Chapter 22
Industrial Engineering and Human Factors

Eye tracking offers a unique measure of human attentional behavior. This is particularly important in evaluating present and future environments in which humans do and will work. The examination of human interaction and behavior within their environments, particularly ones in which humans often perform functions critical to safety, is one topic studied by human factors and industrial engineers. Traditional measurement methods of human performance often include measures of reaction time and accuracy; e.g., how fast a person completes a task and how well this task is performed. These are generally measures associated with performance. To study the steps taken to perform the tasks requires analysis of the individual procedures performed. For this analysis, process measures are often needed. Eye movements are particularly interesting in this latter context because they present measures that can provide insights into the visual, cognitive, and attentional aspects of human performance.

Here, three broad experimental domains are presented in which eye tracking can play an important analytical role: aviation, driving, and visual inspection. With the development of sophisticated simulators for this task, and incorporation of eye trackers into these simulators, analysis of eye movements provides a powerful additional mechanism for measuring human factors.

22.1 Aviation

Since their early use in flight simulators (e.g., see Kocian (1987) and Longridge et al. (1989)), eye trackers have continued their utility in aviation experiments, ranging from testing procedural training to the evaluation of increasingly sophisticated deployment of new (graphical) displays. With improved simulator and eye tracking technology, the use of gaze recording techniques will undoubtedly continue.

An example of a recent combined use of relatively new eye tracking technology in a sophisticated flight simulator was reported by Anders (2001). Eye and head movements of professional pilots were recorded under realistic flight conditions in an

A.T. Duchowski, *Eye Tracking Methodology*, DOI 10.1007/978-3-319-57883-5_22

investigation of human–machine interaction behavior relevant to information selection and management as well as situation and mode awareness in a modern glass cockpit. The simulator used was the Airbus 330 full flight simulator certified for airline training. Microphones and video cameras inside the simulator recorded environmental sounds, voice communication, general cockpit settings, and the actions of the operators. Several flight trials were conducted, with each flight starting from level 210 (21,000 feet altitude) in managed descent mode including an ILS (Instrument Landing System) approach and landing. The captain's eye and head movements were recorded during trials for later Point Of Regard (POR) analysis within a previously defined 3D model of the environment. Areas of interest were chosen to augment eye movement analysis. Major displays and panels were selected, including Primary Flight Display (PFD), Navigation Display (ND), Engine and Warning Display (EWD), System Display (SD), Main Panel (MPn), Glare Shield (GSc), and Overhead + Central Pedestal (OCA). The Flight Control Unit (FCU) is part of the GSc whereas the two Multipurpose Control and Display Units (MCDU) in front were separately defined as CDU. MAP described the chart and other paper information, with OUT representing the window area. The predefined cockpit regions of interest are shown in Fig. 22.1.

Analysis of eye movements illustrates the importance of the PFD as the primary source of information during flight (the PFD is the familiar combined artificial horizon and altimeter display gauges for combined altitude and attitude awareness). Within the PFD, most of the fixation time is attributed to flight parameters presented

Fig. 22.1 A330 cockpit with predefined areas of interest. Courtesy of Geerd Anders

on the speed band (SPD), artificial horizon (AH), and altitude band (ALT). Cumulative fixation time on the heading band (HDG) is fairly low because the ND gives a better indication of horizontal situation. Average attention allocation to areas of interest is very similar for all tested approach flights. Between-subject differences (across pilots) are reportedly not outstanding. Attention shifts from the PFD to OUT shortly before landing. Attentional shifts were also recorded to the GSc in response to simulated Air Traffic Control (ATC) instructions to change flight parameters. Attention is high to the MAP and CDU areas during phases of flight in which the pilot is briefing the approach or when the flight plan changes (e.g., after a “Go direct...” or runway change instructions from ATC). As a proof of concept, this study shows the potential of eye movements for judgment of pilots’ performance and future training of novice pilots.

Besides general scanpath patterns in the cockpit, eye movements have also been used to evaluate the usability of specific instruments such as newly developed electronic maps. Electronic maps such as 3D graphical navigational displays are becoming increasingly viable alternatives to traditional maps in the cockpit (Ottati et al. 1999). In their study, Ottati et al. compared eye movement patterns on different terrain features between experienced and novice pilots during a Visual Flight Rules (VFR) simulation. Based on previous studies of pilots’ eye movements during Instrument Flight Rules (IFR) navigation, the authors expected experienced aviators to spend less time finding and fixating on their navigational landmarks, whereas novices were expected to have greater difficulty finding landmarks and extracting useful data from them, causing greater dwell times. As expected, a greater tendency in novice pilots was found to “fly out of the window” (i.e., devoting visual attention outside the cockpit) than in experienced pilots. Independent of expertise, pilots spent more time out of the window than over instruments.

The primary difference between experienced and novice pilots’ eye movements during VFR flight reported by Ottati et al. was the amount of significant information-gathering fixations exhibited by each group. Experienced pilots employed significantly more fixations, yet each fixation lasted for relatively the same duration. This may indicate that novices were ineffectively scanning their forward field of view (outside the window) to gain sufficient navigational information. Experienced pilots’ performance suggests that their fixations out of the window were more deliberate and informative. The authors suggest that if novices were trained in experienced pilot scan strategies, the evident performance gap between novices and experts could possibly be reduced.

