The knowledge base of the oculomotor system

MICHAEL F. LAND AND SOPHIE FURNEAUX

Sussex Centre for Neuroscience, School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK

SUMMARY

In everyday life, eye movements enable the eyes to gather the information required for motor actions. They are thus proactive, anticipating actions rather than just responding to stimuli. This means that the oculomotor system needs to know where to look and what to look for. Using examples from table tennis, driving and music reading we show that the information the eye movement system requires is very varied in origin and highly task specific, and it is suggested that the control program or *schema* for a particular action must include directions for the oculomotor and visual processing systems. In many activities (reading text and music, typing, steering) processed information is held in a memory buffer for a period of about a second. This permits a match between the discontinuous input from the eyes and continuous motor output, and in particular allows the eyes to be involved in more than one task.

1. INTRODUCTION

Our eyes do not just passively record the scene ahead, but actively seek information from it, and in doing so change their direction of regard up to three times a second. Over the last decade many scientists and engineers involved in machine vision have recognized that there are advantages to be gained from having a moving 'inquisitive' sensor, rather than a fixed camera, and the field of 'active vision' has developed to explore the benefits of systems where the observer—man or machine—interacts with the environment in a purposeful and opportunistic way (Aloimonos 1993; Blake & Yuille 1993). For this to be possible the system responsible for controlling this interaction, the oculomotor system in our case, must act in an informed way—it must have its own knowledge base.

This knowledge is of two kinds: the system that moves the sensor needs to know where in the scene to direct it to find the information required for a particular action, and the sensor's visual nervous system, or electronic equivalent, needs to know what information to extract and pass on. Thus, when checking a road junction the oculomotor system must direct gaze to the appropriate side road, and the visual system must then record the presence or absence of vehicles and form an estimate of their speeds. When reading music, a pianist's gaze must move from note group to note group in an efficient way, and the visual system must extract the pitch, length and timing of each note. It is obvious from these examples that there can be no 'general purpose' strategy for directing gaze. Every visual task has an eye movement pattern that goes with it, and this is likely to be just as true of active machine vision, as it is for human vision.

At present, our knowledge of the way that people use their eyes in real tasks is poorly developed. There is, however, a huge literature on the nature of the eye movements recorded during simplified laboratory tasks, which usually involve locating and tracking small spots of light in the dark. These studies have produced very detailed accounts of the components that go to make up eye movement strategies: the fast saccades that move gaze from one point to another; the stabilizing movements (vestibulo-ocular and optokinetic reflexes) which ensure that the eye does not move during the fixations between saccades, in spite of movements of head and body; the smooth pursuit system that allows fixation to be maintained on a moving target; and the vergence movements which keep both eyes aligned on objects moving in depth (Carpenter 1988, 1992). Physiologists, however, have rarely interested themselves in the patterns of sampling and search that these components make possible (Viviani 1990). A celebrated exception to this was the work of Yarbus (1967) who showed that the patterns of fixations one makes when studying a picture depend both on the features in the picture ('bottom-up processing') and on the questions one is trying to answer from the information contained in the picture ('top-down processing'). Yarbus' examples made it clear that our use of eye movements involves a complex back-and-forth relationship between the parts of the brain that ask questions, those that analyse the incoming visual information, and those that look after the machinery that generates the eye movements themselves.

This paper is about the eye movement patterns that accompany actions, and the roles they play in guiding and executing those actions. A key point that emerges from a variety of studies including reading aloud (Buswell 1920; Geyer 1969) and driving (Land 1996), is that eye movements typically precede motor actions, by about a second in the examples quoted. Like an advance patrol, the eye movement system is out gathering useful

