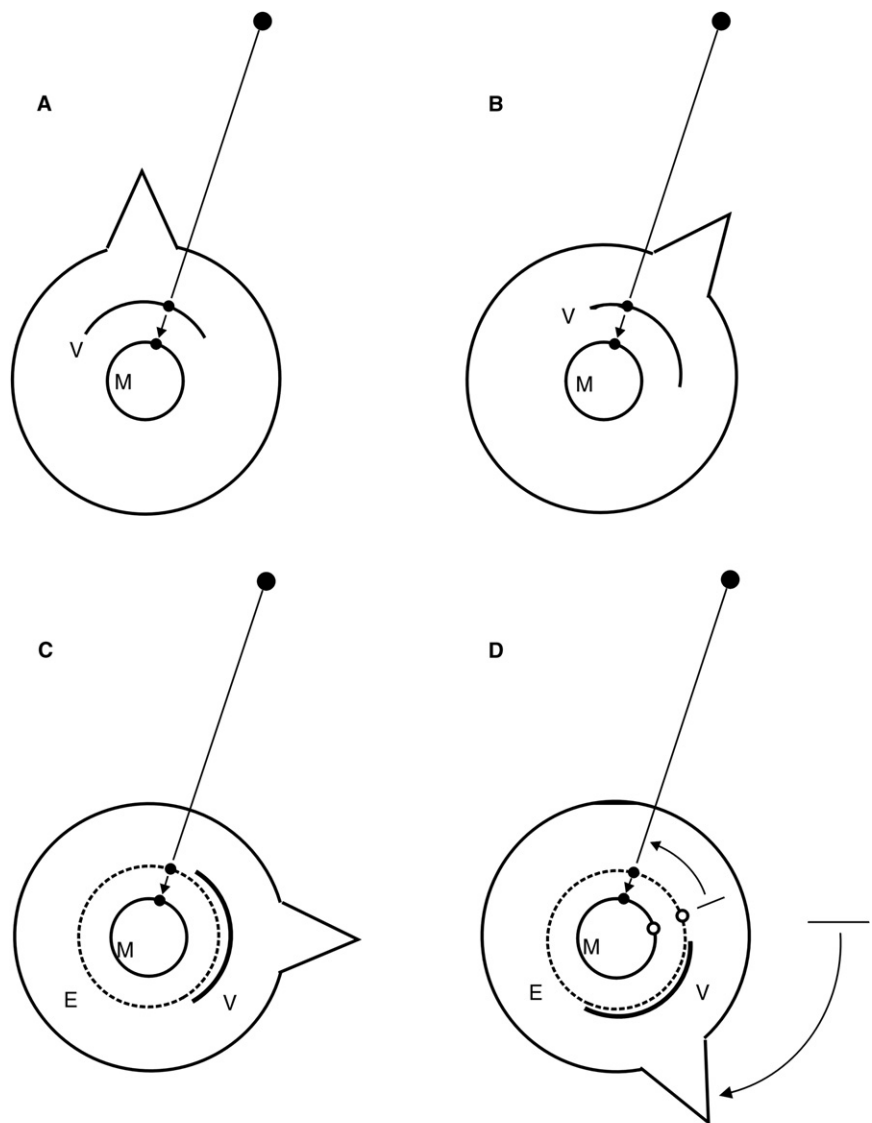
Current Biology

trunk. For the moment I shall ignore this addition and assume that head and trunk axes are in line, and locations on V map directly onto M. Then, if the head rotates 45° right (Figure 2B), the image of the target will move across V, and the mapping onto P will move in line with this to establish a new pointing direction relative to the head (and hence the trunk).

It is also possible, without difficulty, to point to objects that are outside the current field of vision. With eyes closed you can rotate to face the window, and then point to the door or computer screen, and so on for several such rotations with little loss of accuracy. This can all be done on a revolving chair without the feet making contact with the ground. It indicates that there must be a representation of the surroundings that can be independent of immediate visual input, and is updatable by the vestibular system, the only other source of head rotation information in this situation. This egocentric memory representation (E) is shown by the dashed circle in Figure 2C,D.