In a study of electronic maps for taxiing, Graeber and Andre (1999) also suggest that training is necessary to assure proper usage of, and optimal visual attention interaction with Electronic Moving Maps (EMMs). The EMM studied by the authors is a display developed by NASA to increase navigation situation awareness, decrease navigation errors, increase forward taxi speeds, decrease planning time, decrease navigation mental workload, and improve navigation communications of pilots in low-visibility conditions. The main objective of Graeber and Andre’s study was to understand how pilots visually interact with the EMM. This objective was partially motivated by concerns of EMMs disproportionately drawing the pilot’s eyes into the cockpit during taxi, because taxiing is a primarily "eyes-out" task. Pilots navigate the airport and terminal areas using visual cues and signage from the outside environment, and through communication with ground control. Their task is to accurately navigate the cleared route, monitor for potential incursions, and maintain a safe distance from other aircraft, ground vehicles, and obstacles, a task best served by keeping their heads up and eyes out the cockpit windshield.

To examine pilots' visual interaction with the EMM and outside environment, Graeber and Andre examined pilots' eye movements under three different visibility conditions: daytime high visibility (Day Visual Meteorological Conditions, VMC), daytime low visibility (Day 700 Runway Visual Range, or RVR), and nighttime moderate visibility (Night 1400 RVR). Results indicate that visibility significantly affected the amount of time pilots dwelled on the EMM, with dwell time significantly higher in Day VMC than in Day 700 RVR (i.e., under high visibility conditions). That is, as visibility degrades, pilots spend more time eyes-out and less time dwelling on the EMM with no loss in taxi performance. A potential explanation for this surprising result is that pilots need to be eyes-out to maintain lateral and directional loop closure, scan for hazards, and maintain information gathering Out The Window (OTW). This explanation is supported by the increased percent time on route found in the performance data as the visibility decreased. As intended in the design of the EMM, it appears that pilots used the EMM as a secondary navigation aid across all visibility conditions. That is, when visibility degrades, the EMM appears to benefit performance the most, but without incurring additional visual attention requirements.

22.2 Driving

Another type of simulator in which eye trackers are increasingly being incorporated is the driving simulator. It is widely accepted that deficiencies in visual attention are responsible for a large proportion of road traffic accidents (Chapman and Underwood 1998). An understanding of the visual search strategies of drivers is thus extremely important, and much research has been conducted in this area. Several important driving studies incorporating an eye tracker are described in Underwood (1998).

Eye movement recording and analysis provide important techniques for understanding the nature of the driving task and are important for developing driver training strategies and accident countermeasures (Chapman and Underwood 1998). Chapman and Underwood indicate that at their simplest, drivers' fixation patterns on straight roads can be described as concentrating on a point near to the focus of expansion (the point in the visual field in front of the driver where objects appear stationary) with occasional excursions to items of road furniture and road edge markers. This reliance on the focus of expansion in the scene is assumed to be because it provides precise directional information to the driver and is the location near to which future traffic hazards are likely to be first visible. Increasing the complexity of the visual scene (by adding vehicles, road furniture, or irrelevant signing) increases the number of eye movements made and decreases the mean fixation durations on individual objects. This seems to be a natural response to having more objects available in the visual field to look at; it is not clear whether decreases in fixation durations mean that objects are processed less completely or that redundant fixation time is simply reduced.

Chapman and Underwood's study reports on recorded eye movements of relatively large groups of both novice and experienced drivers while watching videos of dangerous situations. The largest overall difference between the groups was that novices had generally longer fixation durations than experienced drivers in this task. Chapman and Underwood argue that this reflects the additional time required by novices to process information in the visual scene. Important differences were also found between the types of situation used in the study. Rural situations, even those chosen to be particularly dangerous, generally evoked fewer responses from subjects and longer fixation durations. In urban films both groups of subjects reported many more dangerous events and had shorter mean fixation durations. It is clear that dangerous events generally evoke long fixation durations. This result demonstrates the dangers in averaging eye movement data over periods of time in dynamic scenes. It is therefore argued that to fully understand the subtlety of such data, and to draw realistic conclusions about the cognitive process underlying observable behaviors, it is necessary to develop a detailed understanding of the moment-by-moment "syntax" of driving situations.

Dishart and Land (1998) discuss previous work showing that experienced drivers obtain visual information from two sections of their view of the road ahead, in order to maintain a correct position in lane while steering their vehicle around a curve. The more distant of these two sections is used to predict the road's future curvature. This section is optimally 0.75–1.00 s ahead of the driver and contains the tangent point. This section of road is used by a feedforward (anticipatory) mechanism that allows the driver to match the curvature of the road ahead. The other, nearer, section is about 0.5 s ahead of the driver and is used by a feedback (reactive) mechanism to "fine tune" the driver's position in lane. As either lane edge approaches the vehicle, the driver steers away from it, correcting his or her road position. Experiments using video-based eye–head tracking equipment have shown that the feedback mechanism is present in most people regardless of their driving experience (although its accuracy is higher in those with experience), but that the feedforward mechanism is learned through experience of steering tasks (that can include riding a bicycle, computer driving games, etc.).