Phil. Trans. R. Soc. Lond. B (1997) **352**, 1231–1239
Printed in Great Britain

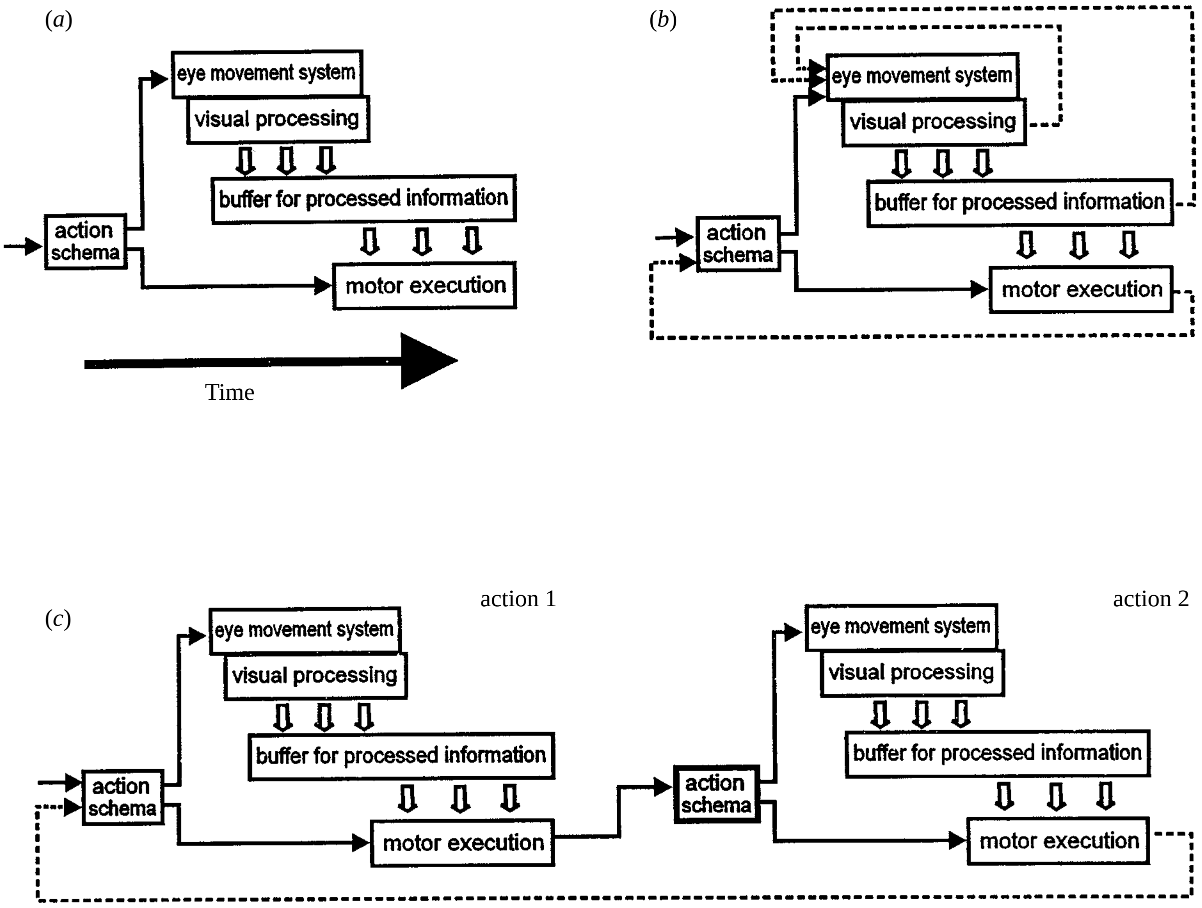

Figure 1. Relations between eye movements and motor acts. (*a*) Basic structure: a *schema* of instructions directs both eye movements and motor action, with the eyes supplying processed information to the motor system via a buffer. (*b*) Self-sustaining activities, such as text and music reading, may re-activate the eye movement system via several different types of loop (dashed lines). (*c*) Many activities form a chain in which one triggers the next. The dashed line indicates that actions may be recursive, as in an assembly task or ball game.

intelligence for the performance of whatever motor strategy is in hand. Thus, in contrast to the impression one gets from laboratory studies where eye movements are treated as 'responses' to 'stimuli', most eye movements in the real world are proactive in nature, not reactive. To study these eye movement patterns it is necessary to take a more naturalistic approach than is customary, and record the eye movements of freely moving people engaged in different activities. Techniques for making these recordings have been available since the 1950s (Thomas 1968), but in the last few years they have become more reliable and less obtrusive, thanks to video technology, and the recordings in this paper (figures 2–4) were made with such a device (Land 1993). They have not been widely used, however, and for most types of action we are ignorant of the nature and role of the accompanying eye movements. The one great exception to this is the study of reading (O'Regan 1990), where research has been driven by a variety of interested parties, including educational psychologists and psycholinguists as well as oculomotor physiologists. Reading, however, is a very specialized activity, and the lessons to be learned from it are not likely to be universally applicable.

In the next section we consider from a general point of view the ways that eye movements are involved in different tasks, Following that, we describe three 'natural' activities that we have studied ourselves—playing table tennis, driving and reading music—which provide varied examples of eye movement strategies. Finally we use these examples, and others in the literature, to look for common features in the different patterns. We will also discuss whether the human 'roving eye' way of doing things is applicable to artifical seeing systems.

2. EYE MOVEMENT STRATEGIES

In many tasks the first visible action is usually an eye movement to the location of some feature of importance in the task's execution. Having found it the eyes fixate it for a fraction of a second whilst the visual system extracts information needed in the performance of the task. This information is passed to the motor system via a memory buffer where it can be held briefly before it is required in the execution of the particular motor act. This idea of an action 'module' is illustrated in figure 1*a*, which reflects the ideas of a number of authors (e.g. Fleischner 1986; Kinsler & Carpenter 1995). It is assumed in this model that outline instructions for both the motor act *and the appropriate eye movements* are supplied by an action 'schema', as suggested by Norman & Shallice (1986), and that matters such as the avoidance of mutually incompatible activities have already been dealt with (Heuer 1996).

Figure 1*a* on its own could apply to 'one-off' actions such as catching a ball or switching off a light. Many actions, however, involve the continuous uptake of information, and so must incorporate some kind of self-sustaining loop, as in figure 1*b*. Examples are reading aloud, sight-reading music and driving on a winding road. What it is that keeps the loop running is not always clear; in reading it may be that the schema itself provides a repetitive trigger, but it is also true that the duration of fixations is influenced by the amount of processing involved (O'Regan 1990), and in music playing mistakes in the motor execution can also act as 'interrupts'. The state of the buffer may also control the fixation rate (Fleischer 1986). These different possible sources of feedback are indicated on figure 1*b*.

Many action sequences are best described by a 'chain' of modules, as in figure 1*c*. Here one action leads to and triggers the next, with a different eye movement strategy associated with each. Most games are of this nature, as are assembly tasks, city driving and arguably most of life! Many multicomponent tasks, such as playing tennis or laying bricks, repeat after a succession of different actions (dashed line). A good example of a recursive two-component task, as shown in figure 1*c*, is the block matching task devised by Ballard *et al.* (1992). This consisted of picking up and moving blocks to build a particular pattern, and it involves the eyes in identifying and selecting appropriate blocks and guiding the hand to the proper place in the pattern. Each subtask (select block and place block) lasted about 0.9 s, with the relevant eye movements preceding the final motor act (lift or drop) by about 0.7 s. Copy typing, described by Fleischer (1986), is a particularly interesting recursive task because the typing itself may be continuous (as in figure 1*b*), whilst gaze alternates roughly every second between reading the page to be copied and checking the copy on the screen (as figure 1*c*). It is the presence of a buffer, replenished during the reading periods, that frees time for the text to be checked.

3. EXAMPLES

In the three examples that follow we shall be looking particularly for answers to the two main questions related to the knowledge base of the eye movement system, namely: (i) what information does the oculomotor system require in order to locate the information needed for the action? and (ii) what information is extracted by the eyes during the fixations involved in viewing the region of interest?

Many other questions are also suggested by point (i). What determines what action is adopted in the first place? What terminates each fixation? How much is stored in the buffer? For how long? What happens during 'multitasking' when two or more activities occur together? We will comment on some of these in the conclusion.

