Primer

The visual perception of motion

Robert J. Snowden and Tom C. A. Freeman

Sensing the movements of the world and the objects within it appears to be a fundamental job for our visual system. In rare cases of brain damage, we find that individuals lacking motion perception live in a very different world of frozen images, where simple tasks like filling a kettle or crossing the road take on alarming difficulties.

That tasks such as driving a fast car down the freeway require a good sense of the movements of yourself and other objects are obvious, but motion information is used in many less obvious ways. For example, it may seem a trivial task to us to follow a moving object with our eyes, but without motion perception these smooth pursuit eye movements are not possible. One way to show this is to have people attempt to move their eyes smoothly along a line etched on a wall. At the same time we place a very bright light just under this line. The bright light burns an afterimage into the retina which can then be examined at leisure. In attempting the eye movement we find that the afterimage is not a smooth line but a series of 'dots'. This is because you find it difficult to smoothly move yours eyes along the line — even more so if it is vertical — instead flicking your eyes in a series of small, fast jumps known as saccades. Each dot represents the alighting point of a saccade and each gap the distance moved by the saccade. Now, if the line is replaced with a moving dot and we try to track this, the resulting afterimage is a smooth line. This shows that our eyes moved at a constant rate so that the very centre of our vision, where it is best at seeing fine detail, remained focused on the target. Not surprisingly, damage to areas of the brain involved in analysing the moving image destroys this ability.

Two routes to motion

The above illustration suggests we can sense the motion of an object as we follow it with our eyes. Of course, if we think about what is happening in the image during an eye movement the situation is somewhat different, because the moving object we are tracking is actually stationary on our retina and the stationary world around it moves across our retina in the opposite direction. Clearly the situation on the retina needs 'interpretation' by the higher centers of the brain if we are to correctly deduce the actual motions about us — we will see later that this is a problem that occurs at many levels.

With respect to our example, it would help a great deal if we knew the way our eyes were moving [1]. Information like this could be gathered either from feedback from the eyes themselves (outflow theory) or from the signals that were sent to move the eyes (inflow theory). Deciding between inflow and outflow theory has led to some of the more adventurous forms of experimentation in human vision science. One trick has been to try and prevent intended eye movements, either by anaesthetizing the eye muscles (or indeed the entire body!) or physically restricting the eye movement using putty. Another trick is to try and induce unintended eye movements, the more extreme techniques including pulling on eye-muscle tendons with forceps or using a hydraulic pulley system attached (by wire) to a suction cap adhering to the eyeball. One group of researchers identified an understandable impediment to this type of research when they lamented that “It is difficult to recruit more subjects”.

In the end a gentle poke in the eye with a finger may suffice and can be used to demonstrate the majority conclusion that inflow theory is more likely correct. As the eye is made to move against its own volition the world loses its stability. In this situation no signal is being sent to the eye to move, hence the motion on the retina caused by the eye moving cannot be compensated for and the movement is interpreted as the world moving. Damage to this signal for the eyes to move leads to a most debilitating condition where the movements of the eyes cause the world to appear to rush by.

Knowing how the eye moves might allow the motion of a tracked object to be judged, or help compensate for the retinal motion created by the eye movement, but it tells us little about how we sense the motion in different parts of the image. If you gaze out of the window it is clear that you can track individual cars on the road, but the very complex movement of pedestrians milling about, or of leaves swaying in the breeze, is also being sensed. No eye-movement-based motion system can hope to track all these objects at once. Most research has actually examined this latter situation, where the retina is stationary and the objects move across it.

A motion detector

Current models of how to extract the early components of retinal motion owe much to the idea that movement involves something being 'here' at one point in time and 'there' sometime later. To spot this 'here-then-there' we need a detector that compares two regions of space at two slightly different times — known colloquially as a 'delay and compare' scheme.

