

# Depth perception by the active observer

Mark Wexler<sup>1</sup> and Jeroen J.A. van Boxtel<sup>2</sup>

<sup>1</sup>CNRS, 11, Pl. Marcelin Berthelot, 75005 Paris, France

<sup>2</sup>Helmholtz Institute, University of Utrecht, The Netherlands

**The connection between perception and action has classically been studied in one direction only: the effect of perception on subsequent action. Although our actions can modify our perceptions externally, by modifying the world or our view of it, it has recently become clear that even without this external feedback the preparation and execution of a variety of motor actions can have an effect on three-dimensional perceptual processes. Here, we review the ways in which an observer's motor actions – locomotion, head and eye movements, and object manipulation – affect his or her perception and representation of three-dimensional objects and space. Allowing observers to act can drastically change the way they perceive the third dimension, as well as how scientists view depth perception.**

## Introduction

Although perception has traditionally been considered only as processing of sensory data, it is really a component of the action–perception cycle. Perceptions inform actions that modify either the world (object manipulation) or our view of it (head and eye movements, locomotion), in turn modifying subsequent perceptions. The boundary between perception and action frequently fades: many actions are undertaken for their perceptual consequences, and perception is often tuned to those aspects of the world that are available for the observer to act on.

What is true of perception in general is especially true of the visual perception of depth. Although sources of 3D information are available on our 2D retinas, we can obtain much richer knowledge of the third dimension by coordinating the gaze directions of the two eyes, by moving the head to produce parallax, by walking to get a different view of a scene, by manipulating an object to better see its shape. To philosophers such as Berkeley and Condillac, perceptions and representations of 3D space could only originate with motor action.

It is evident that actions can modify perceptions *externally* by modifying the world or one's view of it. However, it has recently become clear that motor action and depth vision are *internally* linked as well: executing or preparing a motor action – independent of its sensory

consequences – is often enough to modify the observer's perception and representation of 3D space and shape. In this article we will review the ways in which locomotion, head and eye movements, and object manipulation modify visual perceptions of depth.

## Locomotion and navigation

### *Active exploration and formation of spatial maps*

Perception and representation of 3D space go hand-in-hand with our ability to move around in the environment. Developmental psychologists have shown that a variety of perceptual and conceptual skills related to 3D space undergo dramatic improvement when infants learn to crawl on all fours, usually around the age of 7 months (for a recent review see [1]). For instance, just after learning to crawl infants develop stronger postural responses to large-field OPTIC FLOW (see Glossary), more precise perception of 3D structure from a variety of depth cues including MOTION PARALLAX, and a fear of heights. (These results seem to echo the importance of active locomotion in the development of depth perception in kittens [2].) Another skill that emerges with self-produced locomotion is the formation of ALLOCENTRIC (or observer-independent) spatial maps [3], as illustrated in Figures 1a and b. Pre-locomotor infants, after being shown an object to, say, the left of their midline and then rotated by 180°, tend to incorrectly continue searching for the object to

## Glossary

**Allocentric reference frame:** an observer-independent (possibly earth-fixed) frame of reference.

**Efference copy:** a copy of the neural motor command that, instead of being sent to the muscles, is used for further processing in the brain, such as suppressing reafference (closely related to the notion of 'corollary discharge').

**Extraretinal signals:** sources of information used in vision that do not originate from optical stimulation of the retina.

**Haptic perception:** the combination of tactile perception through the skin and kinaesthetic perception of the position and movement of the joints and muscles.

**Motion parallax:** optic flow in which the motion at each point depends on the corresponding object's distance from the eye.

**Optic flow:** the pattern of movement on the retina caused by relative motion between the observer and the visual scene, caused either by object motion, subject motion, or both.

**Proprioceptive information:** sensory information about the current state of the body's posture and motion, arising from signals from the vestibular organ, muscle spindles, for example.

**Reafference:** afferent sensory information that is systematically dependent on actions initiated by the sensing organism; contrasted to *exafference* which refers to stimulation independent of self-produced movement.

Corresponding author: Wexler, M. (wexler@ccr.jussieu.fr).

their left. Age-matched infants that have learned to crawl, on the other hand, tend to correctly search for the object to their right: they compensate representations of object location by self-motion. Studying infants who learn to crawl either earlier or later than the norm has shown that advances in depth perception and spatial representation are likely to be *caused* by the sorts of action–perception relationships engendered by active locomotion [1].

The tight link between locomotion and the updating of spatial maps persists during adulthood. If we are shown an object and then blindfolded and led to another location, we can quickly and accurately point to the object, without opening the eyes, from our new position and orientation. If we merely imagine walking the path [5], or view the corresponding optic flow [6], we can do this task only with difficulty and with large errors. Therefore, *EXTRARETINAL SIGNALS* produced during locomotion are necessary to induce accurate updating of egocentric spatial maps, but optic flow alone is insufficient to do so. It is not only the directions of objects that are updated by extraretinal signals, but also their 3D orientations with respect to the viewer, as shown by better object recognition after self-motion than after object motion [7].