Dishart and Land (1998) discuss eye–head tracking experiments on learner drivers during their tuition which indicate that use of the section of the road containing the tangent point increases with experience, then decreases as drivers learn to optimize their visual search patterns. This affords experienced drivers to spend more of their visual resources on other visual tasks both related and unrelated to driving.

A fundamental problem in visual search in driving research is defining and controlling demand on the driver as an independent variable (Crundall et al. 1998). Possible confounding factors are increases in visual demand, such as an increase in visual clutter or complexity, and increases in cognitive demand, such as an increase in the processing demands of a particular stimulus perhaps due to an increase in its relevance to a current context. Despite the lack of consistency in the manipulations of visual demand, it seems fairly well documented that general increases in task demands and visual complexity tend to reduce mean fixation duration and increase the sampling rate. Crundall et al. discuss the effects of cognitive and visual demand upon visual search strategies during driving, and whether they can be used to differentiate between novices who have recently passed their test and more experienced drivers. The authors review previous studies with the aim of identifying a process that may account for the effects of changes in demand according to driver experience.

In a driving study examining the effects of clutter, luminance, and aging, but not quite in a driving simulation, Ho et al. (2001) recorded subjects' visual search for traffic signs embedded in digitized images of driving scenes. Example driving stimulus images are shown in Fig. 22.2. In a visual search task, a traffic sign was first presented to subjects, followed by a scene in which subjects were instructed to search for the sign. Analysis of five dependent measures are reported: (a) errors, (b) reaction time, (c) fixation number, (d) average fixation duration, and (e) last fixation duration. For each measure the effects of age, clutter, luminance, and target presence were evaluated. On average, older adults were less accurate than younger ones. Accuracy for daytime scenes was independent of target presence, but errors for nighttime scenes were more common on target-present trials than on target-absent trials. In daytime scenes errors were generally more common in high clutter than in low clutter, whereas in nighttime scenes no consistent clutter effect was detected. Reaction time data indicated that younger adults generally responded more quickly than older adults. Reaction times were also dependent on the amount of clutter and the presence of the target. Fixation number data, a measure that is strongly correlated with reaction time, indicated that older adults made more fixations than did younger adults, and more fixations were needed for high clutter and target-absent scenes than for low-clutter and target-present scenes. Nighttime scenes with high-clutter also required more fixations than did daytime scenes. Examining average fixation duration, it was found that younger adults had shorter fixation durations than did older adults, and low-clutter scenes resulted in shorter fixations than did high-clutter scenes. Relative to daytime scenes, high-clutter nighttime scenes containing a target resulted in longer average fixation durations. The last fixation duration reflects the comparison of the last fixated object with the target representation and the terminal decision regarding target presence. Age differences in this last stage of processing might accrue because of difficulties with the comparison process itself or because the elderly are more cautious in making overt responses. These data showed that age differences were greater on target-present trials than on target-absent trials. In summary:

- High-clutter scenes required longer latencies and more fixations to acquire the sign, were associated with more errors, and had longer fixation durations.
- Luminance effects were negligible, although high-clutter nighttime scenes did impair search efficiency on reaction time and fixation number measures. In addition, search was impaired on target-absent trials for both age groups.

(a) Daytime.

(b) Simulated nighttime.

Fig. 22.2 High-clutter driving stimulus images. Reprinted with permission from *Human Factors*, Vol. 43, No. 3, 2001. Copyright 2001 by the Human Factors and Ergonomics Society. All rights reserved

- Older adults were found to be less accurate and slower than younger adults and executed more eye movements to acquire signs. The age effects on reaction time and fixation number were more pronounced on target-absent trials. Older adults used the visual cues that discriminated targets and distractors to quickly isolate the target on target-present trials. However, they took longer to correctly decide that a target sign was not present on target-absent trials, and they also took longer to make a terminal decision regarding the sign's match to the trial target on target-present trials. Interestingly, older adults were not more adversely affected by increased clutter than were the young.

Considering the rapidly aging driving population, the authors recommend that roadway engineers should consider reducing the number of competing signs (e.g., advertisements), avoiding redundant signs, and making traffic signs that are central to the safety of the driver more conspicuous.

Experiments where gaze is monitored in a simulated driving environment demonstrate that visibility of task-relevant information depends critically on active search initiated by the observer according to an internally generated schedule, which depends on learned regularities in the environment (Hayhoe et al. 2002). The driving simulator used by Hayhoe et al. consists of a steering station mounted on top of a 6 Degree Of Freedom (6 DOF) hydraulic platform. Steering, accelerator, and brake are instrumented with low-noise potentiometers that modulate signals read by a special-purpose 12-bit A/D board on an SGI VME bus. Subjects drive in "PerformerTown", an extensible environment designed by Silicon Graphics. Cars and trucks have been added to the environment that move along predefined or interactively controlled paths. The visual environment is easily modified to simulate lighting at any time of the day or night, along with the addition of fog. Observers view the display in a Head-Mounted Display (HMD) and eye position is monitored using a video-based corneal reflection eye tracker. This simulator has been used to examine "change blindness", a phenomenon where observers are insensitive to many changes made in scenes either during a saccadic eye movement or some other masking stimulus. In one experiment, while subjects drove along a prespecified path following a lead car, in a particular block a No Parking sign changed into a Stop sign for about 1 second as the observer approached it. The retinal transient caused by the change was masked by 100 ms blank screens before and after the change. Results indicate that the likelihood of fixating the sign at any point while it is there is heavily modulated by the task and by the location of the sign. Given a "follow" instruction, observers rarely fixate the sign. When driving normally they invariably respond when it is at an intersection, but miss it two thirds of the time when it is in the middle of the block. This suggests observers must actively initiate a search procedure in order to see particular stimuli in the scene. Markedly different fixation patterns were displayed in the two instructions. In the "drive normally" conditions, subjects spent much more time fixating in the general region of the intersection. This strongly suggests fixation patterns and attentional control in normal vision is learned. Furthermore, the visibility of traffic signs depends on active search according to an internally generated schedule, and this schedule depends both on the observer's goals and on learned probabilities about the environment. The experiment suggests that a fuller understanding of the mechanisms of attention, and how attention is distributed in a scene, needs to be situated in the observer's natural behavior.

