The Scanpaths.org Archive of Eye Movement Data from Everyday Actions: Mission, History and New Data Sets

(January 2008)

Stavri G. Nikolov[1], Jeff Pelz[2], Michael F. Land[3], and Tom Troscianko[4]

[1]Attentive Displays Ltd, UK and Bulgaria
(e-mail: stavri.nikolov@attentivedisplays.com)

[2]Carlson Center for Imaging Science, Rochester Institute of Technology, USA
(e-mail: pelz@cis.rit.edu)

[3]Department of Biology and Environmental Science, University of Sussex, UK
(e-mail: m.f.land@sussex.ac.uk)

[4]Department of Experimental Psychology, University of Bristol, UK
(e-mail: tom.troscianko@bristol.ac.uk)

Summary
This report discusses the Scanpaths.org archive. The mission, history, scope and structure of the archive are briefly presented, followed by detailed descriptions of all the scanpath data sets publicly available at Scanpaths.org, including some recently published eye-movement data sets (December 2007). Cognitive systems have to solve particular problems using specific strategies. A major problem in getting an artificial cognitive system to do such things lies in the difficulty of knowing which of many items of information in a complex scene are relevant to a particular task. A potentially powerful solution to this problem is to provide the cognitive vision community with a resource of natural scenes containing human scanpaths (eye movement traces) when the human observers are carrying out a given task. The Online Archive of Scanpath Data (Scanpaths.org) which was created in 2004 at present hosts a large and diverse collection of such scanpath data sets, including eye movement data of people performing a visual search task, booking airplane tickets online, tracking a target in multi-sensor video, making tea, driving a car in a city centre and in the countryside, playing cricket, drawing a portrait, playing jigsaw puzzle, attending a party and playing table tennis. The Scanpaths.org resource is expected to become increasingly useful to those who wish to construct artificial cognitive systems (such as robots, adaptive computer vision systems or software agents) capable of carrying out complex tasks (such as target selection or navigation to a target) in complex environments.

Scanpaths.org: Mission

The Online Archive of Scanpath Data (Scanpaths.org) is located at www.scanpaths.org and is kindly hosted by the Center for Imaging Science, Rochester Institute of Technology (RIT), USA. Its mission is to provide the eye movement and cognitive science research community with a central online repository of scanpath (eye-movement) data gathered by different laboratories and companies around the world in various experiments and applications. This data, especially when well structured and documented and available in a common eye-tracker independent format, could be used by researchers to develop new gaze-tracking and visual information analysis algorithms and systems, without having to recreate the experiments carried out to collect the data. The data can also be used to train artificial cognitive systems such as robots to perform vision and navigation tasks in various environments. The Scanpaths.org web site was created and is maintained by Dr Stavri Nikolov (Attentive Displays Ltd, UK and Bulgaria) and Dr Jeff Pelz (RIT, USA).

Scanpaths: Definitions, Theory and Applications

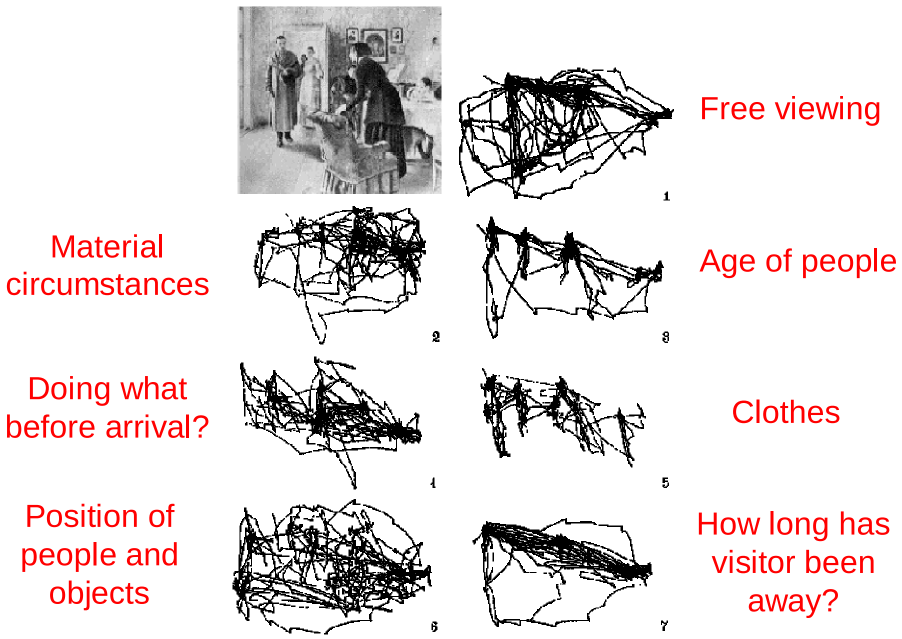
The term "scanpath" was first used by Noton and Stark in their papers published in 1971 in Vision Research and Science, entitled "Scanpaths in saccadic eye movements while viewing and recognising patterns" and "Scanpaths in eye movements during pattern perception", respectively. For the purpose of this archive we will consider a "scanpath" to be any eye-movement data collected by a gaze-tracking device, where information is recorded about the trajectories (paths) of the eyes when scanning the visual field and viewing and analysing any kind of visual information. Such data usually consists of gaze-direction, fixation position and duration, and saccade duration. Our definition of scanpath is broader than the one used by most researchers, including Noton and Stark's original definition, but it allows us to include data collected from different experiments and in various applications in a single framework. Although data from reading, eye-typing and other experiments can also be considered as scanpaths if we use this definition, the data that is included in the Scanpaths.org archive will be primarily from experiments dealing with the visual analysis of natural images and video, with a special focus on data gathered from everyday activities.

Scanpaths.org: Brief History

- March 2004, idea about the creation of the Scanpaths.org archive conceived
- November 2004, web site design started
- December 2004, initial design of web site ready
- March 2005, **Cattaneo, Rosen, Vecchi, & Pelz data set** (Data Set D1) published
- March 2005, Scanpaths.org presented at the Applied Vision Association (AVA) Active Vision meeting in Bristol, UK
- August 2005, Scanpaths.org presented at the 13th European Conference on Eye Movements (ECEM) in Bern, Switzerland
- December 2005, **SR Labs Alitalia data set** (D2) published
- June 2006, **Bristol Eden Project Multi-Sensor Video data set** (D3) published

- February 2007, a **Network Action** proposed by Dr Stavri Nikolov and Prof Tom Troscianko (of Bristol University) to extend the Scanpaths.org archive with some historical eye movement data (Data Sets 4, 5, 6 and 7), gathered by Prof Mike Land (Sussex University), has been approved for funding by the **euCognition** Network Committee
- April 2007, Scanpaths.org presented at the Tobii User Meeting in Frankfurt, Germany
- May 2007, A **blog** associated with Scanpaths.org has been created and is maintained by Sven Laqua of UCL.
- November 2007, Eight new Data Sets (D4-D11) contributed by Prof Mike Land have been published on Scanpaths.org.