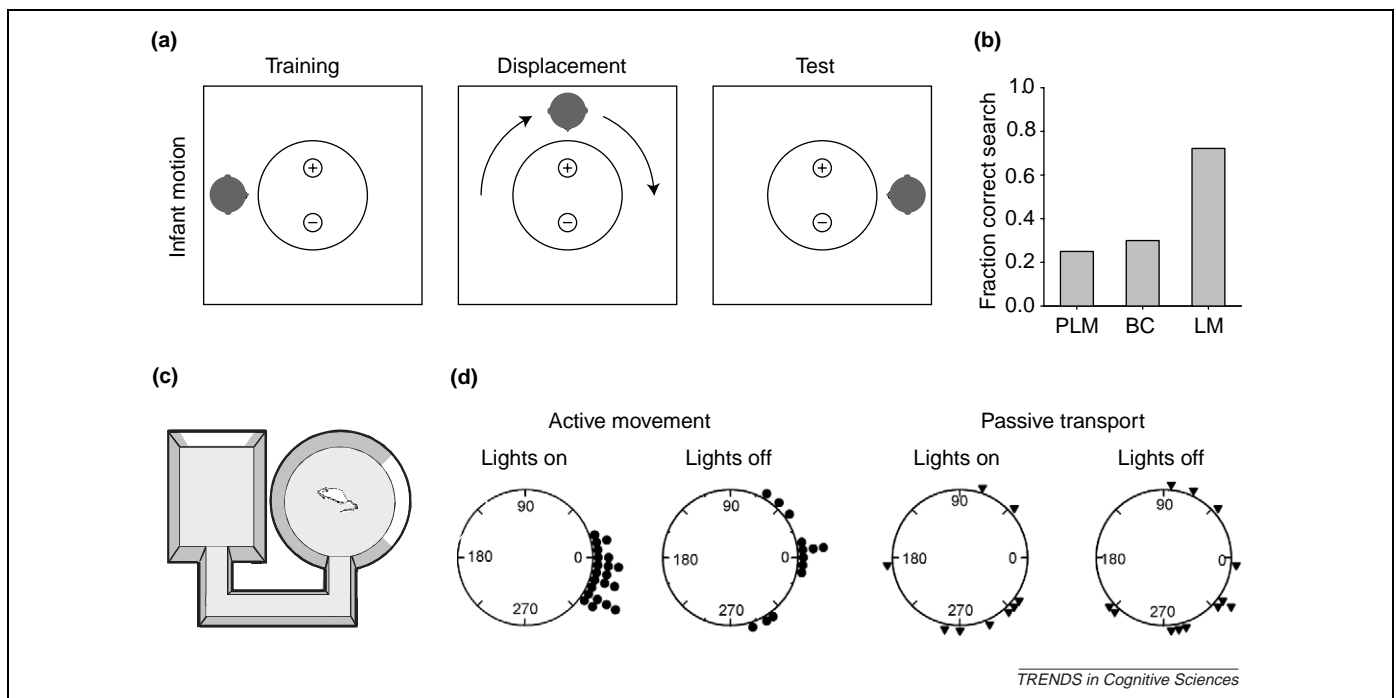
#### Comparing active and passive motion

Possibly underlying our capacity to represent the 3D environment independently of our own position or movement, place cells and head-direction neurons (found in humans, monkeys and rats, in the hippocampus and

thalamus) code the position and orientation of the organism with respect to an allocentric, environment-based *REFERENCE FRAME*. Interestingly, the computations to determine allocentric position and orientation do not seem to be limited to sensory data (visual, vestibular), but also depend on the animal's locomotor action. By comparing actively moving and passively displaced animals, it has recently been shown that the activity of place cells in monkeys [8] and rats [9,10] and of head-direction cells in rats [4,11] depends on the animal's active locomotion, with weaker or disrupted responses during passive displacement (see Figure 1c and d).

Recently, an electrophysiological study of humans actively navigating in a virtual environment (by pressing buttons) has found place cells in hippocampus [12]. However, it is not clear whether the place-cell activity is due to optic flow (as the authors claim) or to subjects' active control of navigation. The latter cannot be excluded, especially because behavioral studies have shown that allocentric spatial coding is more efficient in humans actively controlling their virtual displacement than when merely observing the resulting optic flow [13,14].

To navigate correctly, one has to know where one is with respect to the environment, which is difficult to compute from sensory information that – at best – gives us partial information about the environment with respect to one's sensory organs. Taken together, the above results seem to show that active exploration is necessary to induce accurate updating of spatial maps.



**Figure 1.** (a) Experimental condition in [3] on spatial map formation in infants. An object is hidden (in plain view) in one of two symmetrically arranged positions, say to the infant's left. The infant is then rotated to the opposite side of the table, and searches for the object. All infants tested were ~7.5 months old, but fell into three groups with respect to their locomotion: pre-locomotor (PLM), belly crawling (BC), and crawling on all fours (LM). (b) PLM and BC infants continue searching in the same egocentric direction (i.e. to their left), thereby making a mistake, whereas LM infants correctly update their spatial representation to take into account their movement. (c) The testing set-up in [4]. The preferred directions of rats' head-direction cells were first determined as the animals explored the round chamber. The animals then either actively walked through the passageway to the rectangular chamber, or were passively transported on a cart, with the lights either on or off. (d) The shift in head-direction cells' preferred directions after moving to the rectangular chamber in all four conditions. The cells are strongly disrupted during passive transport (lack of efference copy; see Box 1) and weakly disrupted in darkness (lack of optic flow).

### Perception of heading

An important process in navigation (and therefore for the perception of the 3D environment) is judging one's heading, and one way to do this is through the analysis of optic flow [15]. This analysis is easy to carry out if the gaze is fixed [16], but the gaze is often not fixed, for example when we fixate stationary objects lying off to the side of our path, causing us to turn the eyes and possibly the head. When subjects judge heading from optic flow that includes simulated eye rotation while keeping their eyes still, they often make mistakes. On the other hand, if subjects experience the same optic flow while performing the corresponding eye movement, errors are much smaller [17]. This implies that extraretinal signals participate in the seemingly automatic compensation of optic flow for eye rotations, a compensation that is difficult to achieve without these signals. Similar extra-retinal compensation is found for head [18] and whole-body [19] movements.

Locomotion is most often a goal-directed action: we usually know where we want to go before starting to move. The analysis of optic flow is therefore unlikely to be the only way, or even the main way, by which we judge our heading. The literature on the perception of heading is focused – we feel wrongly so – on how the brain can correct aberrations in optic flow resulting from self-motion. We feel that it would be more fruitful to ask how body, head, and eye movements combine with visual perception in getting us where we want to go [20].