Perhaps the most elegant exposition of this idea has come from Reichardt and colleagues [2]. From observations of the walking behaviour of the beetle *Chlorophanus*, they hypothesised that the signals at two locations on the retina are multiplied, after one has been delayed, to produce a motion signal. Many other models of early perception have been developed, but they are all similar to the architecture or computation of the Reichardt model [3].

Not surprisingly, development of models of motion detection has leant heavily on how this is achieved in the animal kingdom. Cells have been found in many creatures that are sensitive to the direction and/or speed of the movement. Some nerve cells respond vigorously if a spot of light is moved from left to right across their receptive field [4], the area on the retina from where an individual cell receives its primary information; but if the same spot moves from right to left, the same cells remain silent.

In many animals this type of cell is found in the retina, but in primates direction selectivity appears much later. The first cells with this property are in our primary visual cortex, area V1. Even within this area, however, only about 20% of cells show this property, confined to particular layers within V1. These directionally selective cells appear to send their output to just a few specific areas of the brain. In one of these projection areas, the middle temporal area MT (often termed V5), nearly all the cells appear to be directionally selective, so it is not surprising that this area is thought of as our 'motion center'.

A place for motion

The importance of area MT for motion processing has been shown in a number of ways. One is to make a very small lesion within the area: such an animal is still able to see a moving stimulus and report its detail, colour, size and so on; but if the animal has to report its direction or speed of movement, it is severely compromised [5]. Performance recovers if the stimulus is repositioned so that it no longer falls on the site where the lesion took place. Even more ingeniously, we can change the perceived movement of a stimulus simply by electrically stimulating the cells in this area. Here, a stimulus is presented with an ambiguous direction of motion, for example a collection of dots that move in random directions. If the animal has to report whether such a stimulus moved upwards or downwards, it should guess upwards about 50% of the time. But if we locate a cell in area MT that is most responsive to upwards motion and stimulate it when the dots appear, the animal nearly always reports 'seeing' upwards motion. We know the dots must appear to the animal to move upwards, because if the cell is stimulated without the dots present, the animal does not respond.

Experiments in humans agree with this idea of a special place for motion. We can compare brain activity when a person is experiencing motion with that activity when it is not. One clever way of achieving this exploits illusions where the person experiences a sense of motion despite a stationary image (and eye). One of the most celebrated is that of the motion aftereffect (see ViperLib http://viperlib.york.ac.uk/). This is sometimes known as the waterfall illusion after Addams' famous description of the phenomenon he experienced after viewing the Falls of Foyer on Loch Ness [6]. It can be achieved by looking at something that moves in a particular direction for a period of time — the longer the better, hence waterfalls are particularly useful. If you now stare at a stationary object, such as the rocks near the waterfall, they appear to move in the direction opposite to that previously watched. As Addams put it in 1834: "I saw the rocky surface as if in motion upwards, and with an apparent velocity equal to that of the descending water, which the moment before had prepared my eyes to behold this singular deception".

If we could compare the activity in Addams' brain before and after he stared at the waterfall, we would be able to delineate the 'motion centers' of the brain. Indeed experiments using functional magnetic resonance imaging (fMRI) have shown just such activity in the human brain when a person is experiencing the motion after effect [7]. The place of this activity corresponds with an area that is easily activated by motion, and hence is thought of as the human homologue of MT. The activity also seems to last as long as the motion after effect does, so if the person stares at our laboratory waterfall for a long time, we get both motion after effect and brain activity lasting the same amount of time.