(*a*) *Table tennis*

In playing table tennis and other ball games the secret is to know what the ball is going to do as far ahead as possible, in order to allow time for the planning and execution of accurate responses. Simply tracking the ball is not necessarily the best or fastest way of obtaining advance information about its movements, and there is evidence that the eyes are used in more subtle ways. For example, a study of members of the French national table tennis team found that the eyes anticipate the position of the ball before it bounces, making saccades to positions where there is, as yet, no visible stimulus (Ripoll *et al.* 1987). We have found that very ordinary amateur players do the same thing (figure 2*a*, *b*).

Figure 2*b* shows the vertical movements of the gaze direction of a table tennis player and of the ball, both relative to the top of the net. It is clear that the trajectories of ball and eye are similar—the player does generally keep his 'eye on the ball'. However, he frequently anticipates the ball's movement. This is most obvious after the opponent has played the ball and before its first bounce (arrows on figure 2*b*, saccade 4–5 on figure 2*a*). In all three cases the eye reaches the bounce point well before the ball, and the saccade that takes the eye there began between 180 and 220 ms before the bounce. The end points of the saccades are all different and related to the elevations of the actual bounce points, so these are not just stereotyped responses (figure 2*c*). The bounce on the far side of the net that follows the player's strike is also preceded by an anticipatory saccade. This leads the bounce by as much as 400 ms in one case (after 4.1 s), and reflects perhaps the player's expectation of the outcome of his stroke. Other saccades are not anticipated, however. After the ball has bounced, on either side of the net, the subsequent upward saccade is not made until the ball is 2–5° ahead of the fovea, so these are probably examples of ordinary reflexive following.

What information allows the player to make the anticipatory pre-bounce saccade? Its size and direction must be based on some aspects of what was seen in the 200–500 ms period between the opponent's strike, and the saccade itself. We analysed 15 approaching trajectories, as seen by the player, for relevant clues. The timing of the saccade was closely linked to the highest point of the ball's trajectory, which it followed by 174 ± 30 (s.d.) ms. The strike itself was not the timing signal, as the interval between strike and saccade was more variable (415 ± 55 ms). Given that the saccade is initiated by the trajectory apex, it seems likely that the direction and size of the saccade are determined during the subsequent 200 ms pre-saccade period. We found that saccade direction correlated very closely with the ball's direction (in the xy plane; θ in figure 2*d*) during the second half of this period ($r = 0.97$, $p < 0.001$), and we assume that this is indeed the controlling variable. It was more difficult to establish what determined saccade amplitude. An analysis found that only two features of the trajectory produced correlations with amplitude that reached even borderline significance: the height of the ball above the net (h) and its seen velocity (v). The correlation coefficients were weak (0.43 and 0.36 respectively), but the two variables were independent of each other, and their effects added so that together they gave a correlation with saccade size of 0.69 ($p < 0.005$). However, this still left about 70% of