The visual representation V and the egocentric memory representation E differ in one very important respect. When the head rotates, the image of an object moves across V, but there is no equivalent mechanism for this to occur on E, as there is no image. However, if the location of an object in memory simply remained in the same place relative to the head when the head rotated, then the position signal it sent to the motor area M would be wrong (open circles in Figure 2D). It seems that the locations of object traces within E *must counter-rotate relative to the head as the head rotates*, so that they continue to provide M with the appropriate coordinates. Such a rotation is implicit in the parietal window model of Byrne *et al.* [11], and was explicitly proposed in [26]. How this is achieved is not known, but it might make use of a gain-field mechanism, as in other parietal transformations.

Such a rotational updating mechanism would have two interesting consequences. First, the object memories represented in E would always remain in alignment with the surroundings, a point to which I shall return later. Second, that following a head rotation part of the E representation would move into or out of the visual representation V, so that the visual and unseen worlds would interchange continuously to varying degrees. It is probably sensible to regard the E representation as continuous over 360°, as in Figure 2C,D, with the V representation sliding over it.

Rotational Updating of the Egocentric Memory Representation

In its normal operation the rotating spatial memory E provides coordinates for action, but with eyes closed we do not sense a rotation of the surroundings in the vivid way that we do when the visual image rotates. However, such a rotation can become subjectively apparent when the vestibular signal is disturbed, and vision is not available. If you induce mild dizziness by rotating in an office chair for three or four revolutions, then briefly fixate a convenient object such as a doorknob and close your eyes, you will feel yourself rotating in the opposite direction to the original rotation for 10 seconds or more, during which time you will feel that you have rotated through an angle of perhaps 90°. (It is important to have your feet off the ground during this phase, to remove all except vestibular cues). Without opening your

eyes, point to where you think the doorknob now is: on opening your eyes you will then find that you have pointed 90° away from the actual direction of the knob, whose location has, of course, not changed [27].

The reason for this mislocation is that during the initial rotation the fluid in the semi-circular canals catches up with the lumen, so that when the rotation stops the hair cells attached to the cupolae provide an erroneous signal indicating rotation in the opposite direction. This continues to update the egocentric memory representation until the canal fluid stops rotating. Then, when pointing, you accept the new orientation of the memory representation, and act as you would in accordance with Figure 2C,D. A similar pointing mislocation can be induced by caloric stimulation of the labyrinths [28]. With the eyes open this post-rotation updating does not occur. Vision overrides the vestibular input to E, even though the erroneous vestibular signal is still apparent as a transient vestibulo-ocular after-nystagmus.

These observations can be taken to demonstrate that the egocentric memory representation can indeed be updated by the vestibular system, and that the connection between this representation and the motor system for pointing (Figure 2C,D) continues to function. Normally there will be no conflicts between vestibular and visual updating mechanisms. If the memory system is updated as described in Figure 2D, then the locations of objects represented in it will always coincide with those in the head-centred visual representation. Of interest here is that cells in area MSTd in monkeys have inputs that are both visual and vestibular, and their outputs correlate well with the animals' perceptual judgements of shifts in heading direction [29]. A situation in which conflict between vision and the vestibular system does occur is during vertigo, a feeling that the head is turning in a stationary world that occurs as a result of damage to the semi-circular canals. This results in an imbalance between the outputs of the two sets of canals, and a persisting sensation of turning results, which can occur with the eyes open or shut [30]. Under normal circumstances, the relation of the memory representation to the surroundings can be reset by vision. As noted above, if you keep your eyes open, vision can override the vestibular rotation of egocentric memory. It has also been shown that vision alone can be sufficient to update the egocentric memory representation [31].

The egocentric memory representation is almost certainly both sparse in terms of its content, and not particularly accurate. Pointing experiments under conditions where the target is not visible and the subject not deliberately disoriented tend to come up with accuracy estimates of between 5 and 25° [32–34]. Similarly, large (> 90°) saccadic turns to objects required for a task, involving rotations of eye, head and body, and accompanied by a blink preventing vision, tend to miss the intended target by about 10°, and are followed by a second small saccade once vision becomes available [26]. Attempts to reconstruct even familiar places from immediate or long-term memory typically summon up only small numbers of objects, and very little detail. Thus, the egocentric memory is not a high-acuity representation.

