ELSEVIER

Image and Vision Computing 20 (2002) 291–305

image and vision COMPUTING

www.elsevier.com/locate/imavis

Visual search: psychophysical models and practical applications

Guang-Zhong Yang*, Laura Dempere-Marco, Xiao-Peng Hu, Anthony Rowe

Department of Computing, Imperial College of Science, Technology and Medicine, University of London, 180 Queens Gate, London SW7 2BZ, UK

Received 21 January 2001; received in revised form 10 December 2001; accepted 12 January 2002

Abstract

Extensive research into the role of saccadic eye movements in human visual perception has been carried out for many years. Although the search patterns of different observers while studying the same image bear some common characteristics, there are often variations in the temporal order in which fixation points are viewed. During visual search for a defined target, there is evidence for both parallel search, with which all objects are processed concurrently, and for sequential search, in which several fixation points are found leading to the target. In this article, we present a review of both theoretical and experimental research directed towards better understanding of the underlying mechanisms of visual search. We begin by looking at the basic dynamics of saccadic eye movements and some of the major psychophysical models that have been developed over the years. An overview of the practical applications and future trends of visual search is then provided. Visual search is a common task that people perform throughout their daily life, and the number of applications inspired by the human visual search mechanism is potentially large. The purpose of this paper is to highlight some of the key opportunities for the image and vision computing community and promote further interactions between biological and computational vision research. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Eye tracking; Visual search; Visual pattern recognition; Psychophysical modelling; Visual cognition

1. Introduction

Visual search is the act of searching for a target within a scene. If the scene is between 2 and 30°, the eyes will move across it to find the target. If the scene is larger still, the head moves as well. The myriad of visual search tasks performed in a single day is so large that it has become a reactive rather than deliberative process for most normal tasks. The quest for understanding how we look has fascinated both the art and research communities for centuries. Back in the 10th century, Alhazen, the Islamic scholar of optics, divided the process of vision into *aspectus* and *inuitio obtutus*, i.e. the first glance that provides a general idea of the scene, and scrutiny that consists of a series of individual observations leading to clear visual perception [1].

Visual search is a process that takes time, and the nature of the process is governed by our knowledge, interests, and expectations of the scene. Human eyes do not have a uniform visual response, in fact, the best visual acuity is only within a visual angle of 1–2°. This is called foveal vision, and for areas that we do not direct our eyes towards when observing a scene, we have to rely on a cruder representation of the objects offered by non-foveal vision, of which the visual acuity drops off dramatically from the centre of focus. Fig. 1 shows the basic structure of the human visual system and how visual acuity changes. When we try to understand a scene, we do not scan it randomly. Instead, we fixate our eyes on particular areas and move between them. The intrinsic dynamics of eye movement are complex, but the following movements are common:

- *Saccadic*: This is the voluntary rapid eye movement to direct the fovea to a specific point of interest.
- *Miniature*: These are a group of involuntary eye movements that cause the eye to have a dither effect. They include drift and micro-saccades.
- *Pursuit*: It is a smooth involuntary eye movement that acts to keep a moving object foveated.
- *Compensatory*: These movements are similar to the pursuit movement but they maintain a fixation while the head is moving.
- *Vergence*: This is how the eyes converge to a single point, relating to the focusing on a near or far object.
- *Optokinetic*: This is an involuntary saw-tooth movement that the eye performs when observing repeated moving patterns.

---

* Corresponding author. Tel.: +44-20-7594-8441; fax: +44-20-7581-8024.

E-mail address: gzy@doc.ic.ac.uk (G.-Z. Yang).