Scanpaths.org: Relevance to euCognition and the Cognitive Sciences Community



The above picture ("They Did Not Expect Him" by I.E. Repin, 1884), from Yarbus (1967), illustrates one of the main problems we have tried to address with the creation of Scanpaths.org. The same picture was presented to an observer who had different tasks, such as determining the age, or the social status, of the people in the image. The scanpaths were entirely different and task relevant. The difficulty in providing appropriate information to a vision system is well demonstrated by this example. While it is easy to pick out low-level features of a scene according to their "salience" (e.g. Itti and Koch, 2000) it is very difficult to ascertain which particular features are relevant to a given action or task. However, human vision appears to solve this problem in an exquisitely elegant and robust manner, as evidenced by the work of Land, Hayhoe, and others. When carrying out normal everyday tasks, such as making a cup of tea or a sandwich, driving, or playing ball games, almost all fixations are directed to task-relevant locations, just ahead of the action towards/at those locations. It is thus possible to use this high-level knowledge embodied in human scanpaths to train artificial systems to perform specific tasks. The utility of the Scanpaths.org archive will grow with the number and diversity of examples provided – we therefore intend for this to be a rapidly developing resource. While a lot of large collections of facial expression, locomotion, hand gesture, etc., data sets are publicly available, no large eye movement data sets have been made public by research labs in the past. Our central argument is that human fixation data provide a "shortcut" to the *relevant* aspects of such scenes.

Data Set Structure and Information Provided

For each data set in Scanpaths.org there is a separate web page. This includes the information listed below, organised in four different sections. All data sets are sorted by contributor. Both current and future (under development) parts of the archive are discussed.

EXPERIMENT DESCRIPTION

In this section a detailed description of the experiment which produced the data is given. This information is either specially written for the Scanpaths.org web site or a link to a previous paper/report describing the experiment is included.

EYE IMAGES/VIDEOS

The raw eye images/video files are provided where available. Preferred formats are AVI, WMV, and MOV.

SCENE/DISPLAY/STIMULI IMAGES/VIDEOS

The scene or display images/video files are provided. Preferred formats are JPG for stills and AVI, WMV, and MOV for video files. The scene or display videos need to be synchronised with the raw eye images or at least time stamped.

SCANPATH (EYE-TRACKER) DATA

The scanpath (eye-movement) data gathered in the experiment is usually published on Scanpaths.org in raw (native eye-tracker) and TXT formats. The minimum information that needs to be provided includes a timestamp, and horizontal and vertical eye (or gaze) information. Additional data could include pupil diameter, etc. The scanpath data has to be synchronised with the raw eye images and the scene or display images from which it has been extracted. Information about the eye-tracker data format and the algorithm applied to compute this data is also published, where available. Later on, tools to convert the raw eye-tracker specific data into a common eye-tracker independent format will be provided on Scanpaths.org.

Common Eye-Tracker Data Independent Format

A general description of the common eye-tracker independent scanpath data format which we propose will be provided soon at Scanpaths.org. The need for such a format has been argued by many researchers (see COGAIN report 2.2 by Bates et al., 2005). With more conversion tools being created and more labs and eye-tracker manufacturers hopefully joining in the future and using this format, it will evolve in time.

Tools for Analysis and Visualisation of Data in this Format

Different conversion, data analysis and visualisation tools will be published online at Scanpaths.org in the future. We will start with some basic software programs that convert from the specific scanpath data formats used by some of the most widely used eye-trackers, into our common data format. These converters will hopefully be created by the people contributing data in such a format, by the eye-tracker manufacturing companies or by third parties, and will be made publicly available. Simultaneously, we will try to develop software tools that will analyse and visualise any data stored in our common scanpath data format.

Papers about Scanpath Theory and Applications of the Data Sets Available on Scanpaths.org

This section includes a comprehensive list of publications on the subject of eye- and gaze-tracking data collection and analysis in general, and the use of the data sets available on Scanpaths.org in different applications in particular.

Acknowledgments

We would like to thank the Center for Imaging Science at the Rochester Institute of Technology (USA) for hosting the Scanpaths.org archive, euCognition (The European Network for the Advancement of Artificial Cognitive Systems) for the financial support through Network Action NA024-2 to process, annotate and publish Mike Land's data sets on Scanpaths.org, and the University of Bristol (UK) and Attentive Displays Ltd (UK and Bulgaria) for supporting this project. Many thanks also to all the contributors of eye-movement data to Scanpaths.org and to Sven Laqua of UCL (UK) for creating and maintaining the Scanpaths blog.

References

D. Noton and L.W. Stark, Scanpaths in eye movements during pattern perception, Science, 171, pp. 308-311, 1971

A. Yarbus, Eye Movements and Vision. Plenum Press, New York, 1967

L. Itti and C. Koch, A saliency-based search mechanism for overt and covert shifts of visual attention, Vision Research, vol. 40, no. 10–12, pp. 1489–1506, 2000

R. Bates, H. Istance, and O. Spakov, Requirements for the Common Format of Eye Movement Data, Communication by Gaze Interaction (COGAIN), IST-2003-511598: Deliverable 2.2, 2005
M. F. Land, Eye movements and the control of actions in everyday life. Prog Retinal & Eye Res 25: 296-324, 2006
M. Hayhoe and D. Ballard. Eye movements in natural behavior. Trends in Cognitive Sciences, 9, pp. 188–194, 2005

Appendix A: Short Descriptions of Data Sets Available on Scanpaths.org

Eleven scanpath data sets (D1-D11) have been published on Scanpaths.org between 2005 and 2008. They are listed below together with the names of their contributors:

D1: Study and recall of multiple targets on a photograph

Study and recall of multiple targets on a photograph presented on a large-format display.
Dr Zaira Cattaneo, University of Pavia, Italy
Dr Jeff Pelz, Rochester Institute of Technology, USA

D2: Exploring Alitalia's web site and booking tickets

Making an online booking using the web site of Alitalia.
Dr Francesco Maringelli, SR Labs, Italy