Research on the effects of mental activity during driving suggests the convenience of raising drivers' awareness about the possible consequences of driving while their attention is focused on their own thoughts, unrelated to driving (Recarte and Nunes 2000). Recarte and Nunes studied the consequences of performing verbal and spatial-imagery tasks on visual search when driving. On each of four routes (two highways and two roads), participants performed two verbal tasks and two spatial-imagery tasks while their eye movements were recorded. The same results were repeated on all routes. Pupillary dilation indicated similar effort for each task. Visual functional-field size decreased horizontally and vertically, particularly for spatial-imagery tasks. Compared with ordinary driving, fixations were longer during the spatial-imagery task. With regard to driving performance, glance frequency at mirrors and speedometer decreased during the spatial-imagery tasks. Performing mental tasks while driving caused an increased attentional workload on ordinary thought, as shown by pupillary dilation. Mental tasks imposed by the experimenter while the participant was driving, and, therefore, of a more mandatory nature than ordinary thoughts, produced: (a) marked changes in the visual inspection patterns; (b) qualitatively different changes, depending on the type of processing resources required by the mental tasks; (c) the same effects in the four different driving scenarios; and (d) changes in practical driving behaviors, such as inspection reduction of mirrors and speedometer. Spatial-imagery tasks produced more marked effects in almost all the analyzed variables than did verbal tasks. First, an increment in mean fixation durations was found, due to some long fixations and possibly associated with mental image inspection as part of the mental activity of image search or rotating. It is suggested that these eye freezing responses produce impairment of environment perception. Second, a marked reduction of the visual inspection window was found, both horizontally and vertically, possibly associated with narrowing of the attentional focus size. Smaller saccadic size and marked reduction of glance frequency at mirrors and speedometer were also noted. With regard to the implications for driving, Recarte and Nunes suggest that the spatial reduction of the visual inspection window, including the reduction of the inspection of mirrors, could be interpreted as a predictor of decreased probability of detecting traffic events, particularly when performing mental spatial-imagery tasks.

It has been experimentally demonstrated that the pattern of eye fixations reflects, to some degree, the cognitive state of the observer (Liu 1998). Liu suggests that the analysis of drivers' eye movement may provide useful information for an intelligent vehicle system that can recognize or predict the driver's intention to perform a given action. Such a system could improve the interaction between the driver and future vehicle systems and possibly reduce accident risk. Liu reviews numerous studies that have tried to establish the level at which the relationship between eye movements and higher cognitive processes can be modeled. The least controversial conclusions simply state that the pattern of eye movements generally reflects the observer's thought processes, indicating to some degree the goals of the observer and perhaps even the main areas of interest. The strongest conclusions assert that the eye movements are directly observable indicators of underlying cognitive processes, revealing the nature of the acquired information as well as the computation processes. To implement a "smart car," one which would be able to predict or recognize the driver's intentions and then take the appropriate course of action based on that prediction, Liu suggests that it may be possible to determine the current driver state (e.g., centering of the car in the current lane, checking whether the adjacent lane is clear, steering to initiate a change in heading, centering the car in the new lane, etc.) from observed eye movement patterns via hidden Markov dynamic models (HMDMs). The primary consideration for including driver eye movements in HMDMs is how the eye movement behavior will be coded (e.g., by gaze locations or by gaze or saccade direction). Gaze location seems to be the natural choice given that characteristic patterns of driver eye movements have been identified using a Markovian analysis of gaze location. The results from preliminary experiments using HMDMs without utilizing eye movements are promising. The addition of eye movement analysis should enhance the system performance by enabling recognition of driver intentions rather than just recognition of maneuvers as they begin.

22.3 Visual Inspection

In a study of visual inspection of integrated circuit chips Schoonard et al. (1973), state that "visual inspection pervades the lives of all people today. From poultry, meat, and fish inspection, to drug inspection, to medical X-ray inspection, to production line inspection, to photo interpretation, the consequences of inspection directly affect people's lives through their effects on the quality and performance of goods and services."

Important criteria of evaluating the efficiency of human visual inspection include measurements of performance, as indicated by the inspector's speed and accuracy. These performance measures effectively summarize general outcomes of the process of (visual) inspection. In contrast, eye movements captured during visual inspection provide visualization of the inspector's process, and therefore provide an instance of process measures. Important eye movement-related process measures include fixation durations, number of fixations, interfixation distances, distribution of fixations, fixational sequential indices (order of fixations), and direction of fixations. The utility of process measures relies on their relation to, or explanation of, performance. Thus one expects to find a relation between process and performance measures.