### Affordances and perception of the environment

In addition to actual locomotion, an observer's *preparedness* for locomotion can have an effect on 3D perception. The perceived slant of hills or distance to landmarks has recently been shown to depend on a combination of retinal cues and the potential effort required to walk to the landmark. For example, observers who are wearing a heavy backpack tend to estimate landmarks as being farther away [22] and hills as being steeper [21]. These influences of potential motor effort on depth perception perhaps explain why objects are perceived as farther away when a hole in the ground separates them from the observer [23]. These results suggest that even in the immobile observer, some functions thought to be purely visual actually rely on simulation of motor action.

### Head and eye movement

#### *Self-motion and perception of 3D shape*

When an observer moves his or her head, the resulting displacement of the eyes in space generates 'motion parallax', which is optic flow in which the motion of each point in the retinal image depends on the corresponding object's distance from the eye. This 2D optic flow can be used to reconstruct the 3D shape of objects or scenes, without any additional cues to depth [24], a process called structure-from-motion (SfM). SfM also arises from object motion, with the observer still [25] – as shown in the supplementary video (see Supplementary data online).

Traditionally, it was thought that observer motion and object motion lead to the same interpretation of 3D shape, as long as the optic flow at the retina (i.e. the *relative*

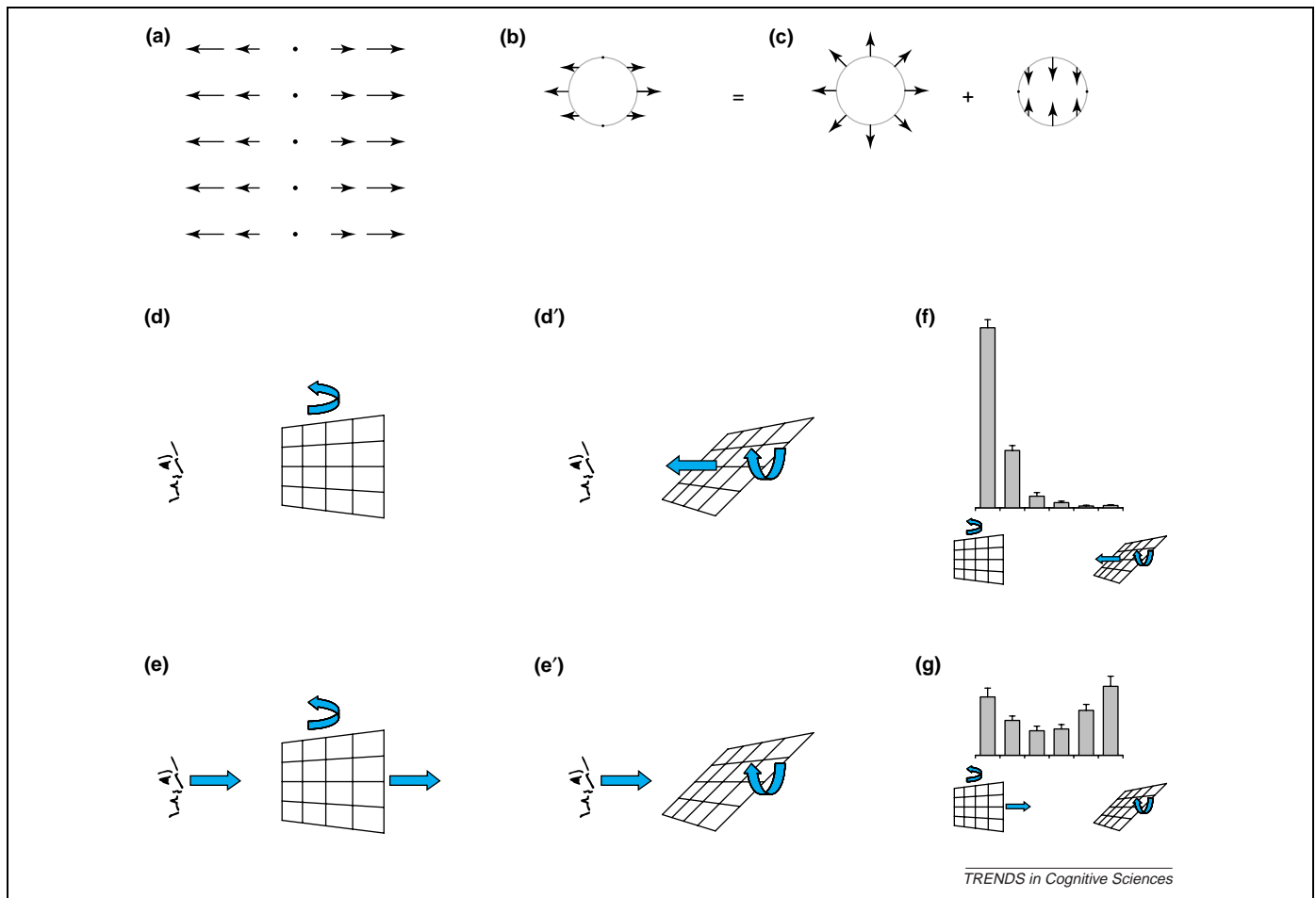
motion between observer and object) is the same [24,26,27]. It turns out, however, that this equivalence is false: the same optic flow can lead to very different perceptions of 3D shape when generated by the observer's own movement than when generated by object motion [28–31]. One illustration of this is shown in Figure 2. The optic flow in Figure 2a is ambiguous, and may be interpreted as different combinations of 3D structure and motion; two of these combinations are shown in Figures 2d and d'. The two solutions correspond to planes slanted in different directions in depth, undergoing depth rotations about different axes; additionally, one of the solutions moves towards the observer. The same ambiguity holds when the observer moves towards the object, instead of the other way around, as shown in Figures 2e and e'.

However, actively moving observers perceive different 3D structure than when they are stationary, despite receiving the same optic flow. When asked to report the orientation of the axis of rotation – which distinguishes solutions in Figures 2d and e from those in d' and e' – observers report predominantly solution d when they are immobile, whereas they report a mixture of solutions e and e' when they are moving (see Figures 2f and g) [29,31]. When the observer's head is moved passively, a result intermediate between active and immobile conditions is obtained [31], showing that both motor and PROPRIOCEPTIVE extraretinal signals can modify the perception of 3D shape.