D3: Target tracking in multi-sensor video

Visually tracking a soldier walking through thick foliage in a multi-sensor (visible and IR) video sequence.
Dr Tim Dixon and Dr Stavri Nikolov, University of Bristol, UK

D4: Tea making

Eye tracker record with head movements. This was designed to find out the temporal relations between eye movements and limb movements during this ordinary, over-learned task. In general eyes locate the next object in the task about half a second before manipulation begins, and leave for the next object about half a second before manipulation is complete.
Prof Mike Land, Sussex University, UK

D5: Car driving/steering

Eye movements of 3 drivers recorded with eye tracker. Head movements also available. This is a winding road, and the intention was to find out what cues drivers use to steer round bends. Much use made of the tangent point on the inside of bends.
Prof Mike Land, Sussex University, UK

D6: Playing cricket

Eye and head movements of three batsmen facing a bowling machine. Main result is that batsmen make saccades to expected bounce point ahead of the ball, thereby establishing the time and place that the ball will reach the bat.

D7: Portrait drawing
Eye movement studies of an artist drawing fast sketches and longer detailed drawings. The observations were designed to work out what the artist extracted from the sitter on each look, and what were the timings of information acquisition and actions.

D8: Jigsaw puzzle
Eye movement studies of two subjects as they completed a 50 piece puzzle of a cartoon scene.
Prof Mike Land, Sussex University, UK

D9: Party
Eye movements of a subject having various conversations with people at a party.
Prof Mike Land, Sussex University, UK

D10: Table tennis
Eye movements of table tennis players.
Prof Mike Land, Sussex University, UK

D11: Urban driving
Eye movements of a driver in the town of Lewes in Sussex.
Prof Mike Land, Sussex University, UK

Appendix B: Detailed Descriptions of Data Sets Available on Scanpaths.org

Eleven scanpath data sets (D1-D11) have been published on Scanpaths.org between 2005 and 2008. They are described on the following pages together with some figures illustrating these data sets and relevant references.

Study and Recall of Multiple Targets on a Photograph (D1)

Description
As part of an experiment probing the picture superiority effect Zaira Cattaneo (University of Pavia, Italy), observers' gaze was tracked during study and recall of multiple targets on a photograph presented on a large-format display. Observers viewed a photograph containing eight objects for 30 seconds. Their task during the 'study phase' presented here was to remember the location of the objects in the photograph. Eye movements were monitored with an Applied Science Laboratories model 501 headband-mounted eyetracker. Head movements were monitored with a Polhemus Fastrak magnetic tracker. The eye and head position were integrated to compute the gaze position on the display. The output of the eyetracker includes video field number (1/60 sec), relative pupil diameter, horizontal and vertical gaze position on the display.

Fig 1 Images were displayed on a Pioneer 50" plasma display (503CMX) with a resolution of 1280 x 768 pixels. Observers were seated 110cm from the 110 cm X 62 cm display The images were 1053 x 556 pixels, subtending approximately 45 x 23 degrees.

Reference:
Cattaneo, Z., Rosen. M., Vecchi, T., & Pelz. J.B., Monitoring eye movements to investigate the picture superiority effect in spatial memory, Perception, (in press)

Exploring Alitalia's Web Site and Booking Tickets (D2)

Description

The present experiment was carried out by the eye-tracking company SR LABS (Milano, Italy). The object of the analysis was the homepage of the web site of Alitalia . The eye-movement registration was divided into three sessions of subject-website interaction:

Fig 1 Display setup (left) and web site stimulus (right).

1st Interaction: The subjects were asked to explore in a free way, for about 15 seconds, the homepage of the web site under analysis. The goal of this kind of test was to investigate the attention distribution among the various areas that constitute the web page and to find possible exploration paths across the web page that appeared to be preferred by the subjects.

2nd and 3rd Interaction: The subjects were asked to surf inside the web site, without time limits, with the goal of achieving specific tasks. The goal of these sessions was to evaluate the visibility and the informational clarity of the busi-ness parts of the homepage while subjects were executing a task with a busi-ness relevance.

In more detail, in the 2nd and 3rd interactions the tasks to be achieved were the following ones:

2nd Interaction:
Task: You have to go immediately to New York for business reasons. Book the round-trip ticket for this destination.
The goal of this task was to evaluate the capacity of the commercial box in the low part of the homepage, on the right, to attract attention to the detriment of the central booking area.

3rd Interaction:
Task: You have some vacation days. You decide to make a travel and you decide to choose a flight among the special offers of "Alitalia".
The goal of this task was to evaluate some properties of the commercial box in the low part of the homepage, on the right. In particular:

- how the interpretation was easy of the link associated with the box, that is, a list of offers and not only the offers that appeared in sequence in the box itself
- the capacity of the box to attract attention to the detriment of the other areas of the homepage in which there are links to special offers.

The subjects under analysis were four. The eye-tracking tool that was used was the eye-tracker Tobii 1750. The Tobii 1750 looks like a common monitor screen.

Reference:
There are no publications associated with this data.

Bristol Eden Project Multi-Sensor Data Set (D3)

Description

The current study is part of a larger project on multi-sensor image fusion – that is the fusion of complementary inputs such as visible light and infrared radiation images. We have been concentrating on the human and computational assessment aspect of this fusion process, and are using participant scanpaths as a new and untapped source of information. This experiment involved participants watching videos of a man dressed in camouflage walking through thick foliage (see Fig 1) whilst performing a secondary button-press task.

Fig 1 The 'soldier' began walking near to the camera (see Fig 1) and to the left of centre, walks down a path away from camera, then left-to-right across a clearing of trees, and finally right-to-left back across the clearing. The video sequence was filmed in both visible light (Viz) and infrared (IR) at the Eden Project Biome in Cornwall, UK, and more details of this can be found in [1].

Fig 2. In the experiment, participants were shown the registered original input videos (Viz and IR) separately, the two videos together side-by-side, as well as three fused versions of the videos. The first fusion method was simply averaging between pixel values of IR and the values of each colour plane of Viz. The second used a basic Discrete Wavelet Transform (DWT) method, and the third utilised the Dual-Tree Complex Wavelet Transform (DT-CWT). The participants were shown each of the videos once, in three separate sessions. In the first session half the participants were shown the order: Viz, IR, Viz+IR, AVE, DWT, DT-CWT; and half were shown the order: DT-CWT, DWT, AVE, Viz+IR, IR, Viz. In the second session, both sets of participants swapped order, and in the third session participants viewed the original order from Session 1.