PII: S0262-8856(02)00022-7

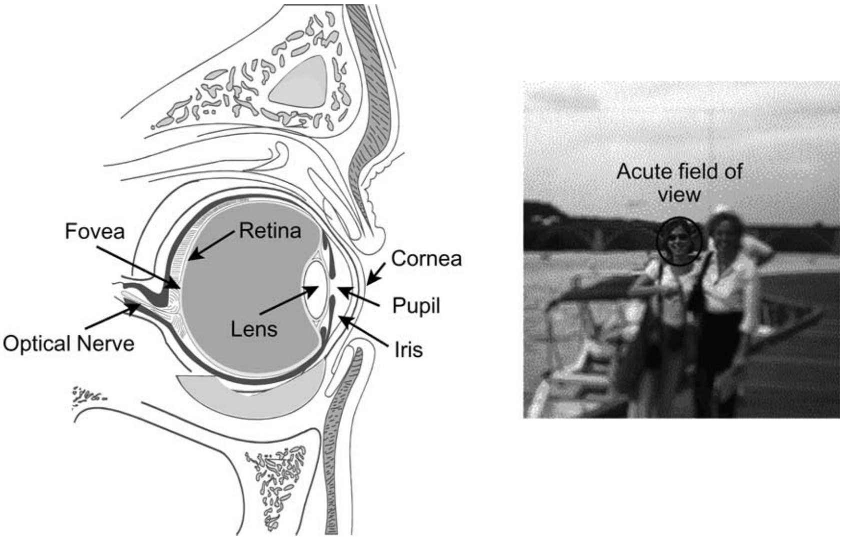

Fig. 1. Left: A schematic illustration of the structure of the human visual system. Right: Our retina does not have a uniform visual response and the best visual acuity is only within a visual angle of 1–2°, outside of which, as indicated by the circle, the visual acuity drops off dramatically.

Saccadic eye movement is the most important to consider when studying visual search. A saccade is a rapid voluntary movement from one point to another. It has a fast acceleration at approximately 40,000°/s^2 and a peak velocity of 600°/s. Saccades are observed in visual searches in the range of 1–40°. The objective of a saccade is to foveate a particular area of interest in a search scene. The fovea has an operation range of 0.6–1°. Miniature eye movements, on the other hand, are small and involuntary movements of less than 1°. There are three different types of miniature eye movements: eye-drift, micro-saccade and eye-tremor. The effect of these small eye movements imposes a lower limit on the accuracy of eye-tracking techniques, and an accuracy of more than 0.5° is generally thought to be not required.

The other types of eye movement are related to a group of involuntary motions that are used in following moving objects and maintaining reference points during head manoeuvres. During a visual search, a saccade moves the gaze of the eye to look at the current area of interest. This area of interest normally needs to be dwelled on for longer than 100 ms so that the brain can register what is in that area. When this happens, the point is called fixation. The objective of eye-tracking systems is to determine when and where fixations occur. The time order of the fixation points represents the actual visual search that takes place.

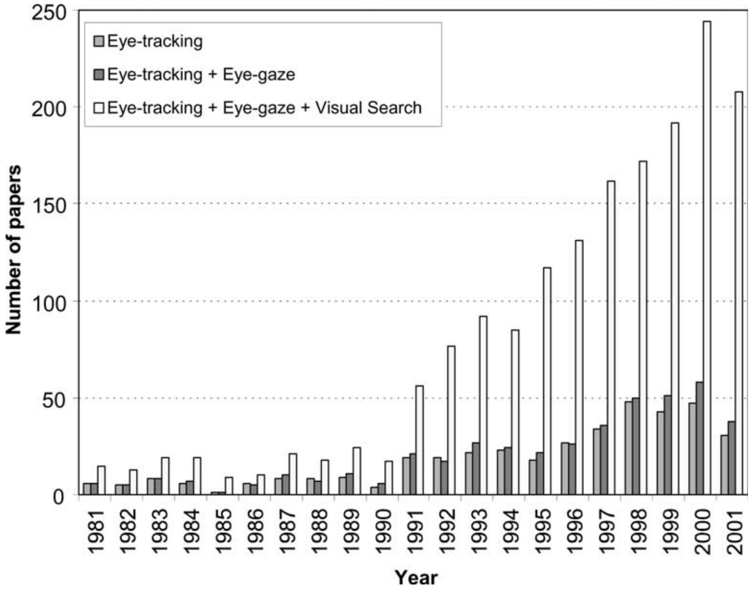

Fig. 2. The trend in research publications in visual search since 1981 based on the search result of Science Citation Database using keywords of eye tracking, eye-gaze, and visual search (the information about the research output in 2001 only covers the first three quarters).