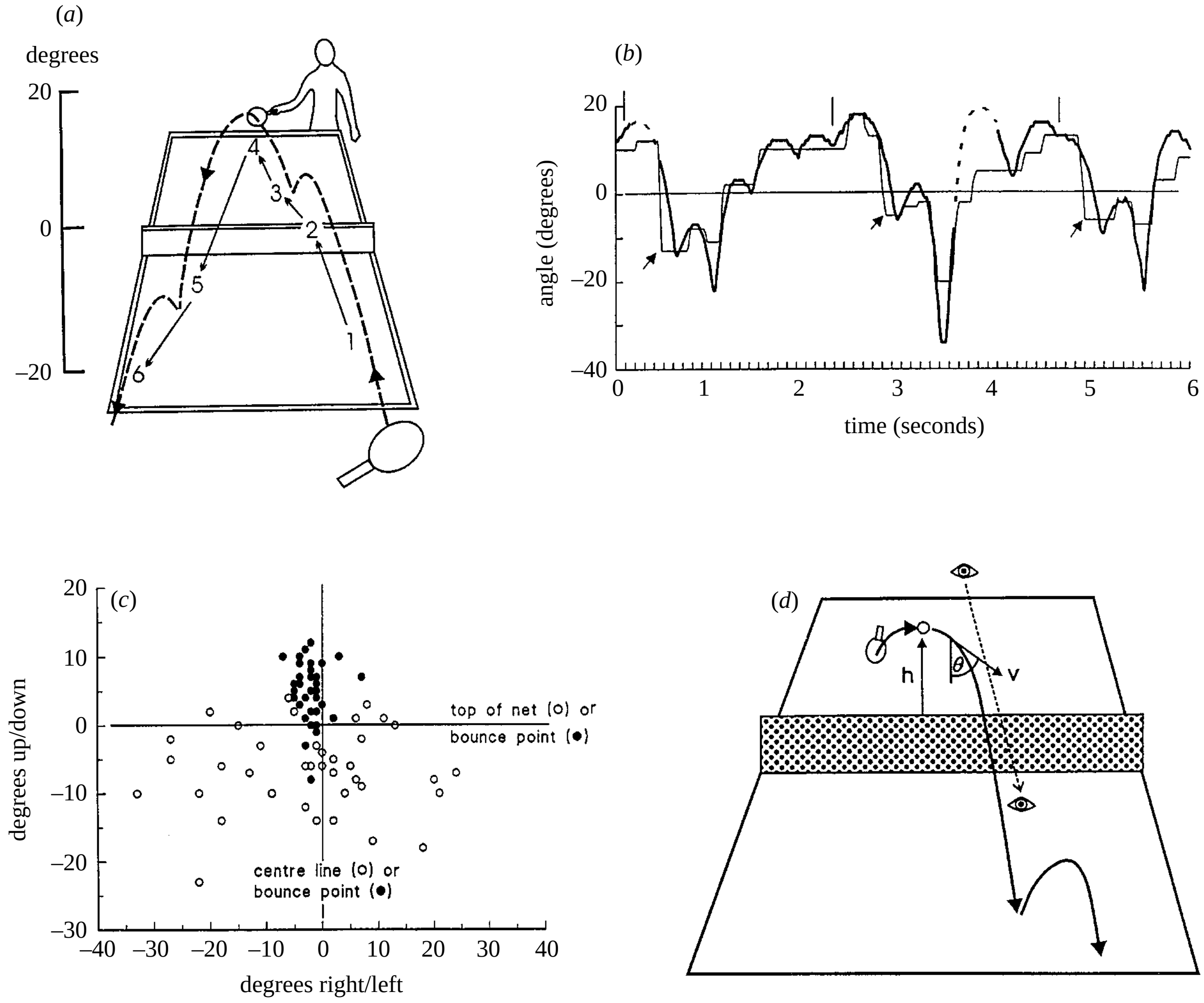

Figure 2. Eye movements in table tennis. (*a*) Typical pattern of fixations (1–6, thin arrows) during a single shot and return. Dashed line: ball trajectory. The vertical scale is in eye coordinates, centred on top of the net. (*b*) Record of vertical gaze movements during three shot/return pairs. Ordinate in eye coordinates relative to top of net. Thick line—ball; thin line—gaze. Short vertical lines: opponent's strike; deep V's: observer/player's strike. Arrows indicate the saccades (4–5 in (*a*)) that anticipate the bounce of the opponent's shot by about 0.2 s. Dashes: ball out of view. (*c*) Positions where anticipatory saccades landed relative to the table (open circles) and the bounce point (solid dots). The saccades are distributed all over the table, but are clearly aimed just above the bounce point. (*d*) Features of the ball's trajectory involved in aiming the anticipatory saccade. h, height of ball above net at the trajectory zenith; θ angle and v velocity of ball along trajectory, just before the saccade is made. The 'eyes' and dotted line indicate a typical anticipatory saccadic gaze shift.

the variance unexplained, which we reluctantly attribute to 'guesswork' by the oculomotor system.

This result is not as unsatisfactory as it seems. It turns out that no aspect of the ball's trajectory, as seen by the player, predicts the bounce point at all well. The problem is that the trajectory angle and ball velocity are confounded in the end-on view of the flight path, which prevents accurate prediction. The two variables on which the saccadic system has placed partial reliance probably provide the best information available, but even so their joint correlation with bounce point is as weak as it is with saccadic end-point. In other words, the eye movement system seems to be doing as well as it can, given the poor available evidence. In spite of this random element in saccade amplitude, the saccades do achieve their apparent purpose: they position fixation at a point whose vertical position above the bounce point is rather variable (mean $+ 5.2°$, s.d. $4.2°$, $n = 39$), but whose lateral position is quite accurate (mean 2.0° left of bounce point, s.d. 2.4°) (figure 2*c*).

Why is this saccade anticipatory? The very fact that the early part of the ball's flight does not predict how far down the table the ball will bounce, and hence the trajectory off the table to the player, suggests the reason. It seems that the eye gets itself to as good a position as it can to observe the bounce, allowing the player to determine the final part of the ball's trajectory, and in particular where it will be 400 ms later when the bat makes contact. Exactly what measurements are made at that time are not at present obvious, but a combination of the flatness of the approach to the bounce and the bounce position itself, probably provide all that is needed, when combined with an estimate of the ball's lateral velocity made before the pre-bounce saccade.

This example provides a nice instance of a saccade whose end-point is determined by identifiable information that is not simply the 'stimulus location'. As their game improves, novice players somehow acquire the rules that enable their eye movement systems to make these anticipatory saccades, and also what to do with the information that the resulting fixations supply. Players are unaware of any of this, yet both kinds of information are crucial for the implicit knowledge base that makes table tennis and other ball games possible.

(*b*) ***Driving***

Driving involves a variety of visual tasks: steering around bends, adjusting speed in relation to road conditions and other vehicles, checking and avoiding stationary and moving obstacles, reading road signs, and so on. Here we consider only steering, and ask the following questions. What visual information does a driver need to turn the steering wheel by the correct amount? From what part of the road does this information come, and what role do eye movements play in obtaining it?

Previous eye movement studies of driving have been performed on relatively undemanding roads, and as a result no clear pattern of fixations emerged (Seraphin 1993). We chose a narrow road with an unpredictable and almost continuous pattern of bends (Queens Drive round Arthur's Seat in Edinburgh), which left little opportunity for gaze to disengage from the steering task (Land & Lee 1994). Three drivers had their gaze direction monitored whilst driving a car fitted with a steering monitor. The pattern of gaze was unexpected and consistent (figure 3*a*, *c*): when approaching a bend all drivers fixated the 'tangent point'—the point on the inside of the bend where the kerb changes direction and the driver's line of sight becomes tangential to the road edge. Between bends gaze returned to the centre of the road, but sought out the next tangent point 2–3 s before the new bend (arrows on figure 3*a*). On two-lane roads the equivalent of the tangent point in right-hand bends (in the UK) is the centre line, and drivers fixated that rather than the far kerb.

Why the tangent point? First and most important, the lateral position of the tangent point relative to the driver's current heading provides a straightforward measure of the curvature of the road ahead: the greater the offset, the tighter the curvature (Land & Lee 1994; Raviv & Herman 1993). Thus, since steering wheel angle converts directly to the curvature of the car's track, tangent point position can be used as a direct input to the steering control loop. Secondly, the tangent point is the only point in the motion flow-field that does not move laterally, except when curvature changes. The eye can 'rest' on the tangent point, whereas all other points in the field will drag the eye around. Typically the driver's eyes make small saccades around the tangent point, but smooth tracking is also seen. Thirdly, by looking at the tangent point the driver has a good view of the road beyond the bend, and so can be aware of distant hazards.

This account shows that the eye movement system 'knows' that the tangent point is the appropriate place to look, and it must also have instructions for finding it (the tangent point is marked by its protrusion into the road, and also perhaps by its relative stationarity). The information that the visual system passes to the motor system that turns the wheel is a measurement of the offset of the tangent point from line of travel, the latter perhaps determined simply as the trunk axis, if the driver is belted to the seat. Drivers may take in other information while steering. Figure 3*d* shows an example where a driver is steering round a bend and checking repeatedly on a cyclist on a parallel course. The points to notice are the switches of gaze every half second, and the fact that steering is related *only* to the direction of gaze when viewing the road, not the cyclist. This seemingly obvious point shows that gaze information can be disconnected from the buffer containing the steering information, presumably at the instant of the saccade to the cyclist.