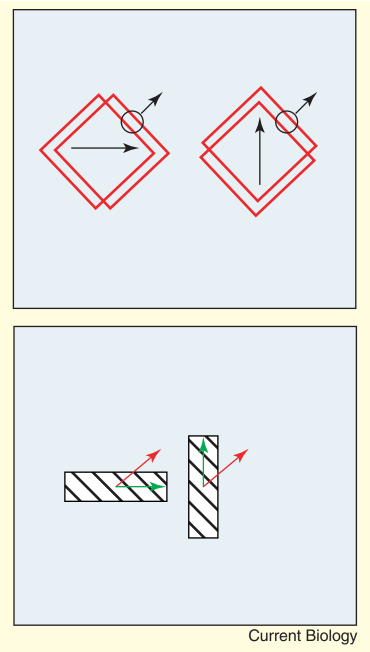
Current Biology

Figure 1. (A) The aperture problem. Consider a diamond moving to the right (left side of figure) or upwards (right side). Its position is illustrated at two points in time. Within the small receptive fields — apertures — the motion is identical for the two very different movements of the diamond. (B) A static illustration of the barberpole illusion.

Cells in area MT respond to simple uniform motion, with each cell responsible for analysing a small patch in the image. But area MT is not the final point of motion analysis in the brain for at least two reasons. The first is that a motion detector with a small receptive field suffers from what is known as the aperture problem. The detectors illustrated in Figure 1 can only signal motion at right angles to the edge of the moving diamond — it cannot see motion parallel to this, because there is no pattern in that direction to stimulate the detector. The movement signaled by a motion detector with a small receptive field is consistent with an edge moving in any one of an infinite number of directions and speeds: the motion information provided by any individual local motion detector is ambiguous.

A number of methods for making sense of these local detectors have been suggested [8], including those that combine

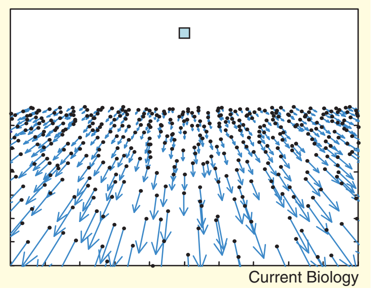

Figure 2. The optic-flow pattern when moving in a straight line across an evenly textured ground-plane. The square indicates the direction of heading. All image movement radiates away from this point with a speed that depends on the distance of the texture element from the eye.

information across detectors, limiting the number of possible velocity interpretations, and those based on motion information carried by image features that do not suffer the aperture problem, such as the points where edges terminate. A compelling demonstration of the importance of terminators is given by the barberpole illusion (Figure 1B). Whilst the stripes within the barberpole physically move up and to the right, we only see the motion along the long axis of the rectangle. This is because the terminators move along the edge and there are more of them on the long edge than on the short edge.

Solutions to the aperture problem suggest the second reason why MT does not have the final say in perceiving visual motion. At their heart lies the idea that motion information must be integrated across detectors in order to reveal the true nature of movement in the world. Integration is also important because detectors in MT cannot tell us about the more complex patterns of motion regularly experienced in daily life. The real world rarely allows us the luxury of simple, uniform movement across the entire image. Instead there is a complex tangle of motion as our eyes and bodies move in a world containing three-dimensional objects that also move. So we are faced with the problem of how to combine the output of our simple motion sensors in order to understand the more complex motion patterns of real life.

Other places for motion

So far we have confined our discussion to fairly simple motions, such as a spot moving across a screen or the cascade of a waterfall. But as we have just seen, motions in the real world are much more complicated than this. Things can move in depth, rotate, speed up or slow down and produce sophisticated patterns such as those associated with the way a human body moves when walking.

Consider a person driving a car over rough ground (Figure 2). The image related to the scene in front of them has a characteristic expanding motion, so that more distant elements on the ground move slowly but get faster and faster as they approach the car. At the same time, the elements in line with the point towards which the car heads move directly toward it, whilst those off to the side move gradually away. Large patterns of image motion such as these are often called optic flow. One can easily imagine other optic flow patterns caused by other types of motion. It has been suggested that optic flow might be very useful for all sorts of tasks where the person is moving in the environment — including driving a car or flying a plane — and might therefore be detected by special cells sensitive to such patterns. The cells of area MT do not seem to appreciate these optic flow patterns, but an area adjacent to this known as the medial superior temporal (MST) area contains cells sensitive to movements that expand, contract, rotate or deform. Again, sophisticated brain imaging in humans, together with cellular recordings in primates, have provided evidence of an equivalent region in the human brain.