The actual tasks the participants were instructed to carry out were as follows. They were told to primarily concentrate on visually tracking the soldier as well as possible throughout the video sequence. In addition, they were asked to press the 'SPACE' bar once whenever the soldier began or finished walking past an electric vehicle positioned in the clearing. That is, when the solder was adjacent to either end of the vehicle.

A Tobii™ x50 remote eye tracker was used to collect eye movement data. This is a table-mounted eye tracker that works at 50 Hz with an approximate accuracy of 0.5 deg. This was run using the ClearView 2.5.1 software package, on a 2.8 GHz Pentium IV dual processor PC, with 3 GB RAM, and twin SCSI hard drives. Stimuli were presented on a 19" flat screen CRT monitor running at 85 Hz, with screen resolution set to 800 by 600 pixels. Participants were required to use a chin-rest positioned 57cm from the monitor screen.

More details of this experiment can be found in [2].

Reference:

[1] J. J. Lewis, S. G. Nikolov, A. Loza, E. Fernandez Canga, N. Cvejic, J. Li, A. Cardinali, C. N. Canagarajah, D. R. Bull, T. Riley, D. Hickman and M. I. Smith. The Eden Project multi-sensor data set. Technical report TR-UoB-WS-Eden-Project-Data-Set, University of Bristol and Waterfall Solutions Ltd, UK, 6 April 2006, ImageFusion.org *(in section publications ¦ reports)*

[2] T.D. Dixon, S.G. Nikolov, J.J. Lewis, J. Li, E.F. Canga, J.M., Noyes, T. Troscianko, D.B. Bull, and C.N. Canagarajah. Scanpath analysis of fused multi-sensor images with luminance change: A pilot study. Proceedings of the Ninth International Conference on Information Fusion, Italy, July 2006

Eye movement recordings by M. F. Land between 1994 and 2004 (Data Sets D4-D11)

General Methods

Eye movement recordings were made with a prototype head-mounted camera that produced a split image in which the top two-thirds showed the scene ahead and the lower third the eye in its socket, imaged via a concave mirror (Fig 1). The recordings were made with a Panasonic GP-LM7-TA camera, onto Sony Hi-8 or Digital tape, and transferred to sVHS tape for analysis using a Panasonic single field AG-7330-B recorder. The tapes were later re-recorded onto disc.

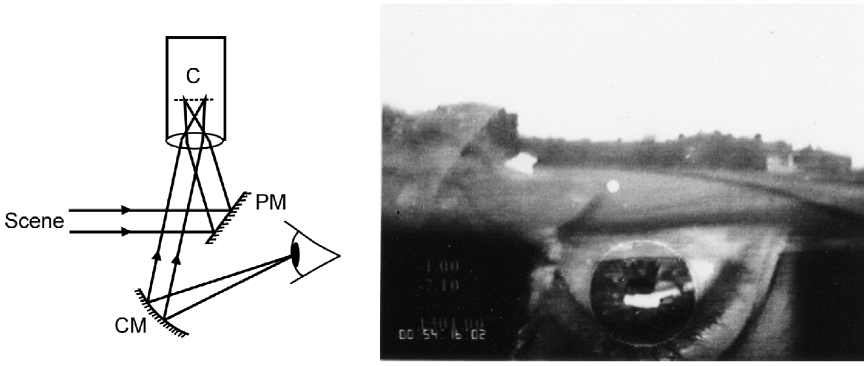


Fig.1 Left: diagram of eye camera (C, camera; PM part-silvered plane mirror; CM, concave mirror). Right: Processed image showing a road with dot indicating foveal direction, and eye in bottom third of the frame.

The location and ellipticity of the iris were used to obtain the coordinates of eye direction, by matching the iris outline to a computer- generated eye model. This was done by hand, frame-by-frame, at 50 f.p.s. The coordinates were used to position a 1º dot on the upper scene view, and each frame re-recorded (Fig 1). The position accuracy was approximately 1º. Head movements could also be obtained by tracking distant background objects in the scene view. The resulting video contains numerical values (in degrees) of the direction of view of the fovea, a frame counter, and a clock. The videos are reversed left to right as a result of the mirror optical system.

Tea Making (D4)

Description

The aim of this study was to determine the pattern of fixations during the performance of a well-learned task in a natural setting (making tea), and to classify the types of monitoring action that the eyes perform. We used a head-mounted eye-movement video camera, which provided a continuous view of the scene ahead, with a dot indicating foveal direction with an accuracy of about 1 deg. A second video camera recorded the subject's activities from across the room. The videos were linked and analysed frame by frame. Foveal direction was always close to the object being manipulated, and very few fixations were irrelevant to the task. The first object-related fixation typically led the first indication of manipulation by 0.56 s, and vision moved to the next object about 0.61 s before manipulation of the previous object was complete. Each object-related act that did not involve a waiting period lasted an average of 3.3 s and involved about 7 fixations. Roughly a third of all fixations on objects could be definitely identified with one of four monitoring functions: locating objects used later in the process, directing the hand or object in the hand to a new location, guiding the approach of one object to another (eg kettle and lid), and checking the state of some variable (eg water level). We conclude that although the actions of tea-making are ‘automated’ and proceed with little conscious involvement, the eyes closely monitor every step of the process. This type of unconscious attention must be a common phenomenon in everyday life.

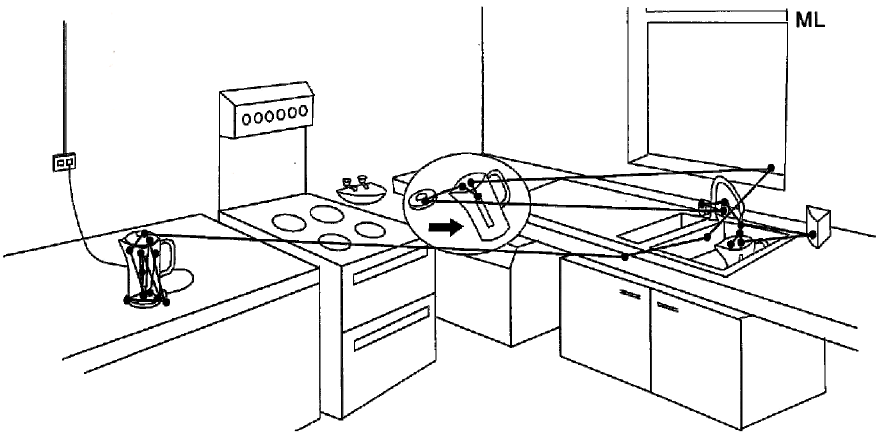


Fig 1 Record of the fixations made by one subject during the first sequence after the kettle is first detected, and during which the kettle is moved from the worktop (left) to the sink (right). Because of the changing viewpoint, the angular relations are only approximate, but fixation positions relative to the objects of regard are accurately represented. Note the associations of fixations with particular objects or other entities - kettle, sink, kettle and lid, taps, water stream - which correspond in time to the actions that relate to them.