Results such as the one shown in Figure 2 [29,31] and others [28,32] show that one way in which the visual system uses information from head motion is to select the perception most compatible with an 'assumption of stationarity', that is, to find the perceptual solution that minimizes object motion in an allocentric, observer-independent and earth-fixed reference frame. Thus, the visual system makes the reasonable assumption that most of the world stays fixed when head movement causes the observer to experience motion parallax – a case of stability being used as a criterion for further visual processing (see [33] for a similar suggestion regarding eye movements). The stationarity assumption can explain other psychophysical results in 3D vision, even ones not involving head movement [34].

However, the contribution of the observer's action to SfM is not limited to the stationarity assumption. When optic flow contains a shear component, the perception of 3D shape by immobile observers degrades rapidly as shear increases [35]. In observers performing active head movements, on the other hand, this performance degradation is much less drastic [30].

An active observer can combine information from head movement together with optic flow to judge absolute distance to stationary objects. Several species of insects, rodents and perhaps birds generate motion parallax by making backwards-and-forwards or side-to-side head movements (for a recent review see [36]). Human subjects can also judge absolute distance, to some extent, from motion parallax [37–39], and one-eyed subjects make such movements spontaneously when absolute distance information is required for accurate reaching [40].



**Figure 2.** Illustration of the effect of head motion on the perception of 3D structure [29,31]. (a) An ambiguous 2D optic flow field that can have different 3D interpretations, discovered by J. Droulez. The arrows represent motion of projections of points in 3D space on the retina. It is fairly easy to see that the 3D configuration shown in (d) will generate this flow. However, the configuration shown in (d') can also generate the flow in (a), and the reason for this is shown in (b) and (c): if the amplitudes of the translation and rotation in (d') are adjusted correctly, the rotation can exactly cancel the expansion flow from the depth translation in one of two dimensions (c). The planes in (d) and (d') have the same slant and angular speed, but different tilts and rotate about different axes. (e,e') As optic flow depends only on the relative motion between object and observer, the same ambiguity holds for an observer moving forward and experiencing the optic flow in (a). If the observer's speed is equal-and-opposite to the translation in (d'), the stationarity of the solutions is reversed with respect to (d) and (d'): it is now the center of (e') that is stationary in space, whereas (e) translates at the same speed as the observer. (f,g) Data re-plotted from [31], showing the frequencies of the perceived solutions for stationary (f) and moving (g) observers, with bars to the left corresponding to solutions (d) and (e), and the bars on the right to solutions (d') and (e'). Although optic flow is the same in the two cases, perceptions of 3D structure are very different, showing the effect of the observer's action.

### Eye movements and 3D shape perception

Probably the most important role of eye movements (meaning rotations of the eye in the orbit) in the perception of 3D shape is to converge the two eyes on objects of interest, allowing binocular disparity mechanisms to reconstruct 3D shape from retinal disparities. However, extra-retinal signals arising from eye-muscle proprioception and *REFERENCE COPY* of eye-movement commands (see *Box 1*) give rise to the perception of absolute distance from accommodation [41] and vergence [42]. Extra-retinal signals also calibrate and modify binocular disparity [43].

As opposed to head movements, eye movements give rise to mainly wholesale shifts in the visual information impinging on the retina, and therefore do not generate 3D information (but see [44]). However, eye movements do lift some visual ambiguities. Stimuli may be generated that, when pursued, cause ambiguous optic flow on the retina (e.g. causing an optic flow illustrated in *Figure 1a*). Nevertheless, it has been found that these stimuli are *not* ambiguous to subjects performing the eye movement [45].

Subjects systematically report the interpretation that is most stationary in an allocentric reference frame, meaning that extraretinal information about the eye movement affects the perception of 3D shape. Reflex eye movements that stabilize gaze on moving objects may also account for the disambiguation of 3D shape [46].

Spatial constancy – how visual directions of stationary objects appear constant, despite their frequent and violent rotations on the retina – is a classical problem in neuroscience. Recently, it has been pointed out that the problem of spatial constancy has a neglected 3D component as well: when the eye rotates in space, the 3D orientations of all surfaces undergo an equal-and-opposite depth rotation, in the reference frame of the eye [47]. How we keep from seeing these rotations has been called the problem of 3D spatial constancy. For saccades, the visual system solves this problem at least partly by anticipating the 3D consequences of the eye movements, as demonstrated by a specific bias in the perception of ambiguous depth rotations during the preparation but before the start of saccades [47].

### Box 1. Action-related signals and their combination with retinal data

There are two main sources of extra-retinal action-related signals in the brain. One consists of a copy of the motor command, known as efference copy or corollary discharge, and exists only in the case of actively generated motion. The other consists of reafferent feedback signals from the vestibular organ, and somatosensory or proprioceptive signals from the muscles. These neural extra-retinal signals related to action are combined with visual signals in multiple brain areas.

Extra-retinal information itself is often already pre-processed before it is incorporated with the visual information [67]. Action is reported to affect a variety of brain areas that also process visual information. However, many of these studies concern spatial constancy during eye movements, which generally do not generate 3D information (but see [34,45–47]). Head and body movements do generate 3D information and have been shown to influence non-motor representations in several brain areas, such as the hippocampus, MST, and in posterior parietal areas such as VIP, LIP, and area 7a [9,68,69]. In these areas representations of space are present in different frames of reference (e.g. eye-, head-, and body-centered) [68], and these reference frames

