PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The Perception of Apparent Movement [and Discussion]

S. M. Anstis and D. M. Mackay

Phil. Trans. R. Soc. Lond. B 1980 **290**, 153-168
doi: 10.1098/rstb.1980.0088

Phil. Trans. R. Soc. Lond. B **290**, 153–168 (1980)
Printed in Great Britain

The perception of apparent movement

By S. M. Anstis
Department of Psychology, York University, 4700 *Keele Street, Downsview, Ontario M3J 1P3, Canada*

When two similar pictures, overlapping but slightly displaced, were projected on a screen in alternation, apparent movement could be seen. How similar must successive pictures be to give apparent movement? This is the 'correspondence problem'. Manipulations of the local and global correspondences between pictures included motion phenomena such as reversed apparent movement; a four-stroke oscillatory cycle which gave an illusion of continuous motion in one direction; edges defined by texture, stereoscopic depth, or flicker, kinetic edges; and wave motion. It was concluded that human motion perception may comprise two separate mechanisms. Local point-by-point correlations between pictures are detected by a relatively peripheral system, probably based on directionally selective neural units. More subtle global correspondences are analysed by a more cognitive system which extracts edges before it processes motion.

1. Introduction

Motion detection is one of the most ancient and primitive forms of vision (Walls 1942). Perhaps this is because it is so important for survival. For any animal, the most important visual stimuli are other animals, which must be detected and recognized quickly and reliably. Potential mates must be approached, links in the food chain above one's own position are predators to be avoided, and links below are prey to be captured. Colour vision and shape recognition are often effective in finding a mate, because they are aided by, for example, mating plumage. However, these systems alone would frequently fail to detect prey or predators because they would be defeated by natural camouflage. Motion detection is more reliable: prey and predators may freeze in a fixed posture to escape detection, but sooner or later they must make a move, and give themselves away to any eye that is equipped with motion detectors.

For humans too, motion gets attention. To catch the eye of a friend in a crowded airport one could wear a red hat, but it is more effective simply to wave a hand and trigger off his peripheral motion detectors. Motion not only gets attention but holds attention. Most people, given a choice between a photograph or a painting which has colour but not motion, and a black and white television, which has motion but not colour, will find their attention held for far longer by the latter than by the former.

To study motion in the laboratory, the stimulus is often simplified from a continuous movement to a set of discrete stimuli which are flashed in sequence. This gives a convenient and tractable display; moreover, and more importantly, it exposes a discrepancy between the stimulus and the percept. Why does the intermittent stimulus look as though it is moving smoothly? This is the phenomenon of apparent movement (a.m.), which was studied by Wertheimer (1912) and by Korte (1915). These authors mapped out the intervals in time and space that gave optimal apparent movement, and this research tradition has been ably extended by Morgan, whose article elsewhere in this symposium explores the spatial and temporal filtering that are produced by the space and time constants of the visual system.

Wertheimer's classic study used simple dot and line stimuli, and the direction of a.m., if it occurred, could only be between the two positions of his flashed spot. His simple stimulus constrained a simple answer. This constraint, however, has been removed in various recent experiments that looked for a.m. between pairs of complex pictures. Now, the use of these stimuli raises a new question which Wertheimer never asked, namely: How similar must successive pictures be for motion to be seen between them? This will be referred to as the 'correspondence problem', because the visual system has to decide which are the corresponding points in the two pictures, between which motion is to be seen. If two identical pictures are presented in sequence, slightly out of register, the perceptual problem seems easy to solve because there is a simple, one-to-one correspondence between every point in one picture and every similar point in the other. Suppose now, though, that a cluster of ten spots in one picture is followed by a cluster of another ten spots, shifted and slightly disarranged, in a second picture. For each initial spot, the visual system must select one of the spots in the second picture as being the true corresponding target, and reject the other nine as being phantom targets (see Julesz (1971) for a discussion of a similar problem in stereopsis). We confront and solve such perceptual problems every time we go to the cinema. It is only in the most boring movie sequences, such as a slow pan over a static landscape, that successive movie frames are approximately identical. As soon as the real action starts, say when a man runs across the screen, then his limbs flex and his body forms a different geometry on each frame, yet one automatically identifies him as the same visual stimulus or object in each frame, and sees a continuous object in motion. The correspondences between successive frames are sufficient to carry the percept of motion. But if the scene cuts from a man standing by his horse to a new shot of a cactus in the desert, one correctly perceives this as a change of scene; not as a man who jumps across the screen and changes into a cactus. In this case the correspondence or similarity between pictures is not sufficient to give a percept of motion. In what follows I shall attempt to specify the degree, and kind, of similarity that will just give apparent movement; our central problem, then, is the correspondence problem. My main technique is to project in alternation on a screen a pair of 35 mm slides, which are almost but not quite in register. The similarity between the two pictures can be manipulated, and the timing of the alternate exposures can be varied: one may either cut between pictures i.e. switch one picture on and the other off at the same time, or else one may dissolve between them, i.e. fade one picture down as the other fades up, as in the Victorian pastime of dissolving views. Although timing is not our main concern, it will be shown that small alterations in timing can radically alter the perceptual organization of apparent movement in some cases. To anticipate my conclusions, I shall suggest that human visual motion perception involves two quite different systems or mechanisms. (Braddick (1974) was the first person to put forward this idea.) Identical or closely similar pictures, with only a small spatial shift between them, are analysed by a simple point-by-point mechanism (system 1) based on fairly peripheral, hard wired motion detectors. More subtle correspondences between pictures will defeat this simple mechanism, and these are analysed by a more 'cognitive' system (system 2) which extracts edges or forms before it processes motion.

A homely illustration of the notion of perceived object similarity as a prerequisite for one type of perceived motion comes from the old story of Brer Rabbit, who challenged Brer Hedgehog to a five mile race. Both animals started together at the first milestone, and Brer Rabbit ran the first mile as fast as he could. But he found Brer Hedgehog was at the one-mile stone ahead of him. He ran even faster over the second mile, but again Brer Hedgehog was there before him. The same thing happened at every milestone, so Brer Rabbit lost the race. What he did not know was that there were really six hedgehogs: Brer, Mrs, and four little hedgehogs. Brer Rabbit's movement was all too real, but the hedgehog's movement was only apparent, and depended crucially on the fact that all the hedgehogs looked sufficiently alike to be taken for the same individual.

Having drawn attention to the existence of the 'correspondence problem', we shall turn now to consider the logical structure of available models of movement detection, to illustrate the magnitude of the 'correspondence problem' as a threat to theoretical orthodoxy.

2. A SIMPLE MODEL

We shall present a simple analogue model of a motion sensor, and then try to apply it to human visual motion perception. We shall suggest that it successfully models the simpler forms of motion perception which depend upon hard-wired neural motion detectors (see Grüsser & Grüsser-Cornehls (1973) for a review), but it fails, in an instructive way, to explain the subtler, more 'cognitive' forms of motion perception.

The model can be embodied in optical or electronic form. In principle, it could also be expressed as a set of equations, or as a computer simulation, but the hardware analogue forms are easier to understand. We shall start with a simple version which senses change but is non-specific for direction, and then we shall add directional selectivity to the model.