With respect to performance, a key issue is the ability to gauge the visual search process with respect to some identifiable baseline measurement. For example, for training or evaluation purposes, it is desirable to be able to compare an inspector's visual search process to that of an expert inspector, or to some ideal model of performance. Visual search can be described by a two-stage process:

1. A time-consuming visual search for a target
2. A decision process identifying the found item as either target or nontarget.

Visual search can be modeled by two idealized processes: a perfect-memory systematic (serial) search, or a memoryless random search. A systematic search differs from a random search principally in the manner of visitation of searchable regions. That is, systematic search, due to its perfect memory, will visit each portion of the entire search region only once. In contrast, under random search, every portion of the search area has equal probability of being visited at any time during the search. Inasmuch as there is no memory of past visitations, regions may be visited more than once (re-inspected).

To model random visual search more formally, consider a visual search area A, with randomly distributed search targets (symbolized by $\times$), and a visual lobe (instantaneous fixation size) with area a, as shown in Fig. 22.3. Let the probability of detecting a target (e.g., a defect) given a fixation at a particular location be s; i.e.,

$$p(\text{detecting defect} \mid \text{fixation}) = s, \;\; s \in (0, 1).$$

Then let the probability that a fixation contains (or lands on) a target be $n_d * (a/A)$; i.e.,

$$p(\text{fixation contains a defect}) = n_d * \frac{a}{A},$$

where n_d is number of randomly distributed defects. Thus, the probability of detecting a defect is:

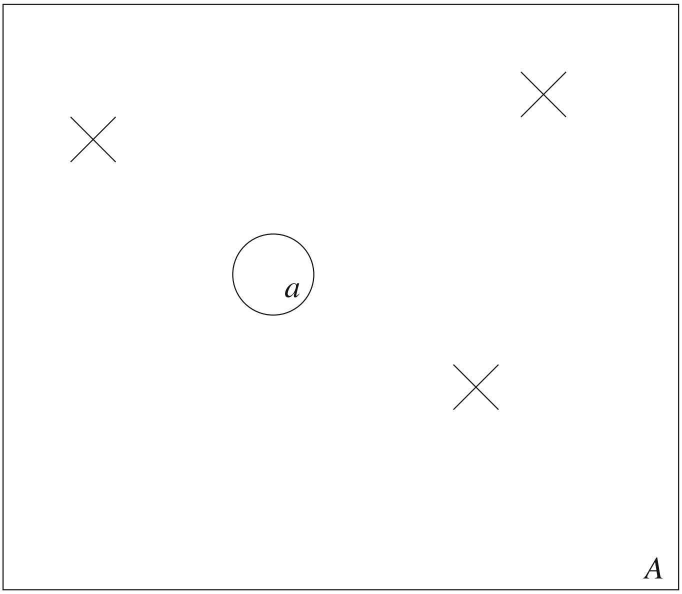

Fig. 22.3 Model of visual search area, visual lobe, and targets

$$p(\text{detecting a defect}) = n_d * \frac{a}{A} * s = Q$$

on any single fixation. The probability of detecting a defect on any ith fixation, in a series of fixations, is:

$$p(\text{detecting a defect on } i\text{th fixation}) = (1 - Q)^{i-1} Q,$$

where $(1 - Q)$ is the probability of not detecting the fault on previous fixations. Thus the cumulative probability of detecting a defect over n fixations is:

$$\begin{aligned} Q \sum_{k=1}^{n} (1 - Q)^{k-1} &= \left[1 - (1 - Q)^n\right] = 1 - e^{-Qn} \\ &= 1 - e^{[n_d*(a/A)*s*n]}. \end{aligned}$$

Considering the time needed to search t, and the time devoted per fixation t', letting $n = t/t'$, the cumulative probability can be rewritten as

$$\begin{aligned} Q &= 1 - e^{\left[n_d*(a/A)*s*t/t'\right]} \\ &= 1 - e^{-\lambda t}, \text{ with } \lambda = n_d * \frac{a}{A} * \frac{s}{t'}. \end{aligned}$$

Q is an exponential distribution that approaches probability 1 as search time increases. A stylized plot of the exponential distribution is shown in Fig. 22.4 (lowest solid curve). The model essentially predicts that with more time, eventually all defects (or almost all defects) will be detected. An important attractive feature of this model is that it predicts the accuracy of detection during the first few seconds of search, meaning that the probability of detection will not change much later on in the search. In other words, an inspector will either see or not see the defect fairly quickly, but the probability of detection will not change much with increased search time.

To illustrate the predictive property of the model, consider three different search durations: $t_1 = 3$ s, $t_2 = 12$ s, and $t_3 = 10$ s. The mean search time is approximately $25/3 = 8$–10 s. The cumulative probability function for this search can be modeled by the function $1 - e^{-t/10} = 1 - e^{-0.1t}$, and is plotted in Fig. 22.5. This particular instance of the visual search model would predict (roughly) that about 10 s are needed to achieve a 63% accuracy rate. A 95% success rate is reached at about 30 s.