Overlap of Visual and Memory Representations

If the memory representation is panoramic, as suggested earlier, then one might expect that within the visual field itself the information needed for action might come from either vision or memory. There is evidence that this is the case. Brouwer and Knill [35] devised a visually guided reaching task, using virtual reality, in which vision and memory were pitted against each other. Two virtual objects had to be picked up and placed in a trash bin. In some trials the position of the second object was moved by a small amount while the first was being moved to the trash. Participants never noticed this, but it did have a measurable effect on their behaviour. In some cases the arm moved in the direction of the current location of the second object, but in others it went to the object's previous location. Thus, the arm was sometimes guided by vision, and sometimes by memory. It turned out that the relative weighting of vision and memory in the reach to the second target depended on its visibility. For high-contrast objects vision was favoured, whereas low-contrast objects produced more memory guidance. This implies that the brain can choose between sources of information depending on their relative reliability. In these studies the distance in visual space from the objects to the trash can was only about 30°, so the actions all took place well within the peripheral visual field.

In a study which led to much the same conclusion, Aivar et al. [36] used a virtual version of a block-copying task [37] to determine whether saccades to blocks seen in peripheral vision were guided primarily by vision or by memory. Changes were made to the layout of the blocks used to make the copy while the subject was looking elsewhere, but with the blocks still visible in peripheral vision. When gaze returned to the blocks, saccades were normally launched to their old locations rather than to the new, visible, locations. Taken with the Brouwer and Knill [35] result, this confirms that visual and memory representations overlap each other, certainly in peripheral vision, and perhaps throughout the visual field, and both can be used in the formulation of actions, whether these are shifts of gaze or of reaching movements by the arm.

Another conclusion that can be drawn from these studies is that visually attended objects can be temporarily attached to and detached from the egocentric representation. It is always possible that this involves another, special kind of working memory, but there seems no need to assume this.

Allocentric and Egocentric Frames of Reference

If we move from one familiar environment to another, the key objects represented in egocentric memory must change. Students of spatial cognition have long accepted that spatial memory is supported by at least two representations: the egocentric representation, discussed here, which is concerned with the orientation of an individual within their immediate surroundings, and an allocentric representation which is more map-like, and independent of the current orientation of the head and body. The former is certainly based in the parietal lobe, possibly the precuneus [38], and the latter in the hippocampus and medial temporal lobe [7].

The two representations interact. The allocentric representation is fed from information that comes from vision via the succession of egocentric viewpoints, and the egocentric representation can be updated from stored information contained in the allocentric archive. It is, for example, easy to imagine oneself facing in a variety of directions in a room of one's house or place of work. The retrosplenial cortex and parieto-occipital sulcus appear to be the locations where allo-egocentric representations interact [8,11]. In rodents these conversions may involve the 'head direction cells' found in many areas of the brain [39], and the 'grid cells' of the entorhinal cortex which are involved in updating

Figure 3. Relations between the external world, the retinocentric representation provided by the visual cortex, and the egocentric representation.