Further evidence for a buffer comes from the fact that steering action does not follow gaze direction immediately, but after an interval of about 0.8 s (figure 2*b* and Land (1996)). This is not simply a reaction time—these are typically in the range 0.3 to 0.4 s—but rather a period inserted to allow the car to catch up with the point 5–10 m ahead, where the driver's estimate of road curvature was made.

Tangent point location is not the only input to the steering mechanism. It provides information about upcoming road curvature, but not the present position of the vehicle with respect to the lane boundaries close to the vehicle. This is done with peripheral vision while the fovea itself is directed further up the road (Donges 1978; Land & Horwood 1995). How this dual information supply is managed is an interesting question; does attention alternate, or are there really two simultaneous input streams?

(*c*) ***Reading text and music***

The pattern of eye movements in text reading has been described many times, and there is much agreement on the main features (Buswell 1920; Yarbus 1967; Rayner 1983; O'Regan 1990). It consists of a series of stepwise saccades, to the right in English, with each fixation covering four to nine letters, the number depending more on the individual than on the actual text. Regressions of three to four letters occur frequently. Rather surprisingly, saccade size adjusts to the size of the print; each fixation takes in a given number of letters and not a constant visual angle. Fixation durations vary little between individuals, with means between 225 and 300 ms, and within a single piece of text reading a typical standard deviation is about 100 ms. Information about where the eye should look next comes from further afield, with evidence that the saccade pattern is affected by word length and shape up to 15 characters ahead, even though these characters are not actually read (Jacobs 1986). It remains hard to predict exactly where a saccade will land on a word, but 'centre of gravity' effects are detectable, and where two saccades are made to a word, the second is affected by the word's lexical structure (O'Regan 1990). Thus what is processed at one fixation can affect the destination of the next.

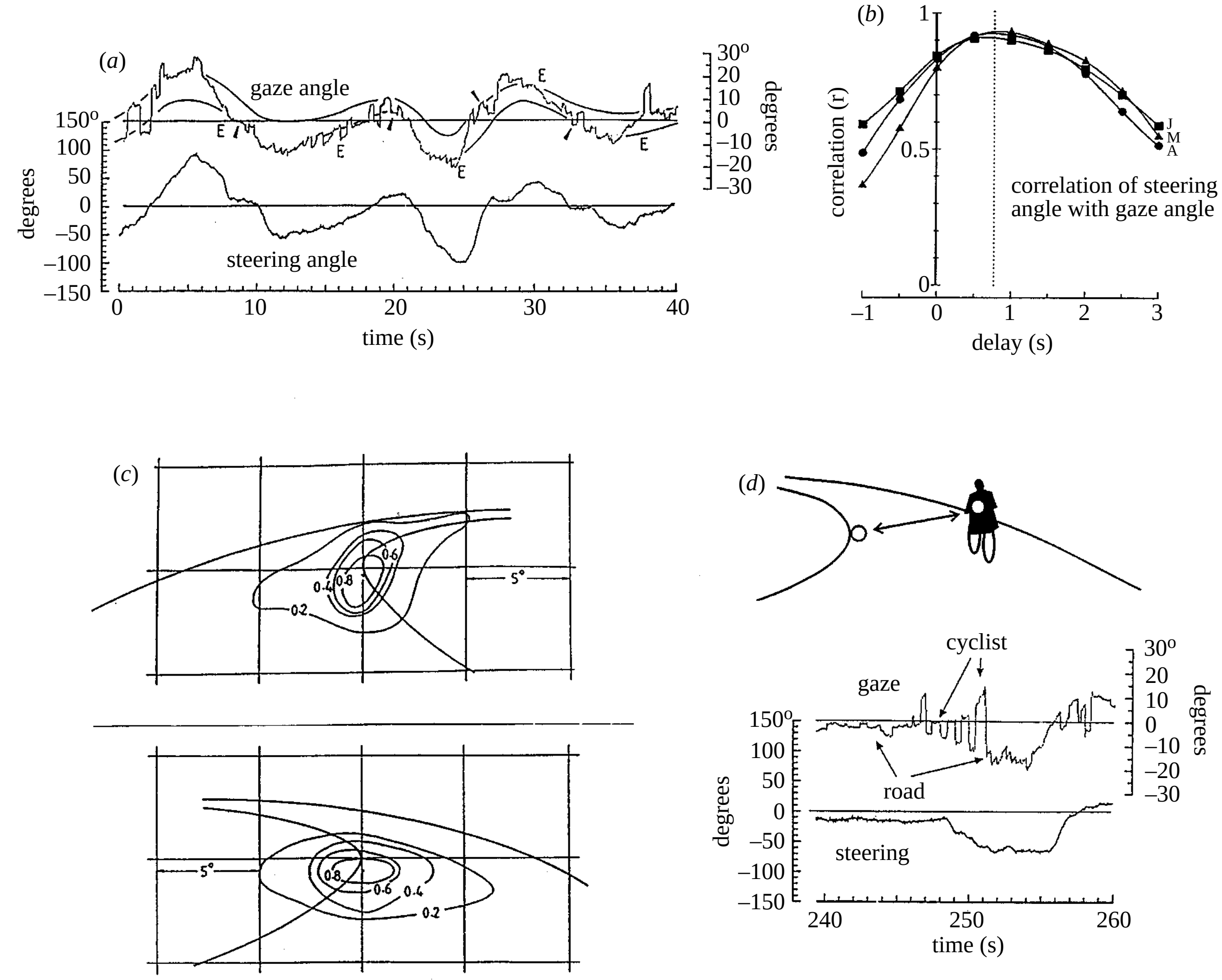

Figure 3. Eye movements when steering a car. (*a*) Record of gaze direction (eye + head) relative to the car's heading when negotiating three right- and three left-hand bends (top) and corresponding steering wheel angles (bottom). The records are very similar. Heavy lines on upper record show the edges of the road. Gaze follows the insides of the bends (tangent points). Short arrows, first saccade to tangent point of new bend; E's, exit from bend and disappearance of tangent point. (*b*) Correlation between gaze angle and steering wheel angle (see (*a*)) for different delays between the two. For all three drivers, steering follows seeing with a delay between 0.5 and 1 s. (*c*) Contour plots of all fixation positions of three drivers negotiating right- and left-hand bends. The contours represent equal numbers of fixations per second per square degree, normalized to their maximum value. Both plots are tightly distributed around the 'tangent points' where the kerb direction changes (Land & Lee 1994). (*d*) Gaze and steering record (as (*a*)) where the driver is dividing gaze between the road tangent point and a cyclist, switching every 0.5 s. Note that steering follows the tangent point position, *not* the cyclist.