Consider the photograph of a clock in figure 1*a*. This photograph has been specially reprocessed in figure 1*b* and *c* so that in figure 1*b* the stationary parts – the case and dial – are still visible, but the moving parts – the hands and the gear teeth – have disappeared. Figure 1*b* is a 'summed picture'. It is a time exposure, in which the static parts of the scene slowly burned themselves into the photographic emulsion, but the moving parts blurred out and disappeared. Conversely, in figure 1*c* the moving parts are clearly visible, but the static parts have been greatly attenuated or have disappeared. Figure 1*c* is a 'difference picture' (Mackay 1959). The clock was photographed twice in succession, and the later picture was optically subtracted from the earlier by sandwiching the negative of one picture with the positive of the other. Thus, in figure 1*b* temporal *summation* accepted static objects and rejected changing or moving objects, whereas in figure 1*c* temporal *differencing* achieved the opposite. Now, the optical process of temporal differencing can distinguish changes over time, but not the direction of movement. Directional selectivity can be added to the model by a 'shift-and-subtract' process: the procedure here would be to take a negative of a movie frame and spatially shift it around over a positive of its preceding frame until the best match was found. The movement represented by the two frames would be directly given by the direction of the required shift. If different regions of the picture contained movements in different directions, one solution would be to cut the picture up into little pieces and slide them around independently. A practical and flexible method for doing this (in effect) is to examine the scene point by point, using an array of detector pairs with a fixed shift or separation between each pair. Of course, many pairs of detectors would be needed. The operation that compares the two picture samples need not be subtraction: multiplication (correlation) or addition could be used instead, with suitable small modifications to the model.

Barlow & Levick (1965) suggested that a directionally selective neuron in the rabbit retina works along these lines. In their model (see figure 2*a*), receptors A and B sample adjacent retinal regions. B's output is delayed and then subtracted from A's output. If a stimulus spot moves from B to A (the null direction) then B inhibits A and there is no output. But if a spot moves from A to B (the preferred direction), B's output arrives too late to inhibit A, and there is an output which signals motion.

In principle, the interaction between A and B need not be subtractive. Barlow & Levick pointed out that addition (excitation) would do just as well, but they found direct evidence demonstrating that in fact the rabbit neurons use subtraction (inhibition). Reichardt (1961) proposed a model of motion detection by the beetle's eye, in which B's delayed output, B′, was *multiplied* by A's output (see figure 2*b*). Multiplication measures the correlation between the patterns seen by A and B, and the product AB′ reaches a maximum value when the same pattern

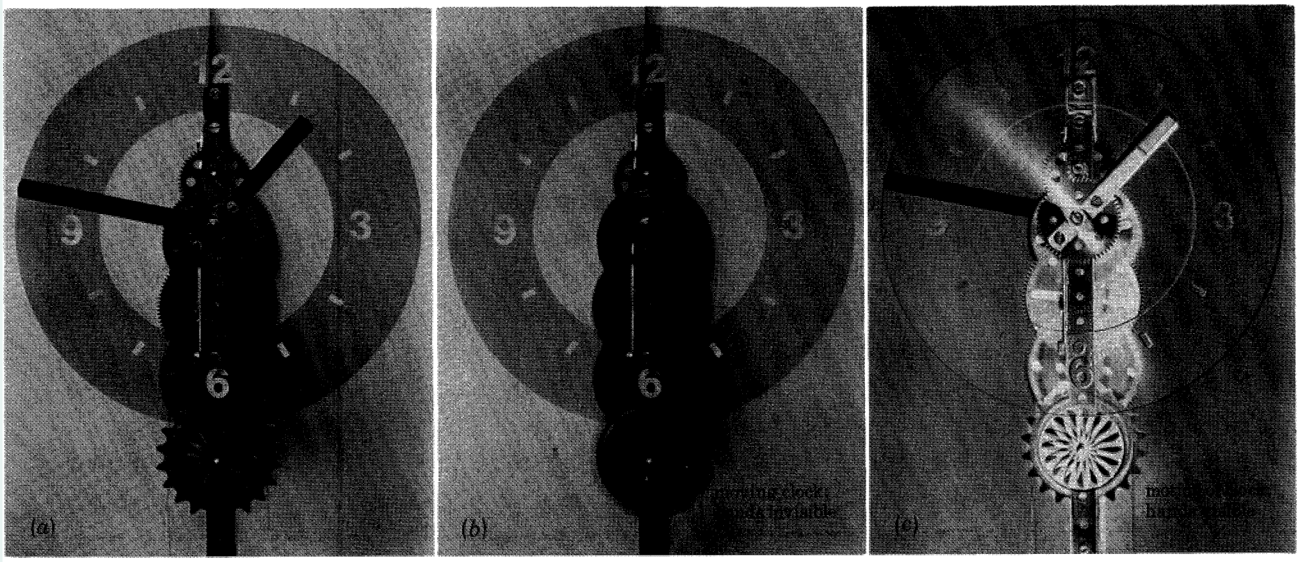

FIGURE 1. Simple model of change (or motion) detector. (*a*) Original photograph; (*b*) summed picture: time exposure shows static parts but moving parts blur out and disappear; (*c*) difference picture made by superimposing negative picture on a positive picture taken earlier: parts that have moved or changed show up clearly (black or white), but static parts cancel out and disappear.

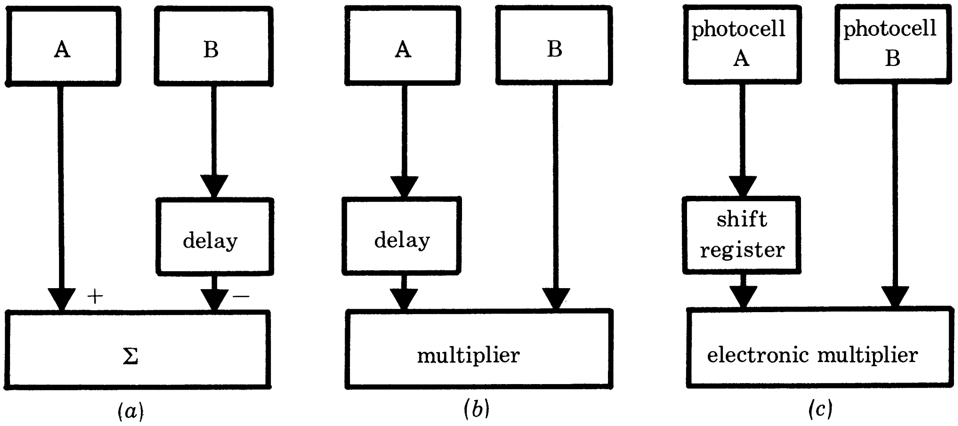

FIGURE 2. Models of physiological motion detectors. (*a*) After Barlow & Levick (1965). Receptors A and B sample adjacent retinal regions. B's output is delayed and subtracted from A's output. Device gives an output when stimulus moves from A to B. (*b*) After Reichardt (1961) (greatly simplified). As (*a*), but B's delayed output is multiplied into A's output, not subtracted from it. (*c*) Working model which is an electronic form of (*b*). Attempts to use this device as a model for human motion perception succeeded for perception of point-by-point random-dot motions (system 1), but failed for perception of moving texture edges or cyclopean edges (system 2).

passes first across B and then across A, with a stimulus transit time equal to B's internal delay time.