Between Figure 3A and 3B the head rotates 45° right, and the contents of the egocentric frame rotate 45° left. Between 3A and 3C the eye rotates 20° right relative to the head. Between 3C and 3D the retinocentric frame is rotated 20° right, converting it to a craniocentric frame, and bringing it into line with the egocentric frame. These two different mechanisms keep the retinocentric and egocentric representations in line with each other and the external world. It is assumed here that such an alignment is the precondition for seeing a stable world.
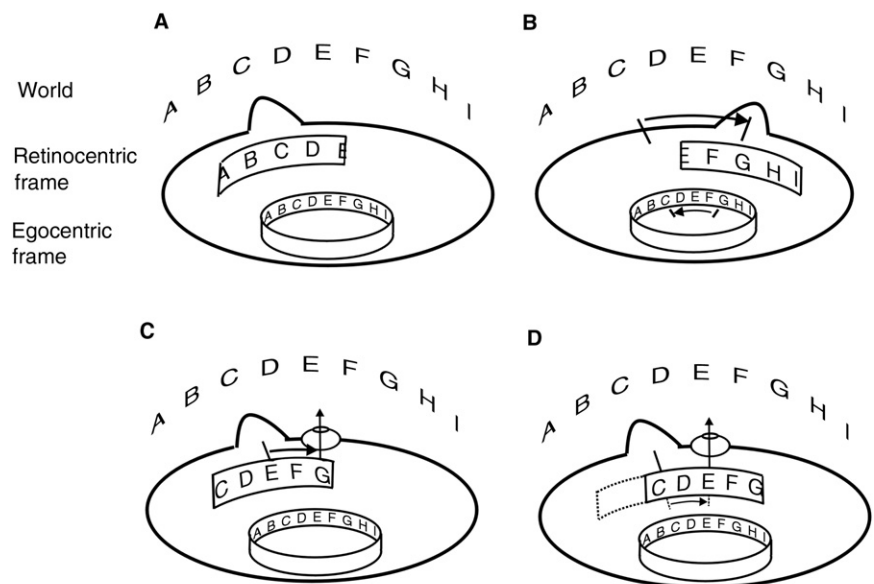
Current Biology

the 'place cells' of the hippocampus during path integration [40]. For present purposes the allocentric representation is only important in that it provides a further source of updating for egocentric memory, in addition to visual and vestibular inputs. This allows the fairly minimal contents of egocentric memory to be changed in an anticipatory way as we move from one part of the environment to another.

The Roles of Eye Movements

As Gordon Walls pointed out: "Their origin (eye movements) lies in the need to keep an image fixed on the retina, not in the need to scan the surroundings" [41]. The saccade and fixate system we have inherited from early vertebrates is really a consequence of the slow response time of photoreceptors (∼20 ms for cones), and the consequent need to stabilize gaze to prevent high-spatial frequency information loss due to motion blur. (The consequences of motion blur can easily be appreciated by observing what happens to the background while tracking a moving finger.) As 'letting the eyes wander' is not an option, the use of saccades to shift gaze is a necessary evil, and their disruptive consequences need to be dealt with. Viewed in this light, the multiple arrangements for minimizing the disruption caused by saccades, which include saccadic suppression, and the neural shifting of image-related excitation from current receptive fields to those that will see the same objects after a saccade [2,13], are perhaps best regarded as the editing processes that remove saccades from what we actually see.

This remapping process may ease the transition from the pre-saccadic to post-saccadic retinal image, but as the receptive field movement is in the same direction as the saccade rather than the opposite (compensatory) direction [42] it is not easy to see how this would help to site the post-saccadic image in relation to external panoramic space. In any event, a purely gaze-centred representation, remapped or not, cannot deal with objects outside the field of view, to which we can still point or orient. Based partly on change blindness studies, some argue that no such panoramic framework exists, or is minimal at best [43]. I argue here that even though its contents are sparse, the egocentric representation is essential for the coordination of a variety of motor tasks.

How Is the Rotatable Egocentric Representation Related to What We See?

Bridgeman [43] has pointed out that we do not have anywhere in the brain "a panoramic high-acuity representation that corresponds to visual experience". What is proposed here is that there exists a low acuity representation which is kept in register with our immediate surroundings, and onto which each high acuity snapshot can be temporarily indexed.

The world we see at any instant has the characteristics of the retinotopic representation provided by immediate vision: colour, motion, detail and so on — all the sub-components of vision provided by the first few stages of the cortical analysis that begins in V1. At the same time it has properties that are much more like the egocentric memory representation discussed above: the absence of saccadic dislocations and the apparent smoothness of movement of gaze when we look around a scene, the evident stability of the seen world and the apparent continuity between currently seen and unseen parts of the surroundings. In addition, there is the requirement for a proper correspondence of the location of seen objects and the motor actions required to deal with them: a correspondence that cannot easily be maintained by a representation that relocates its coordinate system three times a second.