In ordinary reading the 'output' is internal—comprehension and memory—and hard to pin down in time. However, in reading aloud there is a spoken output, which occurs about a second after the fixation that took in the written words (Buswell 1920; Geyer 1969). This implies that the buffer (figure 1) holds around 20 characters, or several words.

Musical sight-reading shares many of the features of text reading, but it is a more complex skill, and has been studied much less intensively (Goolsby 1989). It involves a large amount of processing of the input signal—the notes on the page have to be decoded to give their pitch, duration, timing and dynamics. It also requires great competence in the execution of the motor output. Keeping input and output synchronized and mistake-free is a skill that takes years to perfect.

The pattern of acquisition of notes from the page is a more obviously structured process than in text reading. As first shown by Weaver (1943), sight-reading requires the gaze to alternate between the notes of the upper and lower staves (figure 4*a*). Most players also take time out from note reading to make brief checks on their hand positions. What guides the eye from fixation to fixation? The same questions arise as in text reading and visual search. The pattern of fixations has a predetermined element, demonstrated by stave alternation and the regularity of the progression across the page. Saccades land near note heads, however, showing that note location is seen and used. Fixations are longer and their timing is less regular than in text reading (mean 412 ms, s.d. 309 ms for a variety of sight readers and pieces), the variation implying that this is less of a mechanical process and

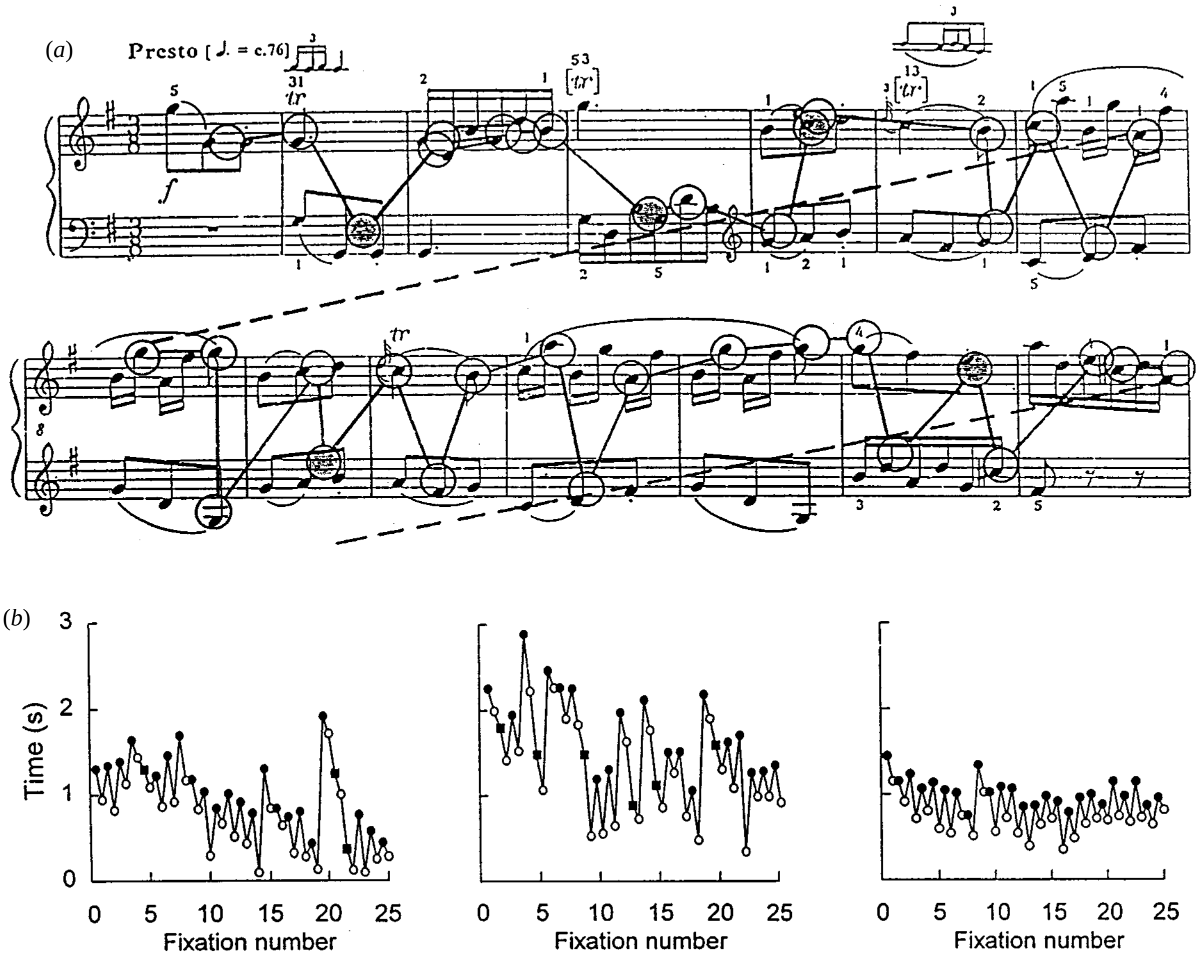

Figure 4. Eye movements and musical sight reading. (*a*) Fixations (circles) made by a professional accompanist sight-reading a Domenico Scarlatti sonata (Presto!). Note the alternation between staves, and the density of fixations (about 1 per 2.5 notes). Shaded circles indicate saccades down to the fingers, followed by a return to the next circled note. These excursions seem to have no effect on the pattern or accuracy of note fixations. (*b*) The time between fixating a note and its execution (eye-hand span) during a fast (left) and a slow (centre) piece played by a novice, and a fast piece (*a*) above) played by a professional (right). Expertise improves consistency, but not preview time, although this is increased for slower tempos. Dots and circles indicate beginnings and ends of fixations, and squares are regressions.

more affected by ongoing cognitive processing. In agreement with this, fixations are generally longer when the music is melodically or rhythmically difficult. From this it seems possible that a new saccade is made only when the information from the previous fixation has been processed and released to the buffer (Kinsler & Carpenter 1995). Thus it appears that the timing and destination of each saccade is determined by an overall strategy, by the seen locations of the notes themselves, and by the state of ongoing cognitive processing. This cognitive component must also involve auditory feedback.