Reference:

*Land M, Mennie N, Rusted J (1999) The roles of vision and eye movements in the control of activities of everyday living. Perception **28**: 1311-1328*

Car Driving/Steering (D5)

Description

Although there had been a number of studies of eye movements during driving prior to this study, they mostly involved freeways or other wide roads, and were not particularly challenging in terms of visuo-motor coordination. We (David Lee and I) therefore chose a narrow winding road - Queen's Drive round Arthur's seat in Edinburgh (Fig 1) – because it required steering to be under continuous visual control. We made simultaneous recordings of steering wheel angle and gaze direction using a head-mounted video-based eye monitor with a resolution of about 1^{o}. Head movements relative to the car were also monitored. The most striking feature of the records were that on bends in either direction gaze sought out and tracked the 'tangent point': the moving point on the inner lane edge that protrudes into the road (Fig 1). Gaze could take 'time out' from monitoring the road (e.g.the jogger in Fig 1), but not for long, and not usually on bends. The significance of the tangent point seems to be that its location, relative to the current heading predicts the curvature of the upcoming bend, independent of distance, according to the formula:

$$\text{curvature} = 1/(\text{bend radius}) = \theta^2 / 2d,$$

where θ is the angle of the tangent point from the current heading (i.e. the gaze angle if the tangent point is fixated), and d is the distance of the driver's head from the inner lane edge. Subsequent simulator studies (detailed in Land, 1998) showed that drivers monitor both θ and d simultaneously. Since at low-to-moderate speeds steering wheel angle is proportional to path curvature, θ and d provide the driver with a direct control signal for steering.

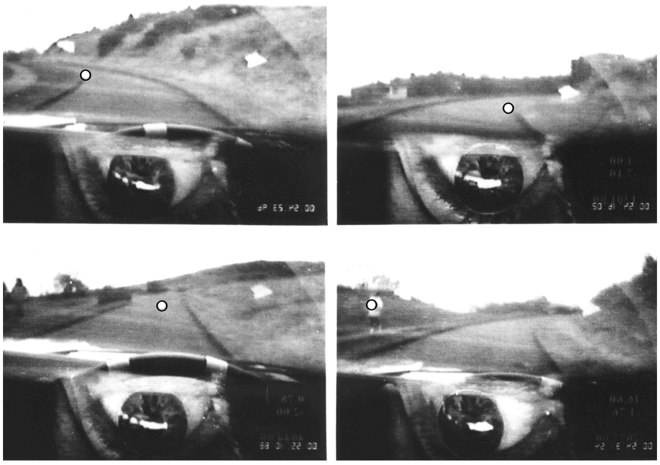


Fig 1. Four frames from a video of a drive around Queen's Drive showing typical gaze directions (white dot) on a left-hand and right-hand bend (top), a straight segment, and while looking off-road at a jogger. Note that the subject's eye occupies the lower third of each frame and unlike the videos themselves the frames are not left-right reversed. The scrap of tape attached to the windscreen allows head movements to be monitored.

Reference:

Land, M.F., 1998. The visual control of steering. In: Harris, L.R., Jenkin, M., (Eds.) Vision and Action. Cambridge University Press. pp.163-180

Cricket Playing (D6)

Description

The aim of this study was to determine where a batsman during the half second or so before the ball reaches his bat. He has to decide in this time what type of shot to make and where and when to make contact with the ball. A key to this is to be in the best possible position to view to ball's bounce. We made eye movement recording of three batsmen of different standards as they faced medium-paced balls from a bowling machine

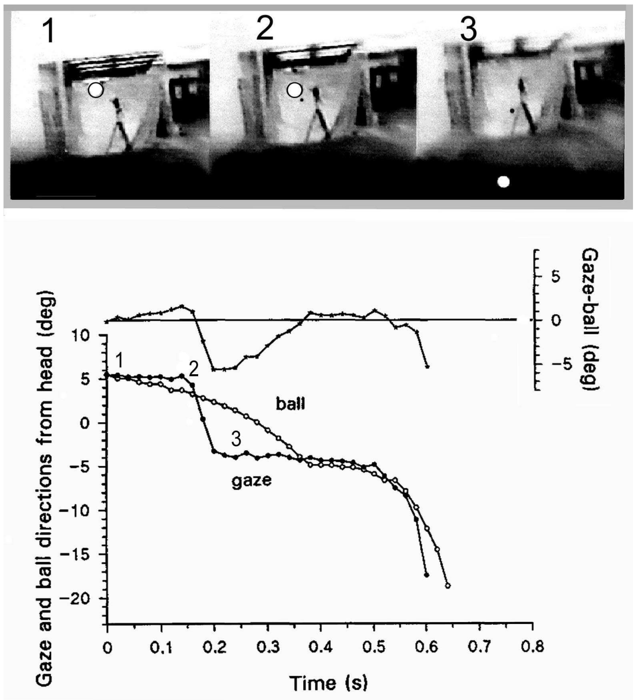


Fig 1. Upper part: The batsman's view of the ball leaving a bowling machine. (1) Ball about to emerge, batsman's gaze (white dot, 1º across) watching the aperture. (2) Ball (small black dot) descending from the aperture with gaze starting to follow. (3) Gaze saccade to a spot close to the bounce point, which the ball will not reach for a further 0.1 s. Object in centre of each frame is a camera tripod. Lower part, main graph. Vertical direction of gaze (●) and ball (○) viewed from the batsman's head. Numbers correspond to photographs above. Note that the saccade after 2 brings gaze close to the bounce point. After the bounce gaze tracks the ball until about 0.6 s after delivery. The ball is struck at 0.7 s. Upper graph: difference between gaze and eye direction. Note that the batsman must take his eye off the ball by about 5º in order to anticipate the bounce.