Vection

One of the most striking illusions caused by image motion is the feeling that it is ourself that is moving rather than the image. Most people will have had the strange feeling when sat in a train, believing that their journey has just begun, only moments later to realise that they are still in the station and it is the train next to them that has pulled away. This feeling of self-motion is termed ‘vection’ and, not surprisingly, is normally felt when we are actually moving.

One way of studying this phenomenon is to place a person in a large drum. When the drum begins to rotate most people report they are sat on a stationary stool inside a rotating drum — quite correct. But after a few seconds the experience changes and people start to report the drum stationary and themselves rotating on the stool; many subjects debriefed after such experiments refuse to believe that the stool cannot rotate at all. Modern cinemas, such as IMAX, often use large screens to reproduce these feelings of self-motion in an attempt to make the experience more ‘real’. Large screens are helpful, but not necessary for this experience. What really appears to govern whether I believe it is myself that is moving is what I believe the background is doing. If I sense the walls are ‘moving’, then my visual system seems wired to interpret this not as moving walls, which are unlikely, but as myself moving past stationary walls.

The issue of what is ‘background’ is less easy to define, but the case is well illustrated by the following experiment [9]. Let us imagine that we look at a large sheet of green dots that are moving upwards. This should give us the feeling that we are moving downwards. Now, imagine that we add in red dots that are moving downwards and we get the observer to concentrate on these red dots — which way would they feel themselves moving? Most people reply that, as they are concentrating on the downward-moving red dots, you should feel as if you are moving upwards. But the answer is the opposite: you feel as if you are moving downwards. This is because the dots you are concentrating on, the red downwards ones, become the foreground, and the ones you are *not* concentrating on, the green upward moving ones, become the background, and vection is dictated by the background. This is a rare case where your actions and/or perceptions are actually governed by the thing you are not trying to pay attention to.

An illusion of motion

Illusions have played a large role in our understanding of perception in all the senses, and visual motion perception is no exception. The illusion illustrated in Figure 3 is by the artist/scientist Akiyoshi Kitaoka. It is as striking in its elegance as it is in its art, and illustrates many of the principles this primer has tried to outline. If you look at the centre of the pattern and rock backwards and forwards, you should see the pattern rotate! Why?

As your head moves towards the picture this causes characteristic expansion motion on the retina (optic flow). If we consider the rightmost element within the pattern, we see that it should move to the right on our retina as the head moves forward. The ‘trick’ to this illusion lies in the oriented line that runs through the circle — note that this element is somewhat away from the true vertical. Now our early motion detectors can only signal motion perpendicular to the element — the aperture problem — and the motion signal is therefore strongest in the detectors that signal ‘right and a bit up’ in this part of the image. One can construct a similar story for all the other elements of the pattern. Our visual system also needs to take into account our own head/eye movements, through inflow type theory or by the fact that our ‘expansion’ detectors are activated, and subtract this expansion component from the retinal image motion. A pure expansion would lead to the ‘right’ part of the image motion being subtracted leaving us with the ‘and a bit up’, thus this element appears to move ‘a bit up’. Once a similar story is constructed for each element we can see that each element moves a little, and together they form a rotation [10].

Conclusion

We understand the basics of motion perception perhaps better than that of any other of human behaviour. This understanding has come from a fusion of psychology, psychophysics, neurophysiology and computational modeling. But this understanding also serves to highlight our areas of ignorance and the many avenues as yet unexplored. The field of visual motion perception has been a test-bed for much blue sky research, and a flagship for the exploration of both human and animal brains. It should not be forgotten that the research has many practical applications, be it issues to do with driving in fog, simulating flight or surgery, or understanding the changes that take place with both disease and age.