The presence of a buffer in the system is more obvious here than in the other examples. The information arrives in a serial form as alternate chunks from the upper and lower staves, yet the notes are played together in parallel as chords or synchronized melodies. Clearly some temporal re-assembly in a buffer of some kind must occur between input and output. The length of time that a particular note spends in the buffer can be estimated by the delay between the fixation on which it was acquired and the time it is performed. As figure 4*b* shows, this 'time index' has a typical value of about a second, but varies considerably through the piece. (This index includes some processing time and some for the organization of motor output, as well as time passing through the buffer). Skilled performers show less variation than beginners, leading to the idea that maintaining a steady throughput of information is an important aspect of sight-reading skill. The buffer contents can also be assessed by the number of notes interposed between fixation and execution (the 'note index'). This varies much more obviously with skill level, from about two notes for beginners to five for skilled sight readers, up to an absolute maximum of about eight (Weaver 1943; Sloboda 1974). This difference allows better players to play more notes in a given time, and so achieve faster tempos.

A final point concerns the role of memory. Except perhaps at the highest levels of sight-reading skill, the performance of a particular piece improves with repetition. This contrasts with text-reading where the mature skill ensures an almost perfect first-time reading. As a keyboard piece is learned, some elements are incorporated into long-term memory, and at the next performance we find that these notes or chords are not fixated. This means that during a performance there is a blending of material from long-term memory with new or unlearned material, taken in *de novo* fixation by fixation. Conceptually it is difficult to imagine how this structured mixing would work, but it does, and it makes the piece easier and more comfortable to play. This seems to be an interesting paradox.

4. CONCLUSIONS

(*a*) *Task specificity*

The eye movement patterns involved in each of the activities discussed above are very different from each other, specific to the particular tasks, and broadly repeatable. Thus it seems sensible to think of the eye movements as an integral part of each skill, as indicated in figure 1*a*. We argue here that however a task is organized—with self-maintaining loops or 'chained' subtasks (figure 1*b*,*c*)—each step has its own set of oculomotor instructions, or subschemas in the Norman & Shallice (1986) model. The eye must be told where to look, what to expect there, and what further observations or measurements to make. Presumably, as a motor skill is learned so too are the eye movement instructions that go with it. It would be interesting to know at what point a novice table tennis player starts to anticipate the ball's bounce point, or a learner driver starts to use tangent points. One is not taught these things.

(*b*) *Where to look*

In the 'reactive' saccades of the laboratory the input to the saccadic system is usually the position of the 'stimulus', i.e. its direction and angular distance relative to the fovea. In the 'proactive' saccades of normal life the input may be many different things. In table tennis it appears to be a complex of variables related to the trajectory of the ball (figure 2*d*). In steering it involves locating and recognizing the tangent point, and tracking it (figure 3*c*). In music it is a note a few notes further on in the score (figure 4*a*). Thus the instructions to the eye movement system are very task-specific, and may involve computation (table tennis), pattern recognition (steering, music) or remembered position information (looking down to the piano keys and back again, or locating the light switch).

(*c*) *What to look for*

What the eye is called upon to extract is also task-specific. It could simply be the answer to a yes/no question: Is there a vehicle at the intersection? Is the traffic light red? It could be a measurement or estimate: what angle does the tangent point make with my line of travel? Where will the ball be in half a second's time? Or it may be a complex coded message such as 'right hand, C and D quavers; left hand, G crotchet'. The temporal unit for information uptake is usually a single glimpse, a fixation lasting 200–300 ms, which varies surprisingly little from task to task. That time is long enough to take in five to eight letters, or the pitch and length of two to six musical notes, or determine the identity of a face or of a great many other objects.

(*d*) *Buffers*

The degree of involvement of a buffer in the control of action varies between tasks, from very little in table tennis to the retention of a complex set of note descriptions in music reading. The chief function seems to be the brief holding of information to allow a match between the episodic input and the continuous motor output. In text and music reading the delay in the buffer may just reflect the slack in the system required for the buffering process itself, but in steering the 0.8 s delay is needed to allow for the vehicle's time of travel. Buffers also permit the eyes to look elsewhere, briefly, without interfering with the ongoing control of action. Thus typists can check their work (Fleischer 1986), drivers can take their eyes from the road to check signs and other traffic, and pianists can check the positions of the fingers. There are many unsolved questions about the nature of buffers. Is there one or are there many; what are their capacities; how is information lost from them (is this just a fast decay process or is the information erased as motor acts are completed); and how is information from long-term memory combined with new visual information, as occurs when reading a part-learnt score?

(*e*) *Multitasking*

In many tasks the eye gathers information for more than one simultaneously executed subtask. Typically, gaze alternates from one information source to another. Switching between the road and potential hazards in driving (figure 3*d*) provides a particularly interesting example because it demonstrates that whilst one subtask (viewing the road) contributes to the information in the steering buffer, the other (viewing the cyclist) does not. Evidently, the routing of information from the eye must switch every half-second, as the subtask changes. So, presumably, must the 'search image' as well. In the example in figure 3*d* the cyclist did not swerve, and no steering action was required. Had he done so, however, the driver would have needed to merge information about the cyclist's position with that from the steering buffer, to produce an appropriate avoidance course. Something similar occurs in music reading; if checking the keys reveals an error then the motor output requires modification, otherwise the score alone determines the output.