By knowing the time from delivery to bounce, and the declination angle of the bounce point, a batsman can determine when the ball will reach the bat, and the height it will be at that time. This information will allow him to produce an accurate, well-timed stroke. We found that all players made an early, reasonably accurate, anticipatory saccade to the bounce point, arriving there about 100ms before the ball. Good players have a latency of only about 130 ms for such saccades, but weaker players may be up to 100 ms later.

Reference

*Land MF, McLeod P (2000) From eye movements to actions: how batsmen hit the ball. Nature Neuroscience **3**: 1340-1345*

Portrait Drawing and Sketching (D7)

Description

The task of producing a picture is very different from than simply looking at one. In drawing a portrait the artist has to acquire information from a sitter, formulate a line to be drawn and execute this on the drawing itself. There is thus a repeating sitter—drawing gaze cycle, with vision employed in different ways in each half cycle. In our laboratory we asked a painter and art teacher, Nick Bodimeade, to make some portrait sketches for us, as well as a longer more measured drawing, whilst wearing an eye tracker. The sketches took about 40 s and the drawing 4 minutes to complete (Fig 1.) The principal findings were that a typical cycle lasted 1.7 s (35 cycles per minute), with 0.8 s on the sitter and 0.9 s on the sketch (Fig. 2b). On average the pen made contact with the paper about 0.1 s after gaze transferred to the sketch, and lasted for the time gaze remained on the sketch.

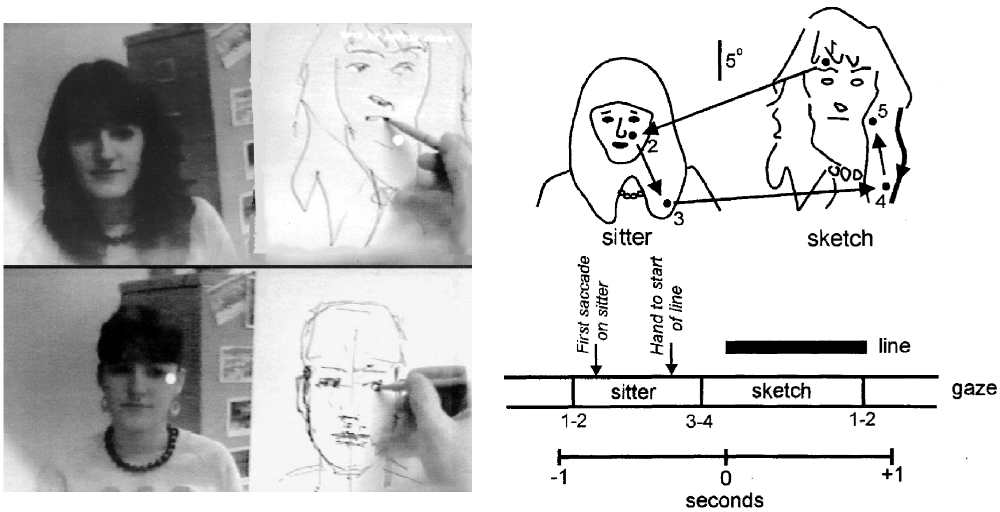


Fig 1 Left: frames taken from the eye camera on the artist head of a scene from the production of a sketch (top: fixation spot is on the sketch (right) about 2° below the pencil) and a more detailed drawing (bottom: fixation is close to the sitter's eye while the artist draws the eye). The sitter (and collaborator) is Genevieve Baker. The fixation spot has a 1° diameter; Right: timing of the events in the production of a sketch. Average drawing cycle derived from all cycles by indexing all events to the beginning of each drawn line. The numbers just below the time-line indicate transitions between the locations shown on the on the diagram of sitter and sketch.

It was possible to work out something of what was happening as the artist formulated his next line. Between one and four fixations were made on the sitter's face (mean 2.3), and by the last fixation the point to be addressed on the sketch had been selected. When gaze left the sitter, it was transferred accurately (< 2° error) to the corresponding point on the sketch. Interestingly, this was not the point that the next line was to be drawn from, but the point drawn to, i.e. the end of the line (Fig.1). This surprised both ourselves and the artist. It does, however, make some sense. In a sketch each line is a new entity, almost unrelated to the last. Thus start of the next line must be determined by some process of internal selection by the artist. The course of the line and its end-point, however, are derived from features on the sitter, once the start of the line has been established. The detailed drawing (Fig 1) was quite different from the sketch, and more complicated. Unlike the sketches, each line was not independent of the last, and often began where the last left off. There was checking in which no line was drawn and alterations to lines. There was also no clear targeting of the end of each line over the beginning, as there was in sketches. Thus whilst the sketches showed a rather clear progression of thought processes, this was much harder to discern in the more complex drawing.

Reference

Land MF (2006) Eye movements and the control of actions in everyday life. Prog Retinal & Eye Res ***25****: 296-324. (section 2.1.5).*

Jigsaw Puzzle (D8)

Description

In jigsaw puzzles the objective is to copy a model – the picture on the box – by assembling pieces drawn from a random source. They require pattern fitting at three levels. The pieces have to be recognised by pattern and colour as belonging to a part of the overall source pattern; they must be chosen so that their pattern fits in with some part of the puzzle already completed; and the outlines of suitable new pieces have to be examined and oriented so as to produce a mechanical fit with other existing pieces in the puzzle. The purpose of this study was to investigate how gaze was apportioned between the different sub acts as the puzzle progressed. The data gathered from the videos was the moment-by-moment location of gaze (50 Hz sampling) and its relation in time and space to the movements of the hands.

We examined the fixation patterns of two subjects as they completed a 50 piece puzzle of a cartoon scene. The overall cycle loosely followed the model-source-copy pattern of the block copying task of Ballard et al. (1993), with the pick-up and place movements of pieces occurring near the end of the fixations of the source and the beginning of those on the copy. Figure1 illustrates an interesting 30 s episode during which the player has two loose pieces (*a* & *b*) which he eventually joins and fits to the completed part of the puzzle.