With R. Hansen, I constructed a simplified electronic version of Reichardt's model (see figure 2*c*). A randomly sectored black and white wheel rotated clockwise or anticlockwise past two photocells A and B, which were mounted side by side so that each contour moved past first one photocell and then the other. To sense motion from A to B, A's output was delayed for a fixed time interval by a shift register, and then multiplied into B's output by an electronic multiplier. To sense motion from B to A, B's output was delayed and multiplied by A's output by a second multiplier. The output voltage from one multiplier was subtracted from that of the other to give a final output voltage which deflected the needle of a centre-zero voltmeter. The meter needle deflected to the left for motion to the left (from B to A) and deflected to the right for motion to the right (from A to B). This device responded to velocity, not to position, and its output was zero whenever the stimulus wheel was stationary. Also, like Reichardt's model, it was tuned to a preferred velocity, giving the maximum output when the transit time of the stimulus from A to B was equal to the dealy time of the shift registers. The output voltage was reduced for velocities above or below the preferred velocity. It was arranged that the internal delay time of the shift registers could be manually switched to a longer time constant, which gave a slower optimal velocity. Further options could be built in if desired; for instance, it would be possible to make the delay times different for A and B, which would give optimal velocities which were asymetrical for the two directions A–B and B–A. Negative feedback might be used to adjust the internal delays on the model so that it would automatically match its optimum velocity to whatever stimulus velocity was applied to it over a period of time.

All of these models – optical, electronic and physiological – are variations on a theme. They all belong to a class of detectors that compare the luminance distributions seen at position A and time t_1 with that seen at position B and time t_2. The comparision is achieved by addition, subtraction or correlation, and despite minor differences between the models the performance of all the devices depends upon the *cross-correlation* between the patterns at (A, t_1) and (B, t_2).

3. APPLICABILITY OF THE MODEL: TWO SYSTEMS IN HUMAN MOTION PERCEPTION

The comparator model, which physiologists have found in retinae and which I have built in electronic form, is perfectly adequate as a model for the detection of apparent movement of a single spot such as Wertheimer studied. But with more complex stimuli such as pictures, the correspondence problem arises. Can a comparator solve the correspondence problem or do we need a new model?

There are two logically possible ways in which a visual system, natural or artificial, could see apparent movement, and they are illustrated in figure 3. Suppose that a human, or a machine, is watching a cine film of a clock with a sweep second hand. In the first frame the hand is vertical, and in the next frame it is rotated a few degrees clockwise. There are two quite different strategies by which he, she or it could detect the movement (Anstis 1970). The first strategy would be to detect each black point in the first picture, and see it as moving to the nearest black point in the second picture (figure 3*a*). This point-by-point strategy could be implemented by a population of the motion detectors which we have been describing. It requires only local comparisons between pictures, and demands no prior analysis of patterns within each picture. One important distinguishing feature of such a system is that the extraction of motion information from the pictures could *precede* the extraction of edges within each picture; indeed, the motion information might in principle be used to define those edges. An alternative strategy would be to segregate those black points within each picture that in some sense 'belong together' as a hand, and only after that to detect hand-to-hand motion between the two pictures (figure 3*b*). On this strategy, edge extraction would necessarily precede motion detection between pictures. It is clear that this process cannot be modelled by the simple luminance comparators described earlier, since they do not have the capacity to extract edges. Does the human visual system employ simply the first strategy, for which we have an acceptable working model, or does it employ – either instead or in addition – the second strategy, for which the model cannot give an account?

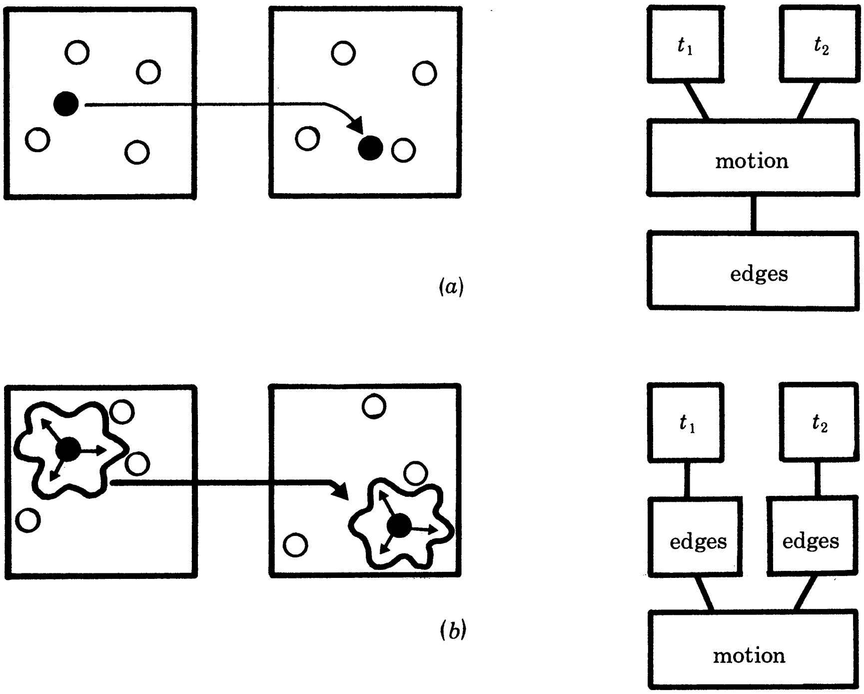

FIGURE 3. Two possible strategies for seeing apparent movement. Left: Two pictures exposed at times t_1, t_2. Though drawn side by side, pictures are really superimposed. Right: Hypothetical perceptual processes. (*a*) Visual system might compare pictures point-by-point (left), for instance, when viewing random-dot stereograms flashed in alternation. A.m. is seen, showing that motion between pictures is seen prior to any extraction of edges within each picture (right). This is system 1 (left column of table 1). (*b*) Visual system might extract edges (e.g. texture edges or cyclopean edges) to segregate zones or regions, and then look for motion between these regions. In this case, edge extraction must precede motion perception (right). This is system 2 (right column of table 1). (Anstis 1970, 1978.)

With most cine film frames or pairs of pictures, there is no way of knowing which strategy is used by the visual system. However, it has proved possible to manipulate the nature of the similarity between pairs of pictures in different ways so that the visual system would be forced to use one strategy in some cases and the other strategy in other cases, if it did indeed have alternative strategies available to it. Pictures such as random-dot stereograms, which when considered separately contain no visible edges, but which considered as a pair show high point-by-point correlations, will signal apparent movement to a detector array and also to a human observer, who must be using the first strategy. As we shall show, however, other pictures can be devised with more subtle correspondences, such as those in which each picture contains subjective contours but there is *zero* point-by-point correlation between them. Here a detector array would be blind to the apparent movement which would nevertheless be clearly visible to a human observer, who must, therefore, be using the second, global strategy. So the comparator device is a successful model for the first type of human motion perception, but not for the second. We conclude, from the demonstration described above and from others described later, that these alternative strategies are not merely logical possibilities; there is good evidence that both strategies are used by the visual systen to perceive motion. Table 1 lists the characteristics of the two supposed visual processes. The table is based on accounts by Braddick (1974), Ramachandran (1978) and Anstis (1978).