The question then becomes: how is it possible to have the benefits of both types of representation, not only for the guidance of action but for visual experience itself? The scheme in Figure 3 is an attempt to specify the needs of a system that will allow this to happen. It contains two elements, a retinocentric representation that corresponds to the view of the scene represented by the activity in the primary visual cortex and subsequent extra-striate areas, and an egocentric frame that corresponds to the memory representation in Figure 2. This counter-rotates when the head rotates, maintaining a constant relationship with the external world. Figure 3A shows the situation when the eye axes are in line with the head. The retinocentric representation centred on C is in line with the egocentric representation. In Figure 3B the

head has rotated 45° to the right, shifting the retinocentric representation by the same amount, so that it is now centred on G. If the egocentric representation remained in the same orientation relative to the head, G in the retinocentric representation would remain aligned with C in the egocentric frame. We must therefore assume that the locations of object traces in the egocentric representation have rotated 45° to the left relative to the head, in part as a result of vestibular input, meaning that the retinocentric and egocentric representations remain aligned, and there is no mismatch.

The situation when the eyes move relative to the head is different (Figure 3C). Here the direction of gaze has been shifted by an eye movement so that, relative to Figure 3A, the retinocentric representation is centred on E, 20° to the right of C, and is no longer in line with the head axis or the egocentric frame. This mismatch cannot be dealt with by the vestibular system rotating the egocentric frame because there is no head rotation, and in any case the orientation of the egocentric representation needs to remain in line with the external world so that actions can be properly organized. Because it is observationally true that eye movements have a minimal disruptive effect on our ability to point, some other mechanism is required to bring the retinocentric frame in line with the head-based arrangement in Figure 3A, which would otherwise still be centred on C.

The extra information to do this is certainly available, in the form of cortical eye-position signals [44]. Such signals are found in a number of cortical areas, including the intra-parietal areas LIP and VIP, the middle temporal area MT and medial-superior temporal area MST. If such signals can be used to re-position the retinocentric frame with respect to the egocentric frame, then again the two representations can be brought into line, as in Figure 3D. A head-centred representation of image space does exist in the ventral intra-parietal area (VIP), in which the receptive fields of the neurons are modified by eye-position signals so that they remain aligned with points in space as the eyes move [45]. There is also evidence from functional magnetic resonance imaging (fMRI) studies that in the motion area MT attended stimuli are mapped in a spatial rather than a retinotopic frame [42,46]. Although it is good to know that such retinotopic to craniotopic conversions do take place, it is not clear whether either VIP or MT is actually the site of the repositioning implied by Figure 3D.

As mentioned earlier, spatial updating is not limited to rotation. There is ample evidence, reviewed in [8,9], that translational movement through the world, with the eyes closed, can be accomplished as effectively as rotation. Memory traces of small numbers of objects can be internally relocated during motion, and their positions can be used to control locomotion, be pointed to, or stopped in front of. Such traces are not limited to those initially stored visually, but can be auditory or haptic in origin, or derived from spatial instructions. These mental images differ from 'pictorial images', which tend to be detailed and located in forward space, in that they are sparse but panoramic, and reflect the parallax changes between objects that occur during locomotion, as would have been perceived with the eyes open. In terms of the formalism for the egocentric frame shown in Figure 3, the ground plane can be represented by radial coordinates in the ring that currently represents only directions. Motion through the world will then change both the directions and distances of objects represented within the mapping. This representation is essentially the same as

that proposed for the 'parietal window', as in Figure 3 of Byrne et al. [11].

Returning to the question of what constitutes the phenomenal visual world, the proposal here is that it is indeed a combination of the series of vivid images provided by the retinocentric frame, and the spatial continuity provided by the egocentric frame. For this combination to work, the two representations must stay at least roughly in register, and for this the two mechanisms outlined in Figure 3 are required. The first rotates the contents of the egocentric frame, so that it maintains congruity with the external world (Figure 3B), and the second compensates the retinocentric frame for eye movements, so that it effectively becomes craniocentric (Figure 3D). What we see is the view presented to our eyes, aligned with the appropriate point in an internal model that represents outside space.