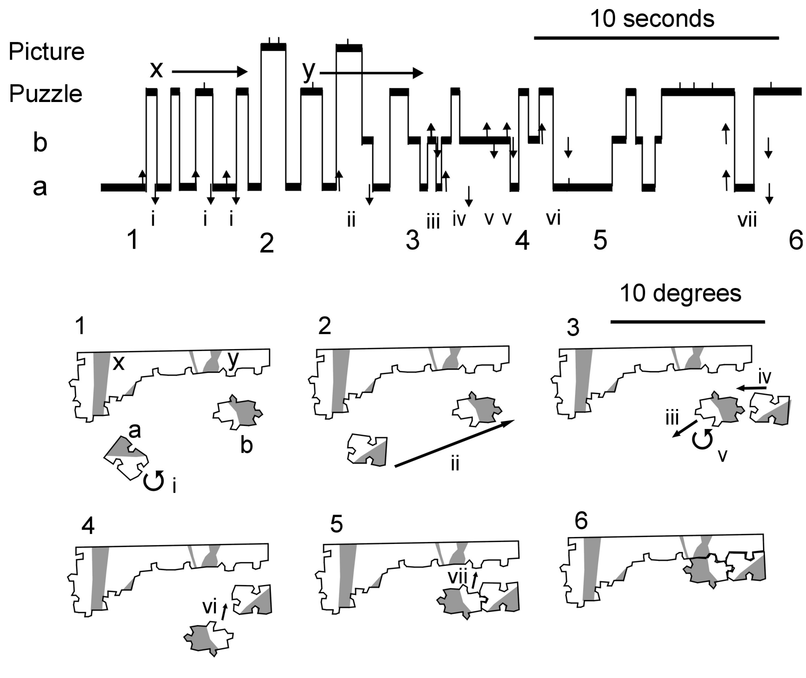


Fig 1 Gaze movements (above) and movements of the pieces (below) during a 30s episode of jigsaw puzzle solving.

This episode illustrates a number of general points. First, the eyes only fixate the parts of the field that are important – the two pieces, the relevant regions of the part-completed puzzle, and on two occasions the relevant region of the picture. When looking at the individual pieces *a* and *b* fixation is accurate: the mean fixation distance from the centre of each piece (approximately 2 by 3º) was 1.4º. Second, during the various movements of the pieces they are either fixated during the move, or during the half second before the move. This seems to bear out both the 'do it where I'm looking' and the 'just in time' rules of Ballard et al (1993). Third, comparisons between patterns and outlines are mostly made by looking from one element to the next and back again, rather than sizing up the situation from a single gaze location. This is particularly clear between 1 and 2 where gaze moves repeatedly from *a* to the completed part of the puzzle and back again as *a* is rotated. This leads to the rejection of the hypothesis that *a* can fit to *x*, and the development of the new idea that it might fit near *y*, which is then checked by consulting the picture. Similar 'aha' moments precede the move at (vi), again accompanied by much to-and-fro checking, and also before move (vii), although here the relevant profiles are now within 2º of each other, and there is a lull in the overt cross-checking seen elsewhere.

Reference

There are no publications associated with this clip. The study was performed in 1998 with the help of two University of Sussex Students, Lynette Owen and Sam Walker.

See also:

Ballard DH, Hayhoe MM, Li L, Whitehead SD (1993) Hand eye coordination during sequential tasks. Phil Trans R Soc Lond B ***337****: 331-339*

Party (D9)

Description

This early video as made for fun in 1996. It documents the subject (M.F.L) having various conversations with people and generally looking around. The idea was to see to what extent the conclusions reached by Alfred Yarbus (1967) about the way we view faces were true for active vision in a real setting. (Yarbus' recordings were of subjects viewing pictures in head-restrained conditions). The main points of interest in the video are first that faces seem to be located effortlessly (in one saccade) from considerable angular distances, and second that they are indeed viewed much as Yarbus had shown. That is, gaze concentrates on the triangle of the eyes and mouth, as shown in Fig 1. It is also interesting that gaze occurs in bouts, with the subject concentrating intently while conversing, but then looking rapidly around the viewing various other people before returning to the same or a different conversation.

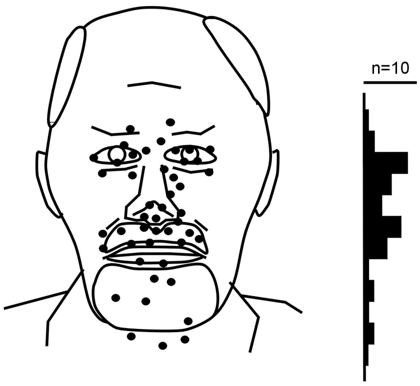


Fig 1. Fixation locations on the head of a colleague with whom the subject was holding a conversation. The total viewing time was 30s, and the head height varied in angular size between 10 and 15 º depending on distance. Graph on the right shows that gaze is concentrated on the regions of the eyes and the mouth, presumably because these are the most informative.

Reference

There are no publications associated with this video. Comparison can be made between Fig 1 and the picture of "The Girl from the Volga" in:

Yarbus, A., 1967. Eye Movements and Vision. Plenum Press, New York.

Table Tennis (D10)

Description

Ripoll et al. (1987) found that international table-tennis players anticipate the bounce and make a saccade to a point close to the bounce point. Land and Furneaux (1997) confirmed this (with more ordinary players) using an eye tracker. They found that shortly after the opposing player had hit the ball the receiver made a saccade down to a point a few degrees above the bounce point, anticipating the bounce by about 0.2 s (Fig. 1c). At other times the ball was tracked around the table in a normal non-anticipatory way; tracking was almost always by means of saccades rather than smooth pursuit. The reason why players anticipate the bounce is that the location and timing of the bounce are crucial in the formulation of the return shot. Up until the bounce the trajectory of the ball as seen by the receiver is ambiguous. Seen monocularly, the same retinal pattern in space and time would arise from a fast ball on a long trajectory or a slow ball on a short one (Fig. 1a). (Whether either stereopsis or looming information is fast enough to contribute a useful depth signal is still a matter of debate). This ambiguity is removed the instant the timing and position of the bounce are established. Therefore the strategy of the player is to get gaze close to the bounce point (this cannot and need not be exact) before the ball does, and lie in wait. The saccade that affects this is interesting in that it is not driven by a 'stimulus', but by the player's estimate of the location of something that has yet to happen. Much the same thing happens in cricket.