The goal of modeling visual search in a visual inspection task is to be able to gauge an inspector's performance versus an idealized model of visual search such as the random search derived above. In contrast to this memoryless model, a perfect-memory systematic search is expected to reach a higher accuracy level faster than the random model. A systematic search model may be described by a piecewise-linear function as shown in Fig. 22.4 (highest solid curve). The systematic search is simply another idealized model, one that happens to be very efficient in comparison with the random model. Human performance is expected to fall somewhere in between

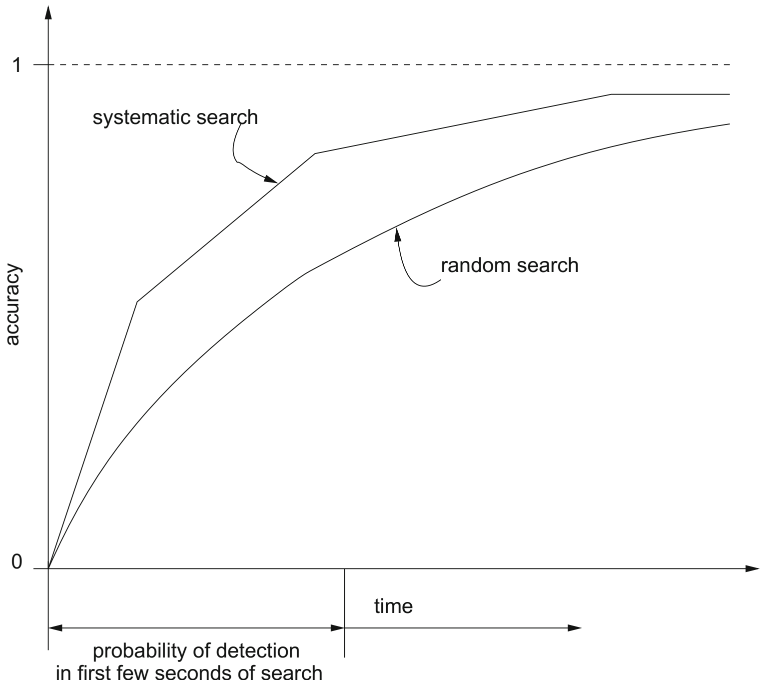

Fig. 22.4 Models of visual search

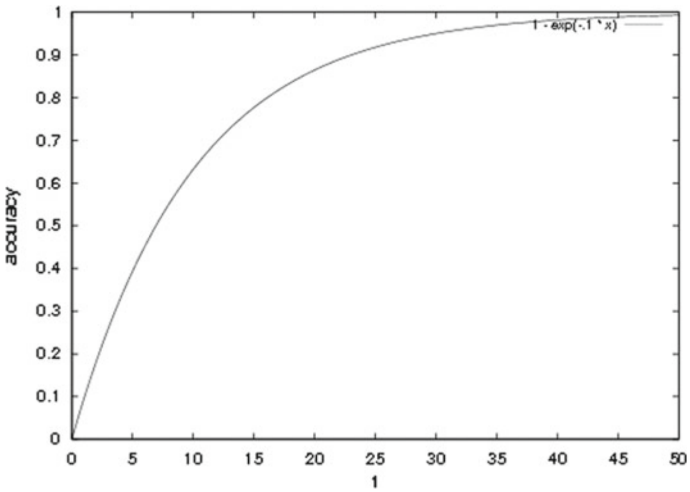

Fig. 22.5 Example of visual search model

the two idealized models. That is, when a cumulative distribution of human visual search performance is plotted, it should appear somewhere between the two curves, as indicated by the dashed curve in Fig. 22.4. Using this methodology, it is then possible to gauge the efficiency of an individual inspector: the closer the inspector's curve is to the idealized systematic model representation, the better the inspector's performance. This information can then be used to either rate the inspector, or perhaps to train the inspector to improve her or his performance.

Tracking eye movements during visual inspection may lead to similar predictive analyses, if certain recurring patterns or statistics can be found in collected scanpaths. For example, an expert inspector's eye movements may clearly exhibit a systematic pattern. If so, then this pattern may be one that can be used to train novice inspectors. In the study of visual inspection of integrated circuit chips, Schoonard et al. (1973) found that good inspectors are characterized by relatively high accuracy and relatively high speed, and make many brief eye fixations (as opposed to fewer longer ones) during the time they have to inspect.

In a survey of eye movements in industrial inspection, Megaw and Richardson (1979), identify the following relevant eye movement parameters.

- Fixation times. The authors refer to the importance of recording mean fixation times (perhaps fixation durations would be more appropriate). Megaw and Richardson state that for most inspection tasks, the expected average fixation duration is about 300 ms. Longer fixation times are associated with the confirmation of the presence of a potential target and with tasks where the search times are short.
- Number of fixations. The number of fixations is a much more critical parameter in determining search times than fixation times and is sensitive to both task and individual variables.
- Spatial distribution of fixations. The coverage given to the stimulus material can be found by measuring the frequency of fixations falling in the elements of a grid superimposed over the display or by finding the frequency of fixations falling on specific features of the display. In both cases these frequencies correlate with the informativeness of the respective parts of the display as revealed by subjective estimates made by the searchers.
- Interfixation distances. With static displays this measure is equivalent to the amplitude of the saccadic movements. It is possible that when a comparatively systematic strategy is being employed, interfixation distances may reflect the size of the useful field of view (visual lobe).
- Direction of eye movements. Horizontal saccades may occur more frequently than vertical ones, which may reflect the elliptical shape of the effective useful field of view (visual lobe).
- Sequential indices. The most popular of these is the scanpath.