References

1. Land, M.F., and Tatler, B.W. (2009). Looking and Acting (Oxford: Oxford University Press).
2. Duhamel, J.R., Colby, C.L., and Goldberg, M.E. (1992). The updating of the representation of visual space in parietal space by intended eye movements. Science 255, 90–92.
3. Merriam, E.P., Genovese, C.R., and Colby, C.L. (2003). Spatial updating in human parietal cortex. Neuron 39, 361–373.
4. Hall, N.I., and Colby, C.L. (2011). Remapping for visual stability. Phil. Trans. R. Soc. Lond. B 366, 528–539.
5. Wurtz, R.H., Joiner, W.M., and Berman, R.A. (2011). Neuronal mechanisms for visual stability: progress and problems. Phil. Trans. R. Soc. Lond. B 366, 492–503.
6. Farrell, M.J., and Robertson, I.H. (1998). Mental rotation and the automatic updating of body-centred spatial relationships. J. Exp. Psychol. Learn. Mem. Cogn. 24, 227–233.
7. Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. Trends Cogn. Sci. 10, 551–557.
8. Burgess, N. (2008). Spatial cognition and the brain. Ann. N.Y. Acad. Sci. 1124, 77–97.
9. Loomis, J.M., and Philbeck, J.W. (2008). Measuring spatial perception with spatial updating and action. In Embodiment, Ego-space and Action, R.L. Klatzky, B. Mac Whinney, and M. Behrman, eds. (New York: Taylor & Francis), pp. 1–43.
10. Loomis, J.M., Klatzky, R.L., and Giudice, N.A. (2012). Representing 3D space in working memory: spatial images from vision, hearing, touch and language. In Multisensory Imagery, S. Lacey and R. Lawson, eds. (New York: Springer), in press.
11. Byrne, P., Becker, S., and Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psych. Rev. 114, 340–375.
12. Crawford, J.D., Henriques, D.Y.P., and Medendorp, W.P. (2011). Three-dimensional transformations for goal-directed action. Annu. Rev. Neurosci. 34, 309–331.
13. Colby, C.L., and Golberg, M.E. (1999). Space and attention in parietal cortex. Annu. Rev. Neurosci. 22, 319–349.
14. Batista, A.P., Buneo, C.A., Snyder, L.H., and Anderson, R.A. (1999). Reach plans in eye-centred coordinates. Science 285, 256–260.
15. Buneo, C.A., Jarvis, M.R., Batista, A.P., and Anderson, R.A. (2002). Direct visuomotor transformations for reaching. Nature 416, 632–636.
16. Crawford, J.D., Medendorp, W.P., and Marotta, J.J. (2004). Spatial transformations for eye-hand coordination. J. Neurophysiol. 92, 10–19.
17. Chang, S., W.C., and Snyder, L.H. (2010). Idiosycratic and systematic aspects of spatial representations in the macaque parietal cortex. Proc. Nat. Acad. Sci. USA 107, 7951–7956.
18. Cisek, P., and Kalaska, J.F. (2005). Neural correlates of reaching decisions in the dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron 45, 801–814.
19. Pesaran, B., Nelson, M.J., and Andersen, R.A. (2008). Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reaching. Neuron 51, 125–134.
20. Zipser, D., and Andersen, R.A. (1998). A back-propagation programmed network that simulates response properties of a subset or posterior parietal neurons. Nature 331, 679–684.

21. Andersen, R.A., Snyder, L.H., Li, C.-S., and Stricanne, B. (1993). Coordinate transformations in the representation of spatial information. Curr. Opin. Neurobiol. *3*, 171–176.

22. Hamker, F.H., Zirnsak, M., Ziesche, A., and Lappe, M. (2011). Computational models of spatial updating in peri-saccadic vision. Phil. Trans. R. Soc. Lond. B *366*, 554–571.

23. Georgopoulos, A.P., Ashe, J., Smyrnis, N., and Taira, M. (1992). The motor cortex and the coding of force. Science *256*, 1692–1695.

24. Scott, S.H., and Kalaska, J.F. (1997). Reaching movements with similar hand paths but different arm orientations. I. Activity of individual cells in motor cortex. J. Neurophysiol. *77*, 826–852.