Table 1. Two systems for motion perception

system 1: peripheral apparent motion	system 2: cognitive apparent motion
can be modelled by comparator device	cannot be modelled by comparator device
short spatial range (up to 15′) (Braddick 1974)	long spatial range (up to tens of degrees)
based on simple point-by-point cross-correlations between pictures as in Julesz random-dot stereograms	subtler correspondences, e.g. cyclopean edges, texture edges, stereo edges
stimulates neural motion detectors (Grusser & Grusser-Cornehls 1973)	does not stimulate motion detectors, which cannot resolve cyclopean contours
adapts neural motion detectors to give motion after-effects	little or no motion after-effects
contrast reversal leads to reversal of a.m. (Anstis 1970)	contrast reversal does not affect perceived direction of motion
colour is not an adequate input for motion perception (Ramachandran & Gregory 1978)	
motion perception between pictures precedes edge extraction within pictures	edge extraction within pictures precedes motion perception between pictures

4. System 1

(*a*) *Julesz random-dot stereograms*

Julesz (1971) has devised random-dot stereo pairs, in which each picture consists of computer-generated random dot patterns looking not unlike photographs of sandpaper. The two pictures in such a stereo pair are identical except that one picture contains a square central region which is slightly shifted to the right. It is important to note that the central square is not visible, indeed does not logically exist, within either picture considered on its own. (The first picture is random and the second picture is obtained by modifying the first. This is like switching some of the columns around in a random number table: the result is merely a second random number table. In the same way, the second picture is also random, and the central square exists only as a correlation between the two pictures.) If the pictures are presented one to each eye in a stereoscope, then binocular fusion gives the percept of a central square floating in depth in front of the surround. If, instead, the pictures are now presented in alternation to one eye only, with the surround regions in register, then the central square is perceived as moving to and fro from left to right (Anstis 1970; Julesz 1971; Lappin & Bell 1972; Braddick 1974). This central square does not exist until *after* the two pictures have been compared and motion perceived, but the edges of the square are clearly seen and are defined by the motion. Clearly, the edge information defining the boundaries of the central square exists only as a relation between the two pictures; it is an emergent property of their juxtaposition and is not a property of either picture taken alone. This fact implies that the perception of the motion between the pictures must be preceding the extraction of edges within each picture (figure 3*a*).

(*b*) *Reversed apparent movement during contrast reversals*

Two identical pictures exposed in sequence, overlapping but with a small spatial shift between them, show apparent motion, which is of course in the direction of the physical displacement. Anstis (1970) and Anstis & Rogers (1975) reported a new effect: if a black and white slide is dissolved to its own photographic negative, overlapping but shifted out of perfect registration by a few arc minutes, then strong and compelling a.m. was reported toward the earlier stimulus, i.e. in the direction *opposite* to the physical displacement. This reversed apparent movement is probably caused by neural blurring or spatial summation which shifts the effective position of superimposed positive–negative contours (Rogers & Anstis 1975). Related effects occur if the positive and negative are held at fixed contrast levels, but the positive makes to and fro real movements, of a few arc minutes, through the in-register position over the stationary negative. Real movements of the positive are apparently enhanced if the negative is dim, but apparently reversed if the negative is bright. Anstis & Rogers (unpublished results) have found that these different effects can be combined into a repetitive four-stroke stimulus cycle, in which a positive picture is initially placed almost in register with an unchanging negative, and then jumps a few arc minutes to the left, grows dimmer, jumps back to the right, grows brighter again, and so on repetitively. This oscillating stimulus creates a compelling illusion of a pattern which drifts continually to the left without ever 'getting anywhere'. The effect is hard to describe but easy to see in films that we have made. The illusion is a form of 'aliasing': it can be shown that some of the low spatial frequency components of the pictures are in fact shifting continually to the left, even though the actual figure contours are not moving through more than a few arc minutes.

Reversed a.m. is probably a system 1 effect occurring early in the visual system. Blakemore & Anstis (unpublished) projected reversed a.m. dissolves on a screen in front of a cat while recording from some of its cortical motion detectors. The cortical units reported reversed motion whenever the human observers perceived it on the screen. In humans, reversed a.m. products motion after-effects in a direction appropriate to the illusory reversed movement, not to the physical displacement.

(*c*) *Colour not important for apparent movement*

In the preceding section, reversed apparent movement arose during contrast reversals, when the black parts of one picture were replaced by the white parts of a succeeding picture and vice versa. The situation is very different, however, during colour reversal, when coloured parts of one picture are replaced by complementary colours in a second picture. It is not satisfactory to use ordinary colour slides and their negatives, because such photographic negatives reverse both the colours and the brightness of their originals.

Anstis (1970) produced apparent movement between a Julesz random-dot pattern with green dots on a red ground which was followed by a second correlated pattern having red dots on a green ground. The luminances of all the colours could be adjusted independently. He found the the direction of the apparent movement was determined by luminance, not by colour. Thus, a red dot on a green ground followed by a geen dot on a red ground showed normal forward a.m. if both dots were lighter (or darker) than their respective backgrounds and they showed reversed a.m. if one dot was lighter than its background and the other dot was darker than its own background. If the dots in each pattern were matched in brightness to their respective backgrounds to give isoluminant coloured dots, then the apparent motion simply disappeared. This loss of apparent movement at isoluminance has been confirmed by Ramachandran & Gregory (1978). So colour does not provide a significant input to human motion perception. This is consistent with Zeki's (1977) finding that 'cells of the movement area [in the monkey cortex] are not concerned with colour'.

(*d*) *Motion after-effects from apparent movement*

The hypothesis that apparent movement in random-dot patterns is sensed by neural motion detectors is supported by the fact that inspection of such a.m. readily produces motion after-effects. Motion after-effects are almost certainly caused by adaptation of such neural units (Barlow & Hill 1963). Moreover, Braddick found that a.m. was not seen in such random-dot patterns if the spatial jumps exceeded 15′. He summarized evidence that 15′ was the maximum spatial range for 'short-range' (system 1) motion perception. It is true that a.m. can still be seen for larger jumps if the stimuli are isolated lines instead of random-dot patterns, but such a.m. no longer gives motion after-effects, so it may be mediated by a different mechanism (system 2). Banks & Kane (1972) found that adaptation to a.m. produced by collapsing circles gave motion after effects only if the spatial jumps were less than about 12.5′.

As we shall describe later, inspection of a.m. produced by shifting texture edges gives slight or no motion after-effects. We suggest that system 1 motion perception is mediated by neural motion detectors, but system 2 is not.

5. System 2

(*a*) *Texture edges and cyclopean edges*

The edges on which system 1 operates are luminance edges. But Ramachandran *et al.* (1973) successfully obtained apparent movement from edges that were defined not by luminance but by texture. The existence of such edges was demonstrated by Pickett (1970), who showed that it was easy to discriminate abutting zones of texture when one zone contained random black and white dots while the other contained random horizontal dashes. Both zones contained 50 % black and 50 % white points, so they had the same mean luminance and the same first order of probability.

