More to 3-D vision than meets the eye

Mark Wexler

Does self-motion affect object recognition? Researchers have long studied how 3-D objects are recognized from different points of view, but have disregarded the observers' own movement by keeping them motionless. A recent study by Simons *et al.* shows that self-motion cannot be ignored, as it changes the way that objects are recognized.

One of the accomplishments of our visual system is to recognize three-dimensional objects from different points of view. This is useful because objects can move and change orientation (either through our own actions, or independently of us), or because we ourselves can move through space and so experience objects from different angles. For every conceivable self-motion in an environment of stationary objects, there exists an equal-and-opposite *object* motion such that the resulting retinal stimulation is identical in both cases (e.g. see Fig. 1). Because object recognition has been thought to depend exclusively on retinal data, these two kinds of transformations have been considered as functionally equivalent, at least as far as object perception is concerned. Experiments involving object motion are much easier to do: projections of objects from different points of view can be shown on a computer screen while the observer remains still. By contrast, experiments involving self-motion are harder to set up – either real 3-D objects or complex virtual reality systems have to be used. As a result, almost all work on object recognition has been done with object motion, and the dependence of recognition on retinal information has not been questioned.

Self-motion versus object motion

In a recently published paper, Simons, Wang and Roddenberry show that it is a mistake to confound viewpoint changes and object transformations in object recognition [1]. They do so by comparing object recognition in self-motion and object-motion conditions (for their previous work on the subject, see [2,3]). In their self-motion condition (Fig. 1a), subjects first view a 3-D block object through a narrow window. Then after walking to a point that corresponds to a 40° rotation about the object's center, they view the object through another window. Subjects do not see the object during their movement. The experimenter, meanwhile, either leaves the object unchanged, or replaces it with a similar but distinguishable object, and the subjects' task is to tell whether the second object seen is the same as the first (despite the viewpoint change), or different. In the object-motion condition (Fig. 1b), the subject remains still while the (unseen) object undergoes an equal-and-opposite 40° rotation, during which it either remains the same or is replaced by a different object; the task is again to tell whether the object is the same in the two views or not.

Thus, in Simons *et al.*'s experiment, the retinal images of the objects are the same in the self-motion and object-motion conditions. Therefore, if objects are recognized from retinal information alone, performance should be the same in the two conditions. Simons and his co-workers, however, found that recognition performance was significantly better for self-motion than for object motion [1]. This remained true even when the visual backgrounds were uniform, so that all retinal data were strictly identical in both views in both conditions. The inescapable conclusion, which comes as a surprise after decades of work on object recognition, is that recognition performance depends on more than the retinal images of objects – self-motion also makes an important contribution.

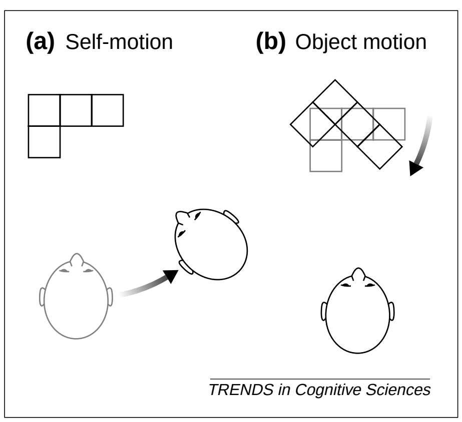

Fig. 1. Self-motion and object motion. This shows that for every viewpoint change arising from self-motion in a stationary environment (a), there exists a rigid object transformation that leads to the same retinal input (b). In spite of the same retinal stimulation, recent results show that object-recognition performance in the two cases can differ, as can the perception of 3-D object shape.

Simons *et al.* have demonstrated that self-motion has an effect on object recognition, but where does self-motion information come from? There are three potential sources. First, if the motion is voluntary, a copy of the motor command can be used to infer movement. Second, once the body is in motion, proprioceptive (vestibular, somatosensory) signals provide self-motion information. Third, if the self-motion is performed with the eyes open in a stationary environment, optic flow also provides self-motion information. In the experiments of Simons *et al.*, all three of these channels are potentially informative. In particular, because the subjects were allowed to keep their eyes open, optic flow was different between the initial and final views in the self-motion and object-motion conditions. It cannot therefore be concluded that object recognition is mediated by extra-retinal information, but only by self-motion information, which – as far as we know from the data presented by Simons *et al.* – may be retinal in origin. One way to exclude this possibility would be, following Rieser [4], to blindfold subjects as they change their viewpoint in the self-motion condition; to the extent that the difference between self- and object-motion conditions persisted in this case, one could exclude optic flow from an explanation of the effect found by Simons *et al.* If optic flow is not responsible for the effect, one could distinguish between the motor command and proprioceptive explanations using Held's technique [5] of moving the subject passively between the two viewpoints, thus suppressing the motor command but leaving proprioceptive feedback mostly intact.