In their survey, the authors review previous inspection studies where eye movements were recorded. These include inspection of sheet metal, empty bottles, integrated circuits, and tapered roller bearings. In an inspection study of tin-plated cans, the authors report that experienced inspectors exhibited smaller numbers of fixations and that each inspector used the same basic scanpath from one trial to the next although there were small differences between inspectors. In an inspection study of electrical edge connectors, it was noted that the number of fixations per connector was much greater for complex items (i.e., more fixations are needed to search over a complex target).

Besides being useful for gauging inspection performance, eye movements may play a part in training visual search strategy (Wang et al. 1997). Visual search strategy training can be effective in adoption of a desirable systematic search strategy. To train visual search strategy, eye movements may be used as both a feedback mechanism and as confirmation of adoption of the new search strategy. In their paper on search strategy training, Wang et al. recorded eye movements over visual inspection of artificial printed circuit boards. Eye movements were used to check how well each subject's search pattern followed the strategy trained. Scanpaths were judged by the experimenter after each training trial and feedback was given to the subject regarding whether she or he had searched in a random or systematic manner.

Because training has been identified as the primary intervention strategy in improving inspection performance (see Gramopadhye et al. 1998), Duchowski et al. (2001) investigated the utility of eye movements for search training in virtual reality. A three-dimensional aircraft inspection simulator was developed at Clemson University for this purpose. The aesthetic appearance of the environment is driven by standard graphical techniques augmented by realistic texture maps of the physical environment. The user's gaze direction, as well as head position and orientation, are tracked to allow recording of the user's fixations within the environment. The diagnostic eye tracking VR system allows recording of process measures (head and eye movements) as well as performance measures (search time and success rate) during immersion in the VR aircraft inspection simulator. The VR simulator features a binocular eye tracker, built into the system's Head-Mounted Display (HMD), which allows the recording of the user's dynamic Point Of Regard (POR) within the virtual environment. A user wearing the HMD and an example of raw eye tracker output are shown in Fig. 22.6. An experiment was conducted to measure the training effects of the VR aircraft inspection simulator. The objectives of the experiment included: (1) validation of performance measures used to gauge training effects, and (2) evaluation of the eye movement data as cognitive feedback for training. Assuming eye movement analysis correctly identifies fixations and the VR simulator is effective for training (i.e., a positive training effect can be measured), the number of detected fixations is expected to decrease with the adoption of an improved visual search strategy (Drury et al. 1997) (e.g., following training). The criterion task consisted of inspecting the simulated aircraft cargo bay in search of defects. Several defects can occur in a real environment situation. Three types of defects were selected to create inspection scenarios:

1. Corrosion: represented by a collection of gray and white globules on the inner walls of the aircraft cargo bay and located roughly at knee level.
2. Cracks: represented by a cut in any direction on the structural frames inside the aircraft cargo bay.

(a) User wearing eye tracking HMD.

(b) Raw eye tracker output (left eye).

Fig. 22.6 Virtual aircraft inspection simulator

3. Damaged conduits: shown as either broken or delaminated electrical conduits in the aircraft cargo bay.

Data for performance and cognitive feedback measures were collected using search timing and eye movement information, respectively. The following performance measures were collected.

1. Search time from region presentation to fault detection
2. Incremental stop time when subjects terminated the search in a region by deciding the region does not contain faults
3. Number of faults detected (hits), recorded separately for each fault type
4. Number of faults that were not identified (misses).

Fixation analysis enabled the collection of cognitive feedback measures, which were provided to subjects during the training session. Cognitive feedback measures were based on the eye movement parameters that contribute to search strategies as defined by Megaw and Richardson (1979), including: (1) total number of fixations, (2) mean fixation duration, (3) percentage area covered, and (4) total trial time. Cognitive feedback measures were graphically displayed off-line by rendering a 3D environment identical to the aircraft cargo bay that was used during immersive trials. This display represented the scanpaths of each trial to indicate the subject's visual search progression. An example scanpath is shown in Fig. 22.7.