(*f*) *Active vision revisited*

If 'active vision' is to mean more than simply obtaining heading from a flow field, or distance from motion parallax, in particular if it is to be used to scrutinize a scene for information that will guide action, then such a system will need many of the properties of the human oculomotor system.

In an artificial system the detectors may or may not have to move. Some advances have been made in the design of mobile 'saccading' camera systems (Murray *et* *al*. 1995; Janssen 1996), but most robotic systems employ fixed cameras. In humans the fact that both resolution in the retina and processing capacity in the cortex fall off dramatically with distance from the centre of gaze (Anstis 1974; Latham & Whitaker 1996), makes it essential for the eyes to move so as to direct the fovea appropriately. In a system with adequate resolution across the field of view it is possible that all the scrutiny could be performed in software with something equivalent to the 'spotlight of attention' interrogating the image. Whatever the exact mechanism for information extraction, however, it seems that something equivalent to the repertoire of vision/action modules (figure 1) that humans possess will have to be built into any robotic device with a claim to versatility.

REFERENCES

Aloimonos, Y. (ed.) 1995 *Active perception*. Hillsdale, NJ: Erlbaum.

Anstis, S. M. 1974 A chart demonstrating variations in acuity with retinal position. *Vision Res.* **14**, 589–592.

Ballard, D. H., Hayhoe, M. M., Li, F. & Whitehead, S. D. 1992 Hand–eye coordination during sequential tasks. *Phil. Trans. R. Soc. Lond.* B **337**, 331–339.

Blake, A. & Yuille, A. (ed.) 1993 *Active vision*. Cambridge, MA: MIT Press.

Buswell, G. T. 1920 An experimental study of the eye–voice span in reading. *Suppl. Educational Monographs no. 17.*

Carpenter, R. H. S. 1988 *Movements of the eyes*. London: Pion.

Carpenter, R. H. S. (ed.) 1992 *Eye movements. Vol. 8 of Vision and visual dysfunction* (general ed. J. Cronly-Dillon). Basingstoke, UK: Macmillan.

Donges, E. 1978 A two-level model of driver steering behavior. *Hum. Factors* **29**, 691–707.

Fleischer, A. G. 1986 Control of eye movements by working memory load. *Biol. Cybern.* **55**, 227–238.

Geyer, J. J. 1969 Perceptual systems in reading: the prediction of the eye voice span. In *Perception and reading* (ed. H. K. Smith). Newark, DE: International Reading Association.

Goolsby, T. 1989 Computer applications to eye movement research in music reading. *Psychomusicology* **8**, 111–125.

Heuer, H. 1996 Dual task performance. In *Handbook of perception and action. Vol. 3. Attention.* (ed. O. Neumann & A. F. Sanders), pp. 113–153. London: Academic.

Jacobs, A. M. 1986 Eye movement control in visual search: how direct is visual span control? *Percept. Psychophys.* **39**, 47–58.

Janssen, H. 1996 Saccadic camera control for scene recognition on an autonomous vehicle. In *Visual attention and cognition* (ed. W. H. Zangemeister, H. S. Stiehl & C. Freska), pp. 293–307. Amsterdam: Elsevier.

Kinsler, V. & Carpenter, R. H. S. 1995 Saccadic eye movements while reading music. *Vision Res.* **35**, 1447–1458.

Land, M. F. 1993 Eye–head coordination during driving. *IEEE Proc. SMC Conf. Le Touquet v3*, pp. 490–494.

Land, M. F. 1996 The time it takes to process visual information when steering a vehicle. *Invest. Ophthal. Vision Sci.* **37**, S525.

Land, M. F. & Lee, D. 1994 Where we look when we steer. *Nature* **369**, 742–744.

Land, M. F. & Horwood, J. 1995 Which parts of the road guide steering? *Nature* **377**, 339–340.

Latham, K. & Whitaker, D. 1996 A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Res.* **36**, 2665–2674.

Murray, D. W., Bradshaw, K. J., McLaughlin, P. F. & Reid, I. D. 1995 Driving saccade to pursuit using image motion. *Int. J. Computer Vision* **16**, 205–228.

Norman, D. A. & Shallice, T. 1986 Attention to action: willed and automatic control of behaviour. In *Consciousness and self-regulation, vol. 4.* (ed. R. J. Davidson, G. E. Schwartz & D. Shapiro). New York: Plenum.

O'Regan, J. K. 1990 Eye movements and reading. In *Eye movements and their role in visual and cognitive processes* (ed. E. Kowler), pp. 395–453. Amsterdam: Elsevier.

Rayner, K. 1983 The perceptual span and eye movement control during reading. In *Eye movements and reading* (ed. K. Rayner), pp. 97–120. Academic Press.

Raviv, D. & Herman, M. 1993 Visual servoing from 2-D image cues. In *Active perception* (ed. Y. Aloimonos), pp. 191–226. Hillsdale, NJ: Erlbaum.

Ripoll, H., Fleurance, Ph. & Cazeneuve, D. 1987 Analysis of visual patterns of table tennis players. In *Eye movements: from physiology to cognition* (ed. J. K. O'Regan & A. Lévy-Schoen), pp. 616–617. Amsterdam: Elsevier.

Serafin, C. 1993 Preliminary examination of driver eye-fixations on rural roads: insight into safe driving behavior. *Univ. Michigan Transportation Res. Lab. Report UMTRI-93-29*, pp. 1–63.

Sloboda, J. A. 1974 The eye–hand span: and approach to the study of sight-reading. *Psychol. Music* **2**, 410.

Thomas, E. L. 1968 Movements of the eye. *Scient. Am.* **21**(2), 88–95.

Viviani, P. 1990 Eye movements in visual search: cognitive perceptual and motor control aspects. In *Eye movements and their role in visual and cognitive processes* (ed. E. Kowler), pp. 353–393. Amsterdam: Elsevier.

Weaver, H. E. 1943 Studies of ocular behavior in music reading. I. A survey of visual processes in reading differently constructed musical selections. *Psychol. Monogr.* **55**, no. 249, 130.

Yarbus, A. 1967 *Movements of the eyes*. New York: Plenum.