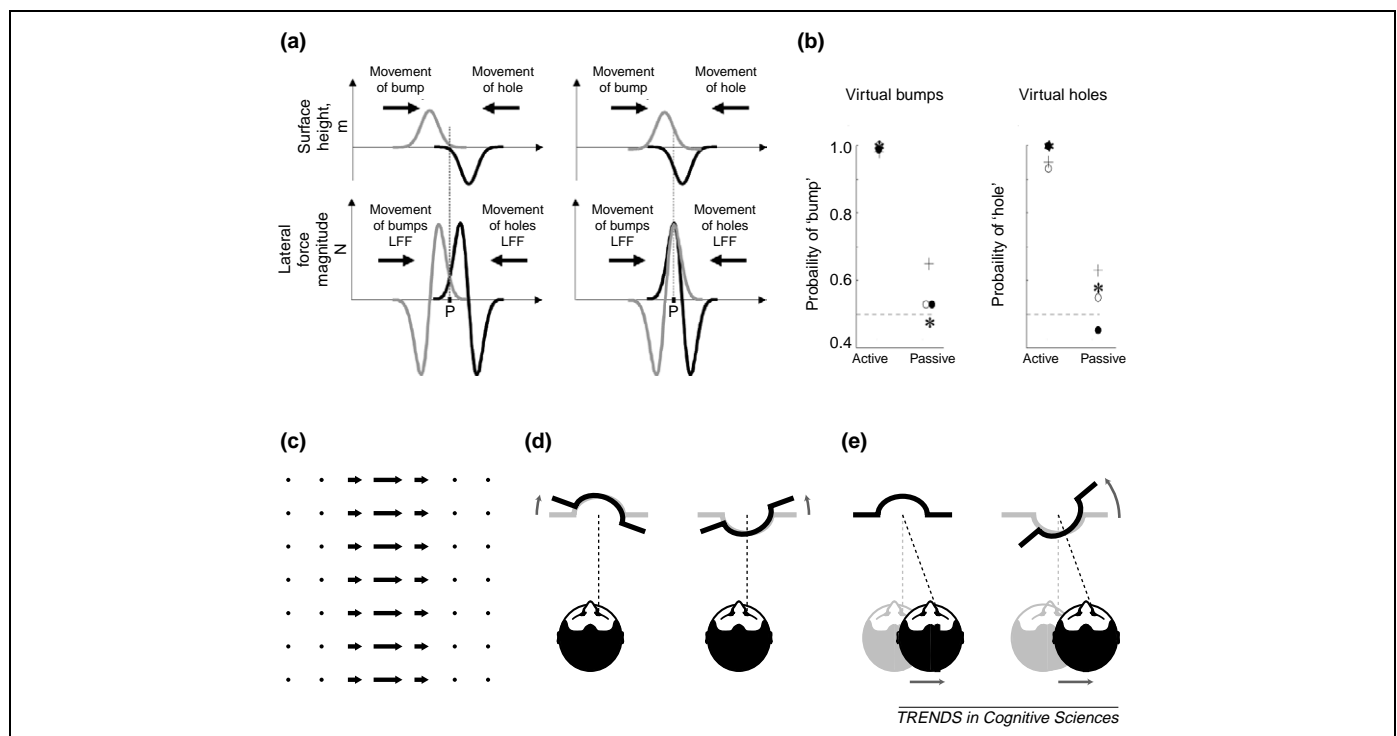
may be used for analyzing and planning actions. It is interesting in this respect that some areas (VIP [69], hippocampus [9]) are found to be differentially activated during active and passive movements, and likewise the anterior superior temporal polysensory area (STPa) responds differently to seen object motion during passive movement and immobile situations [70]. Many of these areas are situated in the dorsal 'where' (action) stream, which fits their assumed role in near extra-personal space representation and navigation [71]. MT/MST is also thought to be important for representing heading during navigation [72], which would explain why it is influenced by extra-retinal information. However MT/MST processes complex motions [72] which are also important for structure-from-motion perception. Accordingly, activity in MT/MST is modulated by 3D shape [73] although it is not considered part of the ventral 'what' stream. As should be expected, shape perception activates ventral areas as well [74,75], and interestingly, haptic information – as a source of actively produced extra-retinal information – is known to influence ventral brain areas [75].

### Object manipulation

#### Perception of 3D shape in active and passive touch

The shape of objects can be HAPTICALLY perceived through either active exploration or by passive sensations of objects that brush against the skin. Comparisons of active and passive touch have a long history, and have often shown superior shape recognition in active touch [48] (but see [49]). Recently, it has been shown that when exploring an object through touch, lateral forces (i.e. forces parallel to the object surface) play an important role in the HAPTIC PERCEPTION of shape [50].

During passive touch lateral forces are ambiguous: a bump moving to the left generates the same lateral forces as a hole moving to the right, and therefore the convex/concave dimension is ambiguous [51], as shown in Figure 3a. In active touch, on the other hand, as the direction of motion is known, at least theoretically the ambiguity can be lifted. Interestingly, the brain does make use of the motion information in active touch to resolve the depth ambiguity [51], as shown in Figure 3b. In fact, this is formally the same ambiguity that faces the passive observer trying to extract 3D structure from



**Figure 3.** (a) The ambiguity of lateral force fields (LFFs) during passive touch. A bump moving under a finger to the right produces the same LFF as a hole moving to the left [50]. (b) When a surface is explored actively, this ambiguity is lifted by the known direction of hand motion. In the case of passive touch, the ambiguity is complete, as it should be (dashed line shows chance level; symbols refer to four different subjects). (c) An ambiguous optic flow field, which is a close analogy from 3D vision to the preceding haptic effect. It is generated by all four of the subject-object movement configurations shown in (d) and (e). (An animated version of this optic flow can be seen in the video in supplementary data.) (d) An immobile observer can interpret this optic flow either as a hole rotating in depth, or as a bump undergoing an equal-and-opposite rotation. (e) The same two solutions are present when optic flow is generated by the observer's head movement. However, observers perceive the solution that is stationary in an allocentric reference frame (left) much more often than the rotation solution (right). As in the haptic case, information about self-motion in the active observer disambiguates the stimulus.

visual motion, as shown in Figures 3c–e, and it is resolved in the same way in active vision [29,30,32,52].