Ramachandran *et al.* used textured patterns that looked deceptively similar to Julesz patterns, but whose structure was in fact crucially different from that of the classical random-dot stereogram. Ramachandran *et al.* used two pictures in which the surrounds contained *uncorrelated* 'noise' or random dots, but with a central square zone that was shifted *en bloc*, being displaced slightly to the right in one picture (figure 4). This square zone was filled with horizontal dashes in one picture, and with vertical dashes in the other. The dashed zones were *not* correlated with each other. When the two pictures were superimposed and exposed in alternation, subjects reported a central square jumping to and fro. Note that there was *no* point-by-point correlation between the two pictures, as there is in the conventional Julesz stereogram pair. The important inference to be drawn from this observation is that in this instance motion perception could *not* be based upon point-by-point matching of images exciting the retina in succession. Instead, the recognition of visual texture, and of a square contour bounding it, must have been visually processed *before* motion was perceived (see figure 3*b*). This is exactly the reverse of the condition described under (*a*) above, and is an example of motion perception by system 2 (see table 1).

Virtually any form of texture edge, provided it is clearly discriminable when at rest, will also give a system 2 apparent movement when abruptly shifted. Figure 5 shows examples of edges based on vernier offset or spatial phase, and on magnification of spatial frequency components Stereoscopic depth edges are also effective. Steinbach & Anstis (1976) generated moving stereo gratings by using dynamic random-dot visual noise. Each eye on its own saw a randomly twinkling snowstorm, like the noise on a detuned t.v. receiver, but when the two monocular views were binocularly fused in a stereoscope, a set of horizontal corrugations in depth were seen, and these bars were made to drift slowly downwards. The downward motion was clearly visible; it was found that the drifting bars could easily be tracked with smooth voluntary eye movements: there was no need for any real movement to brush across the retina to drive the eye movements. The display was not compelling enough to drive involuntary optokinetic nystagmus, but it did generate a very brief movement after-effect (see the next section).

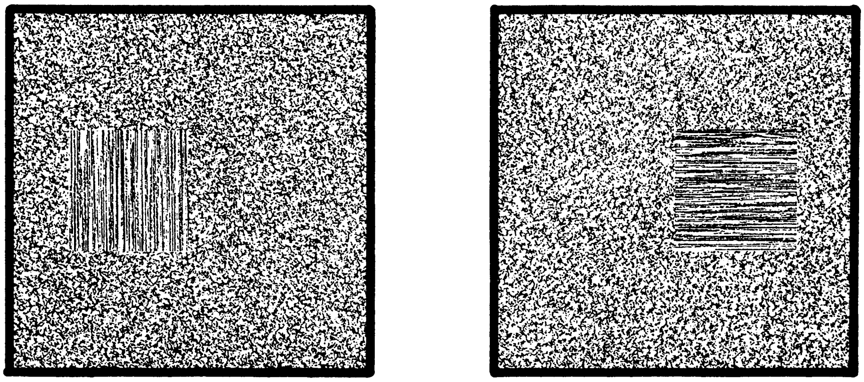

FIGURE 4. (*a*) A zone of horizontal (or vertical) dashes against a background of random dots can be clearly perceived as a texture boundary, even though both zones have the same space-averated luminance (Pickett 1970). (*b*) When the pictures are switched on and off in alternation, the visual system analyses the square contour within each picture, then sees motion between the squares (system 2). Note that there is no point-by-point correlation between the two pictures. (Ramachandran *et al.* 1973.)

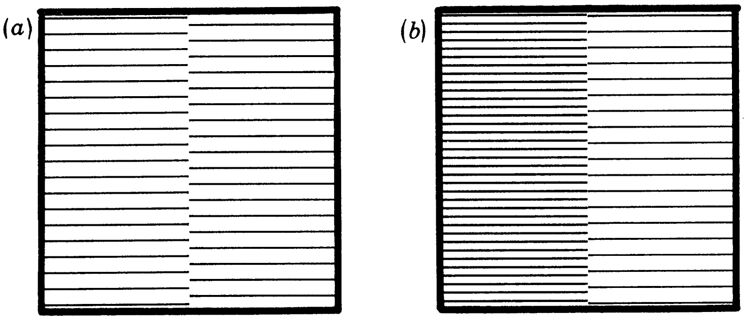

FIGURE 5. Examples of texture edges defined by vernier offset (spatial phase) and by spatial frequency.

What, then, is the nature of these 'texture edges' to which system 2 is sensitive? It would be misleading to call them 'subjective contours', because there is a physical difference on the two sides of the edge. 'Cyclopean edges' is a better term (Julesz 1971). The difference is not usually in the space-average luminance, but it can be in spatial frequency components, spatial phase, binocular disparity, etc. A difference in physical luminance must be translated into some difference in the response of a population of neurons, and so must difference in texture. Very little is known about the visual processing that extracts edges, but computer algorithms for edge extraction and zone segregation are well established (see Rosenfeld & Kak (1976), ch. 8). For instance, a zone of black and white random dots can easily be segregated from a mid-grey background by measuring the contour richness or 'digital gradient'. Presumably, analogous processing occurs in the visual system. With the displays described in this section, such processing must occur in system 2 *before* motion is perceived.

(*b*) *Little or no motion after-effects from cyclopean edges*

Whereas motion after-effects are easy to produce with short-range a.m. from luminance edges (system 1), they are hard to produce with moving cyclopean edges (sytem 2). Either they are very brief or they do not occur at all. Drifting stereo gratings gave motion after-effects (Papert 1964; Steinbach & Anstis 1976) but they lasted for only 1 s or less. The same was true for dichoptic motion produced by widely separate points (Anstis & Moulden 1970). All these motion after-effects must lie central to the point of binocular fusion, and they were all very brief. The drifting kinetic gratings which are described later (§7*b*) gave no motion after-effects, and there appear to be no reports of after-effects from shifting texture boundaries. It seems likely that neural motion detectors adapt readily, but the mechanisms which report the shift of 'computed' or cyclopean edges are stiff systems which show very little adaptation.

6. Patterns that can stimulate either system 1 or system 2

There are some pairs of patterns that give apparent movement in one direction when superimposed and exposed in alternation, but when the timing is slightly altered the apparent movement is radically changed in direction and perceptual organization. Pantle (1973) reported a.m. between two uncorrelated patterns, each containing a square cluster of red spots against a background of green sports (figure 6). The spots in each picture were in uncorrelated random positions. When the patterns were switched on and off alternately at about 0.5 Hz, subjects reported a cluster of red spots moving to and fro *in toto*. However, if one pattern dissolves into the other, with one pattern fading down as the other fades up, the apparent motion is entirely different. There is a random, incoherent movement, with each spot in one picture moving to the nearest spot in the other picture, independent of its colour. Instead of a disciplined square squad of spots jumping globally to and fro, a random rabble of spots swim about locally in all directions. This local incoherent motion between nearest neighbours is signalled by system 1, operating on a point-by-point comparison. The perceived motion of a cluster of elements is signalled by system 2, since it uses the global information about the property (redness) distributed across a number of spots to extract a subjective square contour, which is then seen to move. Anstis (1970) exposed in alternation two gratings that were tilted a few degrees away from vertical, one clockwise and one anticlockwise (figure 7). When they were switched on and off alternatively at about 0.5 Hz, subjects reported a single striped field oscillating to and fro in a rotary manner. But when the alternation rate was raised to 5–10 Hz, or if the fields dissolved into each other instead of switching, then the perceived motion broke up into local, relatively incoherent motion, with horizontal strips of the field rotating independently, and shearing off from other strips along the lines of the moiré fringes. Again the local motion can be attributed to system 1, and the global motion of the whole field to system 2.