25. Graziano, M.S.A., Yap, G.S., and Gross, C.G. (1994). Coding of visual space by premotor neurons. Science *266*, 1054–1057.

26. Tatler, B.W., and Land, M.F. (2011). Vision and the representation of the surroundings in spatial memory. Phil. Trans. R. Soc. Lond. B *366*, 596–610.

27. Land, M.F. (2010). Fairground rides and spatial updating. Perception *39*, 1675–1677.

28. Bottini, G., Sterzi, R., Paulesu, E., Vallar, G., Cappa, S.F., Erminio, F., Passingham, R.E., Frith, C.D., and Frackowiak, R.J.J. (1994). Identification of the central vestibular projections in man; a positron emission tomography activation study. Exp. Brain Res. *99*, 164–169.

29. Gu, Y., Angelaki, D.E., and DeAngelis, G.C. (2008). Neural correlates of multi-sensory cue integration in macaque MSTd. Nat. Neurosci. *11*, 1201–1210.

30. Barton, J.J.S. (2000). Vertigo and vestibular function. Rev. Ophthalmol. *7*, 56–64.

31. Rieke, B.E., Cunningham, D.W., and Bülthoff, H.H. (2007). Spatial updating in virtual reality: the sufficiency of visual information. Psych. Res. *71*, 298–323.

32. Farrell, M.J., and Robertson, I.H. (1998). Mental rotation and the automatic updating of body-centred spatial relationships. J. Exp. Psychol. Learn. Mem. Cogn. *24*, 227–233.

33. Mou, W., McNamara, T.P., Rump, B., and Xiao, C. (2006). Roles of egocentric and allocentric spatial representations in locomotion and reorientation. J. Exp. Psychol. Learn. Mem. Cogn. *32*, 1274–1290.

34. Waller, D., and Hodgson, E. (2006). Transient and enduring spatial representations under distortion and self-rotation. J. Exp. Psychol. Learn. Mem. Cogn. *32*, 867–882.

35. Brouwer, A., and Knill, D. (2007). The role of memory in visually guided reaching. J. Vis. *7*, 1–12.

36. Aivar, M.P., Hayhoe, M.M., Chizk, C.L., and Mruzek, R.E.B. (2005). Spatial memory and saccadic targeting in a natural task. J. Vis. *5*, 177–193.

37. Ballard, D.H., Hayhoe, M.M., Li, F., Whitehead, S.D., Frisby, J.P., Taylor, J.G., and Fisher, R.B. (1992). Hand eye coordination during sequential tasks. Phil. Trans. R. Soc. Lond. B *337*, 331–339.

38. Wolbers, T., Hegarty, M., Büchel, C., and Loomis, J. (2008). Spatial updating: how the brain keeps track of changing object locations during observer motion. Nat. Neurosci. *11*, 1223–1230.

39. Taube, J.J., Muller, R.U., and Ranck, J.B., Jr. (1990). Head direction cells recorded from the post-subiculum in freely moving rats. I. Description and quantitative analysis. J. Neurosci. *10*, 420–435.

40. Moser, E.I., Kropff, E., and Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. Annu. Rev. Neurosci. *31*, 69–89.

41. Walls, G.L. (1962). The evolutionary history of eye movements. Vis. Res. *2*, 69–80.

42. Burr, D.C., and Morrone, M.C. (2011). Spatiotopic coding and remapping in humans. Phil. Trans. R. Soc. Lond. B *366*, 504–515.

43. Bridgeman, B. (2011). Visual stability. In The Oxford Handbook of Eye Movements. S.P. Liversedge, I.D. Gilchrist, and S. Everling, eds. (Oxford: Oxford University Press), pp. 511–521.

44. Morris, A.P., Kubischick, M., Hoffmann, K.-P., Krekelberg, B., and Bremmer, F. (2012). Dynamics of eye-position signals in the dorsal visual system. Curr. Biol. *22*, 173–179.

45. Duhamel, J.R., Bremmer, F., BenHamed, S., and Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex. Nature *389*, 845–848.

46. d'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D.C., and Moroni, M.C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. Nat. Neurosci. *10*, 249–255.