#### *Recalibration of visual signals after manual exploration*

Many surfaces that we see we also touch, and vice versa. Recent experiments have shown that manual exploration can influence the processing of visual depth cues. In one paradigm, the slant of a visual 3D surface is specified by conflicting depth cues (binocular disparity and static cues including motion parallax and texture gradients) [53]. The conflict between the slants specified by the two cues being small, subjects are unaware of any conflict and instead perceive a single intermediate slant. In the experiment, the subjects actively explore the virtual surface with their hands using a force-feedback device, with the feedback being compatible with one of the two depth cues. Following this adaptation phase, purely visual slant judgments are weighted more towards the visual depth cue compatible with the force feedback than they were before the adaptation [53]. Further studies have shown that other visual depth cues, such as the assumption that light comes from above, can similarly be adapted [54]. The force feedback in these experiments gives rise to a mixture of outgoing motor signals, and incoming proprioceptive and haptic signals. It would be interesting to study which of these signals is responsible for the visual adaptation (see Box 2); classic results indicate the importance of the motor command for similar types of sensorimotor learning [55].

#### *Manual action and mental rotation*

Manual action also effects higher visual cognitive tasks, such as mental rotation of 2D and 3D objects and recognition of 3D objects from novel viewpoints. It has been hypothesized that mental transformations of visual images are driven by anticipatory mechanisms that automatically predict the sensory consequences motor actions; according to this theory, mental transformations

are the results of these anticipatory mechanisms being used ‘off-line’ while the corresponding motor action is inhibited [56,57]. In agreement with this theory, it has been found [58,59] that unseen hand rotations improve performance on simultaneous mental rotation tasks when the mental rotation is in the same direction as the motor action, and decrease performance when the two rotations are in opposite directions; this effect has been shown for rotations in the image plane [58] and in depth [59]. Brain imagery and TMS studies have provided further evidence supporting this theory, showing that motor and premotor areas are activated during mental image transformations [60,61]. Actively manipulating virtual objects – as opposed to simply watching the objects undergo the same movements – has been shown to improve subsequent mental rotation [62] and recognition [63] of these objects.

#### **Conclusions**

In this article, we have reviewed ways in which motor actions – such as locomotion, eye and head movements, and object manipulation – influence the perception and representation of three-dimensional space and shape. These influences, many of which have only been discovered recently, go beyond the fact that action and perception are linked in the sense that action accompanies perception in the normal course of things, and that action almost always affects perception by modifying subsequent sensory data. These latter effects of action on perception may be called external, and have been emphasized by several authors [15,64]. The results discussed here go beyond external influence, however, showing that action has an even tighter, internal link to depth perception: action influences depth perception even in the absence of external sensory feedback.

Why would internal action–perception links develop? One possible reason may be, paradoxically, the need to *discount* the effects that one’s actions have on sensation, known as *REAFFERENCE*. The directions of points and the orientations of surfaces should be perceived as constant despite rotations on the retina during eye movements; the positions of objects should appear to remain constant despite relative motion induced by our head and body movements; and object shapes and sizes should appear invariant despite changing projections that result from viewpoint change and object manipulation. The primary role of these spatial invariances is to separate the stream of sensory data into two components, one that is under the subject’s voluntary control (called *reafference*, e.g. the 3D position and orientation of an object held in the hand) and one that isn’t (called *exafference*, e.g. the object’s shape). If the goal of perception is to glean as much new information as possible about the external world, it makes sense to separate sensory data in this way, and to concentrate resources on those variables that are *not* under the subject’s control. Many of the effects of action on depth perception that we have reviewed fall into this category.

Another internal link between action and perception may arise from ‘sensorimotor prediction’: the anticipation of the sensory effects of the motor actions that one is preparing or executing (see [65] for a review). Sensorimotor prediction plays an important role in motor

#### **Box 2. Questions for further research**

- In what cases does manipulating a 3D object influence the way we perceive its shape?
- In studies showing the recalibration of 3D vision by touch [53,54], which signals (motor or sensory feedback) are responsible for the recalibration? Would passive touch also lead to recalibration?
- Comparing responses of neurons to self- and externally produced sensory stimulation, it is found that some neurons have stronger responses to self-produced stimulation (e.g. [11]) than others (e.g. [70]). What is the reason for these two kinds of action modulation?
- The effect of active touch on haptic shape perception [51] suggests the existence of a haptic ‘stationarity assumption’ [28]. How far can this analogy between visual and haptic perception be pushed?
- Is there a general relationship between perceptual invariance and the observer’s motor action?
- Although some authors have claimed that vision is impossible without action [76], there are clearly cases where action, even if present, plays a rather uninteresting role in vision (such as ocular tremor). When sensory input is held constant, in some cases the observer’s action modifies perception and in some cases it doesn’t. Is there a principle that can predict where action effects can be found?
- The principle of genericity [77], said to account for a wide variety of phenomena in 3D vision, relies on a wide sample of views of 3D scenes. Is there a causal connection between active exploration of 3D scenes and the application of the genericity assumption?

planning as well as in perception, where it may be involved in the cancellation of the reafferent effects of motor action on sensory data that we have already discussed. Sensorimotor prediction has also been implicated in more cognitive functions such as motor and visual imagery [58], and the perception and imitation of the actions of others [66]. Some of the internal effects of action on perception might be due to interference between the predicted reafference and actual sensory data. Because sensorimotor prediction is driven by efference copy of the motor command – which is present in voluntary or active motion but absent in involuntary or passive motion (see Box 1) – the cases where active motion has a stronger effect on perception than does passive motion constitute strong evidence for the involvement of sensorimotor prediction in depth perception.

Sensorimotor prediction might play an even more fundamental role in depth perception. It is useful to think of two types of prediction, simple and complex. In simple prediction, the sensory state and planned action at an earlier time are sufficient to predict the sensory state at a later time. For example, knowledge of the location of a point on the retina and the direction and amplitude of an upcoming saccade are sufficient to predict the post-saccadic retinal location of the point. In many cases, however, this information is not sufficient, and ‘hidden’ variables are required. For example, to predict how a head movement will modify the projection of a 3D object on the retina, one must also know the 3D shape of the object, which is not immediately available from the sensory data. In fact, in almost all cases where the depth dimension is involved, complex prediction with hidden variables will prove necessary. Representations of 3D space and shape in the brain may turn out to be hidden variables subserving complex sensorimotor prediction.