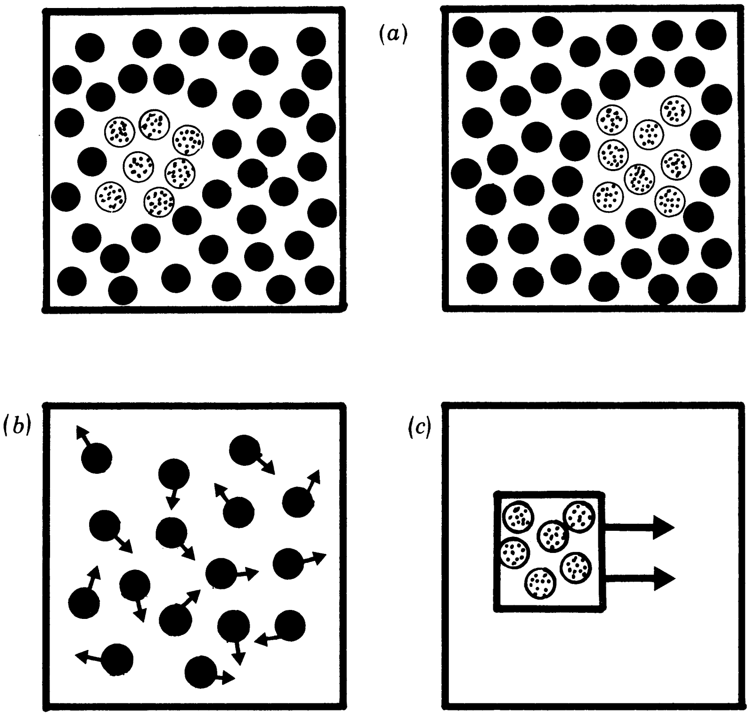

Figure 6. (*a*) Each picture contains a square cluster of red spots against a background of green spots (Pantle 1973). (*b*) With the two pictures superimposed, during slow dissolves, incoherent local motion is seen, with each spot moving to its nearest neighbour irrespective of colour (system 1). (*c*) If the pictures are switched sharply on and off in alternation, the visual system extracts a subjective outline of a square in each pattern and then (as in figure 4) sees motion between two squares (system 2).

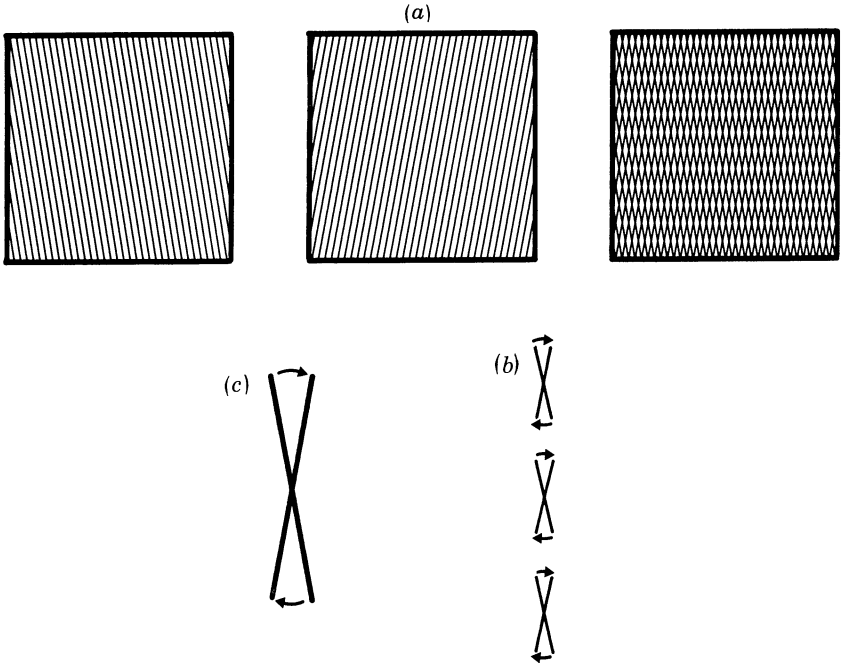

Figure 7. (*a*) Two superimposed gratings give horizontal moiré fringes when both gratings are on continuously. (*b*) During dissolves from one grating to the other the picture appears to split up along the moiré fringes into horizontal strips. Each strip shows local rocking movement (system 1). (*c*) When the gratings are sharply switched on and off in alternation, a single rocking grating is seen (system 2). (Anstis 1970.)

V. Zemon has been independently studying this phenomenon of the crossed gratings (personal communication).

7. Patterns that do stimulate both system 1 and system 2

We have seen that short-range luminance edges are handled by system 1, and most other kinds of edge – cyclopean and the like – are handled by system 2. But there are two special kinds of edge, flicker edges and kinetic edges, that seem to stimulate both systems. Flicker edges stimulate systems 1 and 2 in parallel, whereas kinetic edges stimulate them in series.

(*a*) *Flicker edges*

Consider a vertical edge that alternates in polarity between black–white and white–black at a rate of about 5–10 Hz. Suppose the edge makes a succession of small jumps to the right, each jump being 5–10′ in extent, with a polarity change synchronized to each jump. Thus a black–white edge is replaced by a white–black edge shifted to the right, which in turn is replaced by a black–white edge shifted further to the right, and so on. After an excursion of a degree or so, the edge jumps back to the left and the cycle repeats.

I have found (Anstis 1980) that observers unambiguously perceived a flickering edge moving (somewhat jerkily) to the right. This movement to the right was easy to see, and easy to track with the eyes. If the observers fixated on a stationary point and adapted to the rightward motion of 30–60 s, then they reported a motion after-effect afterwards, as might have been expected. What was quite unexpected, however, was that the direction of the after-effect was to the *right*, i.e. in the *same* direction as the stimulus motion, not in the opposite direction as one would normally expect.

We interpret this as follows. The jumps to the right were picked up by system 2, irrespective of the sign or polarity of the edge which reversed on each jump. Stimulation of system 2 gives a percept of motion, but no after-effect of motion. At the same time, system 1 was also being stimulated: note that the jumps were only 5–10′ in amplitude, well within the spatial range of system 1. But to system 1 each jump produces a *reversed* apparent movement, apparently to the left, as described earlier. It is the adaptation to the apparent leftwards motion that gives rise to a motion after-effect to the right. An interesting feature is that the leftward reversed apparent movement is never consciously perceived in this situation; the edge is perceived as jumping to the right, not to the left. So this is an example of a perceptually invisible motion giving rise to a visible after-effect. We know of no earlier reports of this example, although visible coloured after-images from sub-threshold coloured stimuli have been reported (see review by Anstis *et al.* 1978).

(*b*) *Kinetic edges and wave motion*

Imagine that the fingers of two hands are randomly speckled with luminous paint; suppose the fingers are interlaced and held vertically with one hand higher than the other. In a dark room one would be unable to discern the fingers: one would simply see a random speckle of luminous dots. But suppose the fingers were then slid vertically together and apart: one would immediately be able to distinguish the separate fingers by virtue of their vertical motion or ‘common fate’. Thus when the fingers of the left hand move upwards and those of the right hand downwards, the (stationary) vertical edges between the moving fingers would be readily visible as a kind of standing wave.

Now suppose that the fingers are unlaced and the fingers of one hand are drummed on the table top. Each finger moves only up and down as before, never sideways. But a beetle on the table would perceive a thunderous creeping barrage of fingertips apparently advancing horizontally across the table. What moves horizontally is not, of course, the fingers themselves, but a *sequence* of fingers or a phase velocity, in other words a travelling transverse wave.