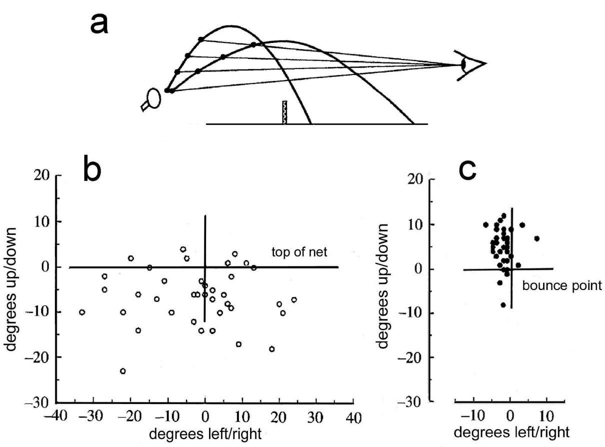


Fig 1. (a) The visual ambiguity in the trajectory of an approaching ball before it bounces. The vertical motion of a slow ball bouncing short, and a faster ball bouncing long will appear similar to an observer. The ambiguity is removed when the ball bounces; (b,c) The locations in the field of view of the receiver of 38 fixations which follow the first saccade after the ball has been struck by the opponent: (b) relative to the table top, and (c) relative to the bounce point. The receiver mainly fixates a point a few degrees above the expected bounce point, independent of where that is on the table.

Reference

Land MF, Furneaux S (1997) The knowledge base of the oculomotor system. Phil Trans R Soc Lond B 352: 1231-1239

Urban Driving (D11)

Description

In contrast to driving on country roads, town driving does pose great demands on steering to get round bends, but it does require vigilance in avoiding obstacles, other traffic and pedestrians, as well as attention to signs, traffic lights etc. We used an eye tracker to determine the visual strategy of a driver (MFL) negotiating a rather typical urban situation in the UK (Fig 1).

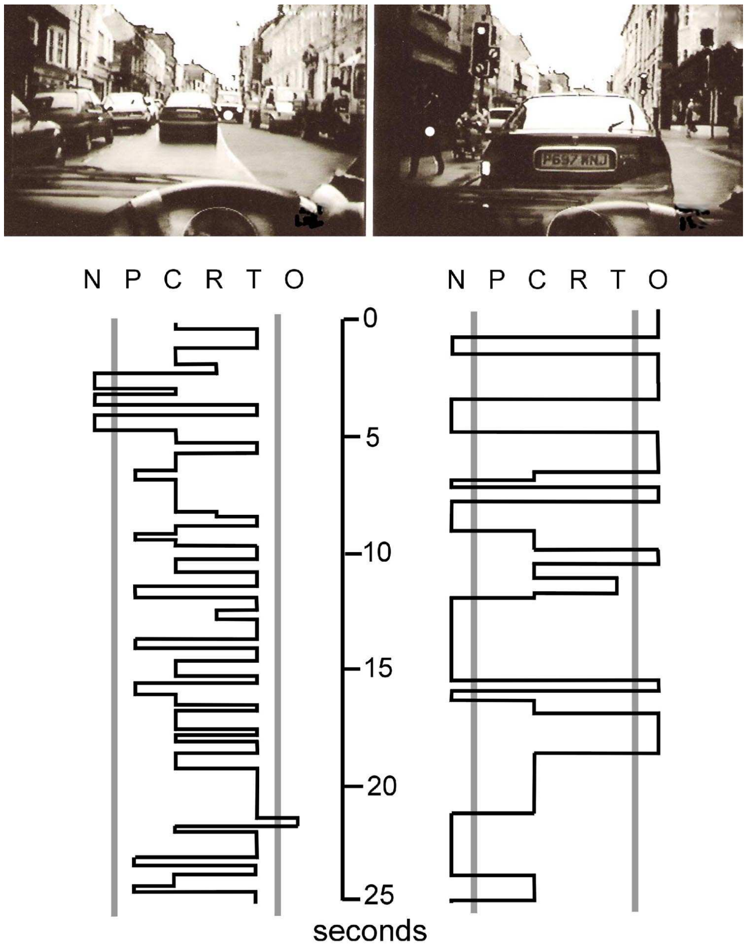


Fig 1. Record showing driver gaze distribution during two phases of the drive up the High Street: steering through dense traffic (left, gaze is on the oncoming car) and when stationary at traffic lights (right, gaze is on a pedestrian on the left). When driving almost all fixations are on objects on the roadway; when stationary almost all are off-road. N, near-side off-road; P, parked vehicles on near-side; C, car in front; R, on or above road surface; T, oncoming traffic; O, opposite side off-road. Grey lines are edges of roadway (kerbs). Each 'look' at a particular region usually involves more than one fixation.

Like many towns, the High Street in Lewes in Sussex is relatively narrow, has parked cars or delivery vans on either side, and can just manage two lanes of moving traffic. Even at low speeds, it requires considerable concentration for the driver to stay within the 'field of safe travel'. In the record in Fig.1 (left) the car is travelling through the traffic at about 7 m.p.h. Gaze movements alternate rapidly and irregularly between the car in front (35% of the time), the oncoming vehicles on the right (42%), and the edges of the parked cars on the left (10%). Hardly any time is spent looking off road (6 %) or at the road surface (7%). In the period between seconds 10 and 20 gaze shifts between different regions every half second, with either one or two fixations on each region. In Fig.1 (right) the car is stationary at traffic lights, and the fixation pattern is completely different. Most of the fixations (75%) are off-road on the left or right pavements, looking at shops and pedestrians. 22% are on the car in front and 3% on another vehicle. In this situation viewing is recreational, unlike the serious vigilance apparent when driving in traffic. The leisureliness of stationary viewing is also shown by the doubling of the time spent looking at each region (mean 1.14s) compared with the equivalent time when moving (0.54s). Other traffic situations produce different patterns of gaze distribution, and it is hard to make any overall generalizations. Drivers have a range of strategies to match the current driving conditions, varying between continuous two-per-second sampling of the potential obstacle field, to more relaxed road following in lighter traffic, with many more glances to the off-road surroundings.

The other videos in this compilation are of steering round right-angle corners on suburban roads. What is critical in getting these manoeuvres right is the timing of the steering action, both when entering and exiting from the corner. Using the view provided by the eye-tracker, it was possible to examine what timing cues were available in the half-second or so before the driver began to steer into and out of the bend. The changes in the appearance of the road-edge (kerb) seemed to be the only cues to provide useful timing information, and which also correlated reliably with the initiation of the steering action. In a left hand turn (nearside in the UK) the tangent point slips leftward as the corner approaches and steering starts when this angle reaches between 30º and 40º. The cue for straightening up at the exit of the bend seems to be rotation of the nearside kerb in the visual field . Just before the end of the bend the kerb angle rotates through the vertical in the driver's view, going from acute to obtuse. The change of steering direction occurred when this angle reached between 140 and 150º.

Reference

Land MF (2006) Eye movements and the control of actions in everyday life. Prog Retinal & Eye Res ***25****: 296-324. (section 2.3.4).*