### Supplementary data

There is a video illustrating optic flow in the supplementary data to this article, available online at: [doi:10.1016/j.tics.2005.06.018](https://doi.org/10.1016/j.tics.2005.06.018)

### References

- Campos, J.J. *et al.* (2000) Travel broadens the mind. *Infancy* 1, 149–219
- Held, R. and Hein, A. (1963) Movement-produced stimulation in the development of visually guided behavior. *J. Comp. Physiol. Psychol.* 56, 872–876
- Bai, D.L. and Bertenthal, B.I. (1992) Locomotor status and the development of spatial search skills. *Child Dev.* 63, 215–226
- Stackman, R.W. *et al.* (2003) Passive transport disrupts directional path integration by rat head direction cells. *J. Neurophysiol.* 90, 2862–2874
- Rieser, J.J. *et al.* (1986) Sensitivity to perspective structure while walking without vision. *Perception* 15, 173–188
- Klatzky, R. *et al.* (1998) Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychol. Sci.* 9, 293–298
- Simons, D. *et al.* (2002) Object recognition is mediated by extraretinal information. *Percept. Psychophys.* 64, 521–530
- Nishijo, H. *et al.* (1997) The relationship between monkey hippocampus place-related neural activity and action in space. *Neurosci. Lett.* 226, 57–60
- Foster, T.C. *et al.* (1989) Spatial selectivity of rat hippocampal neurons: Dependence on preparedness for movement. *Science* 244, 1580–1582
- Song, E.Y. *et al.* (2005) Role of active movement in place-specific firing of hippocampal neurons. *Hippocampus* 15, 8–17
- Zugaro, M.B. *et al.* (2001) Active locomotion increases peak firing rates of anterodorsal thalamic head direction cells. *J. Neurophysiol.* 86, 692–702
- Ekstrom, A.D. *et al.* (2003) Cellular networks underlying human spatial navigation. *Nature* 425, 184–187
- Péruch, P. *et al.* (1995) Acquisition of spatial knowledge through visual exploration of simulated environments. *Ecol. Psychol.* 7, 1–20
- Sun, H.J. *et al.* (2004) Active navigation and orientation-free spatial representations. *Mem. Cognit.* 32, 51–71
- Gibson, J.J. (1950) *The perception of the visual world*, Houghton-Mifflin
- Warren, W.H. and Hannon, D.J. (1988) Direction of self-motion is perceived from optical flow. *Nature* 336, 162–163
- Royden, C.S. *et al.* (1992) The perception of heading during eye movements. *Nature* 360, 583–585
- Crowell, J.A. *et al.* (1998) Visual self-motion perception during head turns. *Nat. Neurosci.* 1, 732–737
- Bertin, R.J. and Berthoz, A. (2004) Visuo-vestibular interaction in the reconstruction of travelled trajectories. *Exp. Brain Res.* 154, 11–21
- Wilkie, R.M. and Wann, J.P. (2003) Eye-movements aid the control of locomotion. *J. Vis.* 3, 677–684
- Bhalla, M. and Proffitt, D.R. (1999) Visual-motor recalibration in geographical slant perception. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1076–1096
- Proffitt, D.R. *et al.* (2003) The role of effort in perceiving distance. *Psychol. Sci.* 14, 106–112
- Sinai, M.J. *et al.* (1998) Terrain influences the accurate judgement of distance. *Nature* 395, 497–500
- Rogers, B. and Graham, M. (1979) Motion parallax as an independent cue for depth perception. *Perception* 8, 125–134
- Wallach, H. and O’Connell, D.N. (1953) The kinetic depth effect. *J. Exp. Psychol.* 45, 205–217
- Wallach, H. *et al.* (1974) The compensation for movement-produced changes in object orientation. *Percept. Psychophys.* 15, 339–343
- Rogers, B. and Graham, M. (1982) Similarities between motion parallax and stereopsis in human depth perception. *Vision Res.* 22, 261–270
- Wexler, M. *et al.* (2001) Self-motion and the perception of stationary objects. *Nature* 409, 85–88
- Wexler, M. *et al.* (2001) The stationarity hypothesis: an allocentric criterion in visual perception. *Vision Res.* 41, 3023–3037
- van Boxtel, J.J.A. *et al.* (2003) Perception of plane orientation from self-generated and passively observed optic flow. *J. Vis.* 3, 318–332
- Wexler, M. (2003) Allocentric perception of space and voluntary head movement. *Psychol. Sci.* 14, 340–346
- Cornilleau-Pérès, V. and Droulez, J. (1994) The visual perception of three-dimensional shape from self-motion and object-motion. *Vision Res.* 34, 2331–2336
- Deubel, H. *et al.* (1998) Immediate post-saccadic information mediates spatial constancy. *Vision Res.* 38, 3147–3159
- Naji, J.J. and Freeman, T.C.A. (2004) Perceiving depth order during pursuit eye movement. *Vision Res.* 44, 3025–3034
- Cornilleau-Pérès, V. *et al.* (2002) Visual perception of planar orientation: dominance of static depth cues over motion cues. *Vision Res.* 42, 1403–1412
- Kral, K. (2003) Behavioural-analytical studies of the role of head movements in depth perception in insects, birds and mammals. *Behav. Processes* 64, 1–12
- Gogel, W.C. and Tietz, J.D. (1979) A comparison of oculomotor and motion parallax cues of egocentric distance. *Vision Res.* 19, 1161–1170
- Bingham, G.P. and Pagano, C.C. (1998) The necessity of a perception-action approach to definite distance perception: Monocular distance perception to guide reaching. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 145–168
- Panerai, F. *et al.* (2002) Contribution of extraretinal signals to the scaling of object distance during self-motion. *Percept. Psychophys.* 64, 717–731