According to the approach outlined here, the sequence of visual events is as follows. The up and down movements of the fingers are picked up by system 1. The common motions seen by system 1 serve to segregate the edges of the fingers. Since these edges are defined by motion information without any luminance cues, we shall call these 'kinetic edges'. When the moving fingers are interlaced they define stationary kinetic edges, but when they are drumming they define travelling kinetic edges, which move at right angles to the fingers' motions. The displacements of the kinetic edges, i.e. the transverse waves, are picked up by system 2. We have produced travelling kinetic edges and waves both mechanically (Steinbach & Anstis 1976) and electronically (K. Nakayama, C. W. Tyler & S. M. Anstis, unpublished).

The mechanical display consisted of 50 long thin metal rods held in a frame and arranged in a vertical sheet somewhat like the teeth of a giant comb. Each rod was free to move independently of its neighbours along its own length but in no other direction. The bottom of each rod rested on a helicoidal barley-sugar twist Jacobean table leg, which lay with its long axis horizontal at the bottom of the sheet of rods. When this table leg was rotated about its own axis by means of a handle, it imparted a sinusoidal up and down motion to each rod; moreover, each rod had a slight phase lead over its right-hand neighbour. The resulting display was like 50 long thin metal fingers drumming on a table top. Each rod was randomly banded in black and white so that its vertical motions could be clearly seen. The travelling waves moving to the right defined a drifting kinetic grating. The wave motion to the right was easily perceptible, and provided an adequate stimulus for smooth tracking eye movements (Steinbach & Anstis 1976). If the eyes fixated a stationary point on the frame while an observer adapted to 60 s of wave motion to the right, then *no* motion after-effect could be seen thereafter.

In the electronic display, a static pattern of about 200 × 200 static random dots was set up on a television raster. Then, by means of a sinusoidal signal applied to the X (horizontal) inputs, it was arranged that each horizontal line of the raster moved horizontally to and fro, slightly lagging the line above it. All of the dots on the screen were moving to and fro horizontally, but the percept was a series of sinusoidal waves or ripples running downwards across the screen; in other words, a drifting kinetic grating. The stimulus was like the mechanical display just described laid on its side. Once again, the downward drift of the kinetic grating was easy to perceive, even though no dots or luminance borders were moving downwards. The downward motion gave an adequate stimulus for smooth pursuit eye movements, but it could not be used to generate a motion after-effect.

We conclude that system 1 responds to the horizontal motion of the dots, and that system 2, operating at a later level in the visual system, responds to the wave motion.

REFERENCES (Anstis)

Anstis, S. M. 1970 Phi movement as a subtraction process. *Vision Res.* **10**, 1411–1430.
Anstis, S. M. 1978 Apparent movement. In *Handbook of sensory physiology* (ed. R. Held, H. W. Leibowitz & H.-L. Teuber), pp. 655–673. New York: Springer-Verlag.
Anstis, S. M. 1980 Direction of apparent movement of a moving flicker edge. (In preparation.)

Anstis, S. M. & Moulden, B. P. 1970 Aftereffect of seen movement: evidence for peripheral and central components. *Q. Jl exp. Psychol.* **22**, 222–229.

Anstis, S. M. & Rogers, B. J. 1975 Illusory reversal of visual depth and movement during changes of contrast. *Vision Res.* **15**, 957–961.

Anstis, S. M., Rogers, B. J. & Henry, J. 1978 Interactions between simultaneous contrast and coloured afterimages. *Vision Res.* **18**, 899–911.

Banks, W. P. & Kane, D. A. 1972 Discontinuity of seen motion reduces the visual motion aftereffect. *Percept. Psychophys.* **12**, 69–72.

Barlow, H. B. & Hill, R. M. 1963 Evidence for a physiological explanation of the waterfall phenomenon. *Nature, Lond.* **200**, 1345–1347.

Barlow, H. B. & Levick, W. R. 1965 The mechanism of directionally selective units in rabbit's retina. *J. Physiol., Lond.* **178**, 477–504.

Braddick, O. 1974 A short range process in apparent motion. *Vision Res.* **14**, 519–528.

Grüsser, O. J. & Grüsser-Cornehls, U. 1973 Neuronal mechanisms of visual movement perception and some psychophysical and behavioral correlations. In *Handbook of sensory physiology*, vol. 7 (3*a*): *Central processing of information* (ed. R. Jung), pp. 333–430.

Julesz, B. 1971 *Foundations of cyclopean perception.* Chicago: University of Chicago Press.

Korte, A. 1915 Kinematoscopische Untersuchungen. *Z. Psychol.* **72**, 193–206.

Lappin, J. S. & Bell, H. H. 1972 Perceptual differentiation of sequential visual patterns. *Percept. Psychophys.* **12**, 129–134.

Mackay, R. S. 1959 Display of moving parts of a scene. *Science, N.Y.* **130**, 223–224.

Pantle, A. J. 1973 Stroboscopic movement based upon global information in successively presented visual patterns. *J. opt. Soc. Am.* **63**, 1280A.

Papert, S. 1964 *M.I.T. Quarterly Progress Report.*

Pickett, R. M. 1970 Visual analyses of texture in the detection and recognition of objects. In *Picture processing and psychopictorics* (ed. B. S. Lipkin & A. Rosenfeld). New York: Academic Press.

Ramachandran, V. S. 1978 Studies in binocular vision. Ph.D. thesis, University of Cambridge.

Ramachandran, V. S. & Gregory, R. L. 1978 Does colour provide an input to human motion perception? *Nature, Lond.* **275**, 55–56.

Ramachandran, V. S., Rao, V. M. & Vidyasagar, T. R. 1973 Apparent movement with subjective contours. *Vision Res.* **13**, 1399–1401.

Reichardt, W. 1961 Autocorrelation: a principle in motion perception. In *Sensory communication* (ed. W. Rosenblith), pp. 303–317. New York: Wiley.

Rogers, B. J. & Anstis, S. M. 1975 Reversed apparent depth from positive and negative stereograms. *Perception* **4**, 193–201.

Rosenfeld, A. & Kak, A. 1976 *Digital picture processing.* New York: Academic Press.

Steinbach, M. J. & Anstis, S. M. 1976 Paper presented to the Annual Conference of the Association for Research in Vision and Ophthalmology, Sarasota, Florida.

Walls, G. 1942 *The vertebrate eye and its adaptive radiation.* Michigan: Cranbrook Press.

Wertheimer, M. 1912 Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.* **61**, 161–278. (Excerpted and transl. in *Classics in psychology* (ed. T. Shipley). New York: Philosophical Library 1961.)

Zeki, S. M. 1977 Colour coding in the superior temporal sulcus of rhesus monkey visual cortex. *Proc. R. Soc. Lond.* B **197**, 195–223.

Discussion

D. M. MacKay (*Department of Communication and Neuroscience, University of Keele, Keele, Staffordshire ST5 5BG, U.K.*). I should like to append some further evidence indicating that we must expect motion to be represented in more than one neural subsystem.

In the first place, there is an additional mode of motion perception, not so far mentioned, in which the retinal image is not displaced at all. During 'smooth pursuit' of a moving target, for example, there is a compelling perception of motion even when neither of the two subsystems mentioned by Dr Braddick is activated.