The results of Simons *et al.* seem to be in agreement with a growing consensus that at least some kinds of visuo-spatial transformations are intimately connected with motor processes, and possibly driven by the motor system. Behavioral evidence for this has been provided in both mental rotation [6–8] and object recognition [9] paradigms; imagery data supporting this view have also been presented [10–13].

 1364-6613/02/$ – see front matter PII: S1364-6613(02)02023-5

One possible explanation is that the anticipatory mechanisms of the motor system, when run 'off-line' as in motor imagery, geometrically transform internal visual representations. Interestingly, in one of the object-motion conditions of Simons *et al.*, either the subject or the experimenter turned the table (and therefore the object). Although the anticipation hypothesis would predict that recognition would be facilitated in the subject-initiated move condition, Simons *et al.* found no significant effect – but perhaps their accuracy measure was not sensitive enough to reveal this.

Are spatial representations allocentric?

Another way of viewing the results of Simons *et al.* is in terms of reference frames of visual representations. For simplicity, consider two types of reference frame: 'egocentric', or centered on the observer, and 'allocentric', or fixed to the earth and independent of the observer. In an egocentric frame, the self- and object-motion conditions are identical. In an allocentric frame, on the other hand, the object remains fixed in the self-motion condition, but turns relative to the earth in the object-motion condition (as depicted in Fig. 1b). Thus, if object representations are allocentric, the object recognition task would be easier in the self-motion condition (where no rotation of the object representation would have to take place) than in the object-motion condition (where the representation would have to undergo rotation) – precisely the results found by Simons and his co-workers.

The only problem with the allocentric explanation is that visual inputs and retinal images are, of course, egocentric. In order to convert these egocentric data to an allocentric frame, the visual system would have to compensate for the egocentric sensory data – which provide information on the *relative* motion between object and object – by the observer's own movement. Indeed, in the context of the perception of 3-D shape from motion, it has also been shown recently that self- and object-motion yield different perceptual results, in spite of identical optic flow [14]. The same study demonstrated that allocentric representations appear early in spatial vision, and that extra-retinal information is used to compensate egocentric retinal inputs to make this possible [14]. Neurophysiologically, the existence of allocentric reference frames in vision is supported by the discovery of a hierarchy of spatial reference frames – ranging from egocentric to allocentric – in the parietal cortex of monkeys [15].

In conclusion, the results of Simons, Wang and Roddenberry demonstrate that vision researchers cannot study self-motion fully by merely simulating its effects through object motion: the observer's actual movement changes the way he or she recognizes 3-D objects. This study contributes to a growing body of evidence for the importance of motor action for visual perception.

References

1 Simons, D.J. *et al.* (2002) Object recognition is mediated by extraretinal information. *Percept. Psychophys.* 64, 521–530
2 Wang, R.F. and Simons, D.J. (1999) Active and passive scene recognition across views. *Cognition* 70, 191–210
3 Simons, D.J. and Wang, R.F. (1998) Perceiving real-world viewpoint changes. *Psychol. Sci.* 9, 315–320
4 Rieser, J.J. (1989) Access to knowledge of spatial structure at novel points of observation. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 1157–1165
5 Held, R. and Freeman, D.J. (1963) Plasticity in human sensorimotor control. *Science,* 142, 455–462
6 Wexler, M. *et al.* (1998) Motor processes in mental rotation. *Cognition* 68, 77–94
7 Wohlschläger, A. and Wohlschläger, A. (1998) Mental and manual rotation. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 397–412
8 Schwartz, D.L. and Holton, D. (2000) Tool use and the effect of action on the imagination. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1655–1665
9 James, K.H. *et al.* (2001) Manipulating and recognizing virtual objects: where the action is. *Can. J. Exp. Psychol.* 55, 113–122
10 Richter, W. *et al.* (2000) Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *J. Cogn. Neurosci.* 12, 310–320
11 Ganis, G. *et al.* (2000) Transcranial magnetic stimulation of primary motor cortex affects mental rotation. *Cereb. Cortex* 10, 175–180
12 Lamm, C. *et al.* (2001) Evidence for premotor cortex activity during dynamic visuospatial imagery from single-trial functional magnetic resonance imaging and event-related slow cortical potentials. *NeuroImage* 14, 268–283
13 Vingerhoets, G. *et al.* (2001) Regional brain activity during different paradigms of mental rotation in healthy volunteers: a positron emission tomography study. *NeuroImage* 13, 381–391
14 Wexler, M. *et al.* (2001) Self-motion and the perception of stationary objects. *Nature* 409, 85–88
15 Snyder, L.H. *et al.* (1998) Separate body and world-referenced representations of visual space in parietal cortex. *Nature* 394, 887–891

Mark Wexler
Laboratoire de Physiologie de la Perception et de l'Action, CNRS/Collège de France, 11, Place Marcelin Berthelot, 75005 Paris, France.
e-mail: wexler@ccr.jussieu.fr