- 40 Marotta, J.J. *et al.* (1995) Adapting to monocular vision: Grasping with one eye. *Exp. Brain Res.* 104, 107–114
- 41 Wallach, H. and Floor, L. (1971) The use of size matching to demonstrate the effectiveness of accommodation and convergence as cues for distance. *Percept. Psychophys.* 10, 423–428
- 42 Viguier, A. *et al.* (2001) Distance perception within near visual space. *Perception* 30, 115–124
- 43 Backus, B.T. *et al.* (1999) Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Res.* 39, 1143–1170
- 44 Bingham, G.P. (1993) Optical flow from eye movement with head immobilized: “ocular occlusion” beyond the nose. *Vision Res.* 33, 777–789
- 45 Freeman, T.C.A. and Fowler, T.A. (2000) Unequal retinal and extra-retinal motion signals produce different perceived slants of moving surfaces. *Vision Res.* 40, 1857–1868
- 46 Nawrot, M. (2003) Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Res.* 43, 1553–1562
- 47 Wexler, M. (2005) Anticipating the three-dimensional consequences of eye movements. *Proc. Natl. Acad. Sci. U. S. A.* 102, 1246–1251
- 48 Heller, M.A. and Myers, D.S. (1983) Active and passive tactual recognition of form. *J. Gen. Psychol.* 108, 225–229
- 49 Chapman, C.E. (1994) Active versus passive touch: factors influencing the transmission of somatosensory signals to primary somatosensory cortex. *Can. J. Physiol. Pharmacol.* 72, 558–570
- 50 Robles-De-La-Torre, G. and Hayward, V. (2001) Force can overcome object geometry in the perception of shape through active touch. *Nature* 412, 445–448
- 51 Robles-De-La-Torre, G. (2002) Comparing the role of lateral force during active and passive touch. In *EuroHaptics Conference Proceedings*, pp. 159–164, publ. online (<http://www.eurohaptics.vision.ee.ethz.ch/2002/roblesdelatorre.pdf>)
- 52 Rogers, S. and Rogers, B.J. (1992) Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Percept. Psychophys.* 52, 446–452
- 53 Ernst, M.O. and Banks, M.S. (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433
- 54 Adams, W.J. *et al.* (2004) Experience can change the ‘light-from-above’ prior. *Nat. Neurosci.* 7, 1057–1058
- 55 Held, R. (1965) Plasticity in sensory-motor systems. *Sci. Am.* 213, 84–94
- 56 Kosslyn, S.M. (1994) *Image and brain: The resolution of the imagery debate*, MIT Press
- 57 Grush, R. (2004) The emulation theory of representation: Motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–442
- 58 Wexler, M. *et al.* (1998) Motor processes in mental rotation. *Cognition* 68, 77–94
- 59 Wohlschläger, A. and Wohlschläger, A. (1998) Mental and manual rotation. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 397–412
- 60 Windischberger, C. *et al.* (2003) Human motor cortex activity during mental rotation. *Neuroimage* 20, 225–232
- 61 Zacks, J.M. *et al.* (2003) Selective disturbance of mental rotation by cortical stimulation. *Neuropsychologia* 41, 1659–1667
- 62 James, K.H. *et al.* (2001) Manipulating and recognizing virtual objects: Where the action is. *Can. J. Exp. Psychol.* 55, 111–122
- 63 Harman, K.L. *et al.* (1999) Active manual control of object views facilitates visual recognition. *Curr. Biol.* 9, 1315–1318
- 64 Findlay, J.M. and Gilchrist, I.D. (2003) *Active vision. The psychology of looking and seeing*, Oxford University Press
- 65 Wolpert, D.M. and Flanagan, J.R. (2001) Motor prediction. *Curr. Biol.* 11, R729–R732
- 66 Wolpert, D.M. *et al.* (2003) A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B* 358, 593–602
- 67 Cullen, K.E. (2004) Sensory signals during active versus passive movement. *Curr. Opin. Neurobiol.* 14, 698–706
- 68 Scherberger, H. and Anderson, R.A. (2003) Sensorimotor transformation in the posterior parietal cortex. In *The visual neurosciences* (Chalupa, L.M. and Werner, J.S., eds), MIT Press
- 69 Klam, F. and Graf, W. (2003) Vestibular signals of posterior parietal cortex neurons during active and passive head movements in macaque monkeys. *Ann. N. Y. Acad. Sci.* 1004, 1–12
- 70 Hietanen, J.K. and Perrett, D.I. (1996) A comparison of visual responses to object- and ego-motion in the macaque superior temporal polysensory area. *Exp. Brain Res.* 108, 341–345
- 71 Bremmer, F. *et al.* (2002) Visual-vestibular interactive responses in the macaque ventral intraparietal area. *Eur. J. Neurosci.* 16, 1569–1586
- 72 Vaina, L.M. (1998) Complex motion perception and its deficits. *Curr. Opin. Neurobiol.* 8, 494–502
- 73 Sugihara, H. *et al.* (2002) Response of MSTd neurons to simulated 3D orientation of rotating planes. *J. Neurophysiol.* 87, 273–285
- 74 Murray, S.O. *et al.* (2002) Shape perception reduces activity in human primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15164–15169
- 75 Amedi, A. *et al.* (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 4, 324–330
- 76 O’Regan, J.K. and Noë, A. (2001) A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–1031
- 77 Nakayama, K. and Shimojo, S. (1992) Experiencing and perceiving visual surfaces. *Science* 257, 1357–1363

### Elsevier.com – Dynamic New Site Links Scientists to New Research & Thinking

Elsevier.com has had a makeover, inside and out.

As a world-leading publisher of scientific, technical and health information, Elsevier is dedicated to linking researchers and professionals to the best thinking in their fields. We offer the widest and deepest coverage in a range of media types to enhance cross-pollination of information, breakthroughs in research and discovery, and the sharing and preservation of knowledge. Visit us at Elsevier.com.

**Elsevier. Building Insights. Breaking Boundaries**