Secondly, the existence of separate subsystems for signalling *continuous drift* and *discontinuous displacement* is suggested by the fact that, for small enough displacements of the retinal image, a continuously illuminated target in a stroboscopically lit surround can be seen to move while the surround appears to remain at rest (MacKay 1976) came across a striking illusion that seems to show the possibility of rivalry between these two subsystems. The subject views an electronically generated field of dynamic visual noise with a movable rectangular 'window' (electronically) cut out of it so as to reveal a stationary under-layer of random texture. When the window is moved smoothly across the dynamic noise field, it seems to move in a succession of jerks which increase in size with eccentricity. The reason seems to be that the texture seen through the moving window is at rest relative to the fixated frame, and therefore gives rise to signals indicating zero drift, whereas the moving edge of the window itself generates signals indicating displacement. The conflict between the two sets of signals issues in a jerky succession of revisions of the perceived location of the window.

When we come to look for physiological candidates for these various functions the position becomes still more complex. My colleague P. Hammond and I have recently been using the same electronic display to generate textured stimuli for investigating the response characteristics of simple and complex cells in area 17 of cat (Hammond & MacKay 1975, 1977; Groos *et al.* 1976). As Gibson (1951) has particularly emphasized, textured bars camouflaged against similarly textured backgrounds become clearly perceptible when moving over the background. We have found no simple cells, however, which respond to such moving stimuli. Somewhat to our surprise, the cells that do respond are complex (presumably receiving an input independent of simple cells); but even these show no specific sensitivity to the shape or orientation of the textured bars. To make matters worse, the polar diagrams of responsiveness to a moving textured field generally have maxima in different directions from those for a moving black or white bar (Groos *et al.* 1976). Thus although the firing of such complex cells may serve to indicate motion, it cannot of itself represent either the shape or the velocity vector of the moving stimulus.

More recently, we have found that although the simple cells do not respond to texture motion, they are not uninfluenced by it (Hammond & MacKay 1978). A black bar on a textured background generally evokes a weaker response if the background moves with it. When only a small exploratory patch of texture is moved synchronously with the bar, its effect varies with its location relative to the conventionally defined receptive field, and can be significant far outside its boundaries. Along the axis parallel to the preferred bar orientation, the effect can in some cases vary from suppression (at the field centre) to facilitation (well outside the receptive field).

Even at this relatively peripheral level, then, it would seem that motion finds a diversity of representations in terms of neural activity, and that none of these is adequate to signify all forms of commonly perceived motion, let alone those generated in illusions. Certainly any hope of identifying one class of cell as *the* 'detector of motion' now seems illusory.

References

Groos, G. A., Hammond, P. & MacKay, D. M. 1976 Polar responsiveness of complex cells in cat striate cortex to motion of bars and of textured patterns. *J. Physiol., Lond.* **260**, 47*P*–48*P*.

Gibson, J. J. 1951 *The perception of the visual world.* Boston: Houghton Mifflin.

Hammond, P. & MacKay, D. M. 1975 Differential responses of cat visual cortical cells to textural stimuli. *Expl Brain Res.* **22**, 427–430.

Hammond, P. & MacKay, D. M. 1977 Differential responsiveness of simple and complex cells in cat striate cortex to visual texture. *Expl Brain Res.* **30**, 275–269.

Hammond, P. & MacKay, D. M. 1978 Modulation of simple cell activity in cat by moving textured backgrounds. *J. Physiol., Lond.* **284**, 117*P*.

MacKay, D. M. 1976 Perceptual conflict between visual motion and change of location. *Vision Res.* **16**, 557–558.

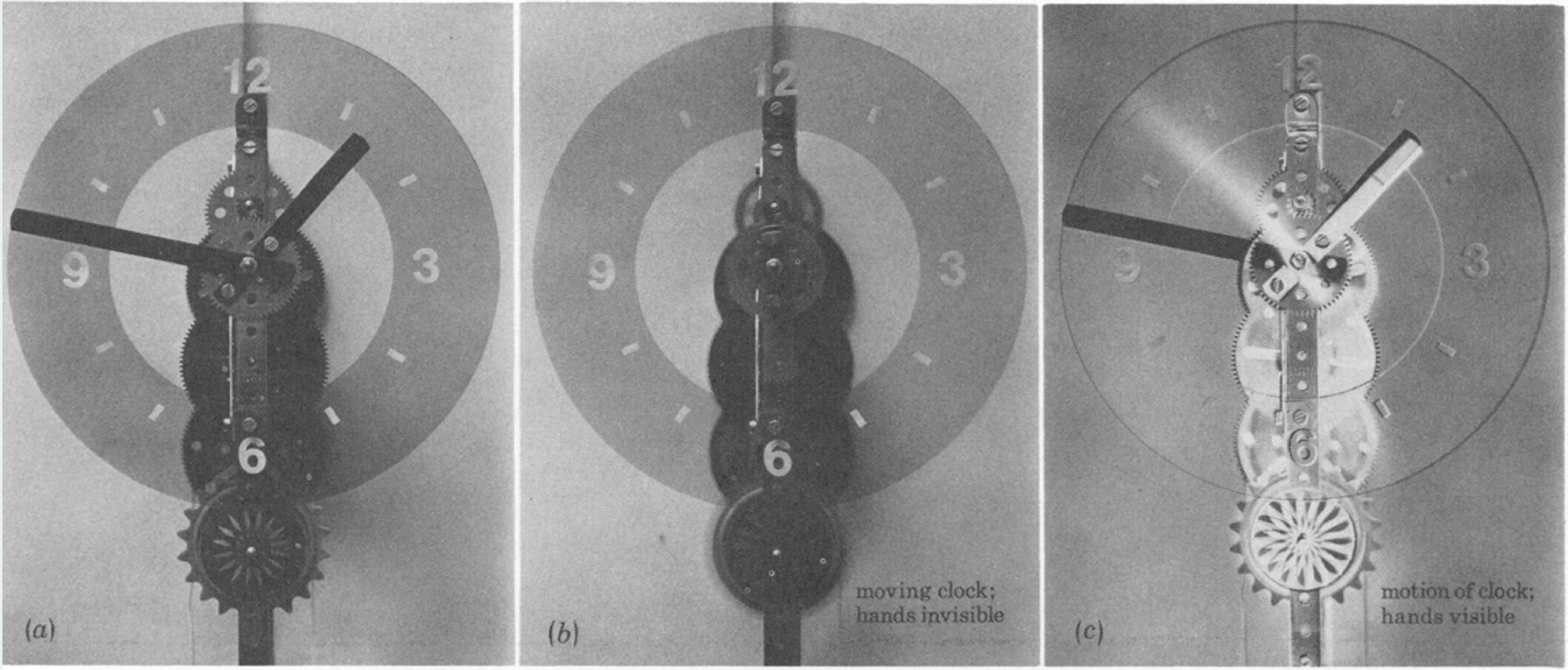

FIGURE 1. Simple model of change (or motion) detector. (*a*) Original photograph; (*b*) summed picture: time exposure shows static parts but moving parts blur out and disappear; (*c*) difference picture made by superimposing negative picture on a positive picture taken earlier: parts that have moved or changed show up clearly (black or white), but static parts cancel out and disappear.