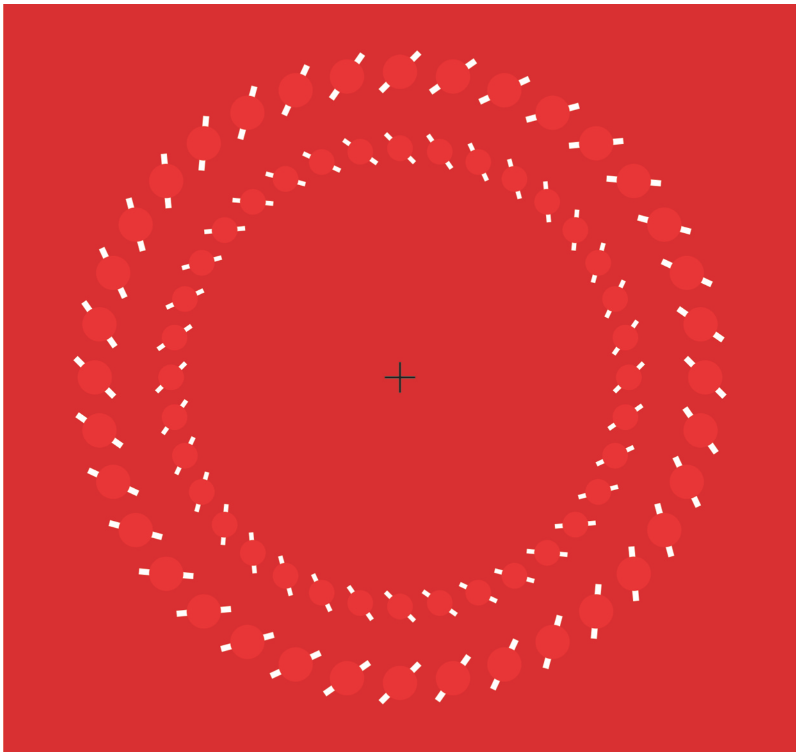

Figure 3. The illusion ‘Candies2’ by the artist Akiyoshi Kitaoka (© 2002). If you move your head backwards and forwards you should find that the outer circle of elements rotate in one direction, whilst the inner circles rotate in the opposite direction. The rotations will alternate depending on whether you move towards or away from the picture. See text for a possible explanation.

References

1. Sumnall, J.H., Freeman, T.C.A., and Snowden, R.J. (2003). Optokinetic potential and the perception of head-centred motion. Visi. Res. *43*, 1709-1718.
2. Hassenstein, B., and Reichardt, W. (1956). Systemtheoretische analyse der zeit-, reihenfolgen- und vorzeichenauswertung bei der bewegungspwezeption des rüsselkafers Chlorophanus. Z. Naturforsch *11b*, 513-524.
3. Borst, A., and Egelhaaf, M. (1989). Principles of visual motion detection. Trends Neurosci. *12*, 297-306.
4. Lennie, P. (2003). Receptive fields. Curr. Biol. *13*, R216-R219.
5. Newsome, W.T., and Paré, E.B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). J. Neurosci., *8*, 2201-2211.
6. Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body etc. London and Ed. Philos. Mag. J. Sci. 3rd series, *5*, 373-374.
7. Tootell, R.B.H., Reppas, J.B., Dale, A.M., Look, R.B., Sereno, M.I., Malach, R., Brady, T.S., and Rosen, B.R. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic imaging. Nature *375*, 139-141.
8. Snowden, R.J., and Verstraten, F.A.J. (1999). Motion transparency: making models of motion perception transparent. Trends Cogn. Sci. *3*, 369-377.
9. Kitazaki, M., and Sato, T. (2003). Attentional modulation of self-motion perception. Perception *32*, 475-484.
10. Morgan M. J. (The Guardian 24 January 2002)-http://www.guardian.co.uk/Archive/article/0.4273,4341518,00.html

School of Psychology, Cardiff University, Cardiff CF10 3YG, UK.
E-mail: Snowden@Cardiff.ac.uk