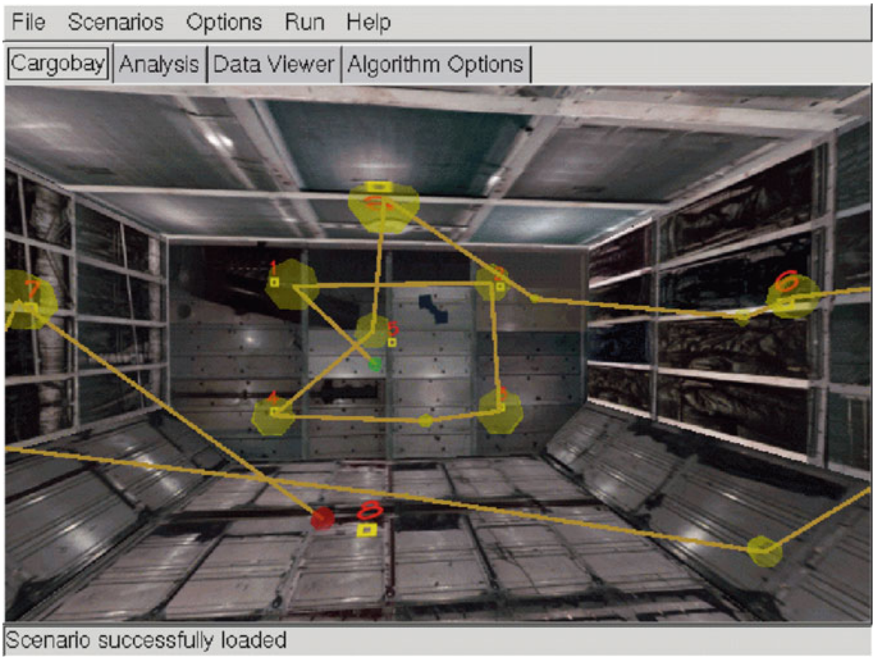

Fig. 22.7 Visualization of 3D scanpath in VR

Analysis indicates that, overall, training in the VR aircraft simulator has a positive effect on subsequent search performance in VR, although there is apparently no difference in the type of feedback given to subjects. Cognitive feedback, in the form of visualized scanpaths, does not appear to be any more effective than performance feedback. It may be that the most effective common contributor to training is the immersion in the VR environment, that is, the exposure to the given task, or at least to the simulated task. Whether the eye tracker, by providing cognitive feedback, contributes to the improvement of inspection performance is inconclusive. Users may benefit just as much from performance feedback alone. However, the eye tracker is a valuable tool for collecting process measures. Analysis of results leads to two observations. First, mean fixation times do not appear to change significantly following training. This is not surprising because eye movements are to a large extent driven by physiology (i.e., muscular and neurological functions) and cognitive skill. In this case the search task itself may not have altered cognitive load per se, rather, prior experience in the simulator may have facilitated a more efficient search. Second, the number of fixations appears to decrease following training. These results generally appear to agree with the expectation of reduced number of fixations with the adoption of an improved visual search strategy (e.g., due to learning or familiarization of the task). The implication of reduced number of fixations (without an increase in mean fixation time) suggests that in the posttraining case, subjects tend to employ a greater number of saccadic eye movements. That is, an improved visual search strategy may be one where subjects inspect the environment more quickly (perhaps due to familiarity gained through training), reducing the time required to visually rest on particular features.

Alternatively, fixation duration may be influenced by feedforward training, as follow-on studies have shown (Sadasivan et al. 2005). These more recent results suggested that given feedforward training (e.g., "You should look here"), participants significantly increased their mean fixation duration over those who did not benefit from training. Training benefit was observed in improved accuracy, leading to a classic speed–accuracy tradeoff. The trainees' slower-paced inspection strategy may have stemmed from a more deliberate target search/discrimination strategy that required more time to execute (see the relevant case study report in Chap. 19).

Recently, Reingold et al. (2002) reviewed the motivation behind using chess as an ideal task environment for the study of skilled performance. Because pioneering work showing perception and memory to be more important differentiators of expertise than the ability to think ahead in the search for good moves, chess research has been instrumental in enhancing our understanding of human expertise and in contributing to the study of artificial intelligence. The chess master is thought to use recognizable configurations of pieces, chunks, and templates as indices to long-term memory structures that, in association with a problem-solving context, trigger the generation of plausible moves for use by a search mechanism. Search is thereby constrained to the more promising branches in the space of possible moves from a given chess position. Hence, Grandmasters, the best human players, can find excellent moves despite generating only a small number of potential states (perhaps 100 or so for a few minutes of search). Such constrained search differs sharply from the enormous space explored by computer chess programs, which typically explore millions to hundreds of millions of alternatives in the same timeframe.

Reingold et al. study visual span as a function of chess skill (expert vs. intermediate vs. novice) and configuration type (chess configuration vs. random configuration) using a gaze-contingent window technique. The paper extends classic work demonstrating that after viewing structured, but not random, chess positions for five seconds, chess masters reproduced these positions much more accurately than lesser-skilled players. The authors document dramatically larger visual spans for experts while processing structured, but not random, chess positions. In a check detection task, where a minimized 3×3 chessboard containing a king and potentially checking pieces were displayed, experts made fewer fixations per trial, and had a greater proportion of fixations between individual pieces, rather than on pieces. These results provide strong evidence for a perceptual encoding advantage for experts attributable to chess experience, rather than to a general perceptual or memory superiority.

22.4 Summary and Further Reading

To summarize the potential impact of eye trackers in human factors research, this chapter focused on the eye tracker's capability of recording human visual process measures during exemplar tasks in aviation, driving, and during visual inspection. Performance measures typically quantify how a person performed (e.g., with what speed and accuracy), however, process measures cannot only corroborate performance gains, but can also lead to discoveries of reasons for performance improvements (i.e., what the subject performed). In particular, tracking the users' eyes can potentially lead to further insights into the underlying human cognitive processes under varying conditions and workloads.

Good sources for further information dealing with eye tracking and human factors include the Human Factors and Ergonomics Society (HFES) journal and annual conference proceedings, and the proceedings of SIGCHI, the European Conference on Eye Movements (ECEM), and the U.S.-based Eye Tracking Research and Applications (ETRA).