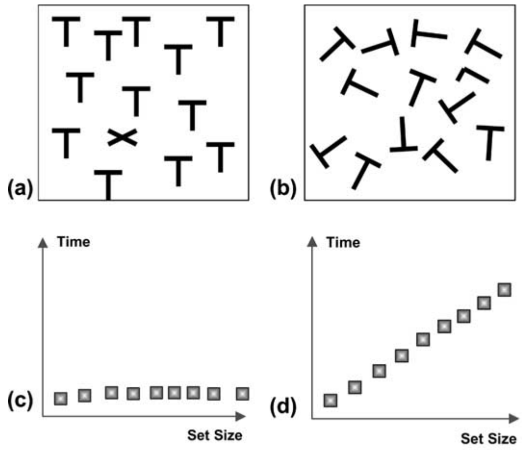

Fig. 3. (a,b) Two typical examples of test patterns for experiments using feature based models for visual search, and their corresponding RT/set-size plot (c,d). The steeper RT/set-size slope is indicative of serial searches.

Although the mechanism of visual search has been researched for many years, accurate measurements of eye movements have only become possible in the last few decades. Recent advancement in tracking hardware design has allowed the implementation of visual search experiments in both laboratory settings and public displays [1,2]. This has resulted in a widespread use of eye-tracking techniques for a whole range of new applications that include art, engineering, psychology, cognitive science, behaviour science, man–machine integration design, market research, and medical imaging.

Fig. 2 illustrates the trend in research publications in visual search, where three key words, *eye tracking*, *eye gaze*, and *visual search*, are used to search the Science Citation Database (http://wos.mimas.ac.uk/). Albeit being illustrative, the graph signifies the growing trend in the number of publications in recent years. With the future introduction of consumer products for tracking eye movements, the applications of visual search are likely to accelerate even further. The purpose of this review is to outline different psychophysical models for visual search, current eye-tracking hardware design, and potential applications of visual search.

2. Psychophysical models for visual search

Like all research into human perception, the construction of psychophysical models for visual search is a challenging task. Hitherto, visual search models commonly involve the use of visual features to measure the response time of observers in identifying them among different distractors. The visual features to be examined can include colour, orientation, motion, size, curvature, depth, vernier offset, gloss, intersection, and spatial position/phase. The results of these experiments have led to perceptual and cognitive models for describing low-level mechanisms of visual search [3–13]. In such experiments, the number of objects in the search space is called the *set size* and a typical analysis of these experiments is to plot the reaction time (RT) against the set size. Fig. 3 demonstrates an example that has been used for over 30 years to investigate human visual responses. It is designed to evaluate the RT that one takes to single out a symbol amongst other similar shaped distractors. The experiment shown in Fig. 3 will, in general, produce two types of result. The first, as shown in Fig. 3(c), which corresponds to Fig. 3(a), is the very flat plot of time taken against set size. It appears that the RT is independent of the set size. The other, demonstrated in Fig. 3(d), which corresponds to Fig. 3(b), is a steeper plot of time taken against the set size, which increases in proportion to the number of objects within the image. This implies that the RT is strongly influenced by the set size.

The above experiments provide an objective way of assessing general visual search responses. To elucidate the underlying mechanics of how visual search takes place, however, one needs to consider in detail the directed attention that our cognitive resources rely on. The attentive selection may be modelled in different ways; it is common, however, to distinguish between *early selection* (or *passive*) models (i.e. selection occurs before recognition) and *late selection* (or *active*) models (i.e. selection occurs after recognition) [14]. Early work on focused attention by Broadbent [15] and Treisman [16] resulted in the Early Filtering Model and the Attenuation Model, which are regarded as a cornerstone in cognitive science.

In the 1980s, Triesman and Gelade [17] introduced the Feature Integration Theory (FIT), which was a visual search model that consisted of two different stages, i.e. parallel search and serial search. In parallel visual search, the object is considered to be significantly different to the distractor and all objects can be processed in parallel in order to identify the target. The parallel description implies that all of the objects are processed concurrently. This means that the time taken in a parallel search is largely independent of the set size. Because of this, it produces a flat *RT/set size* slope that is indicative of parallel searches. The serial search occurs when that target object is not significantly different from the distractors. Each object in the search space has to be attended to separately and a decision made as to whether it is the target or not before moving onto the next object. A steeper RT/set size slope is indicative of serial searches. Although the distinction between serial and parallel searches had a long history by then, the FIT provided for the first time a systematic framework for combining the two.

The serial/parallel dichotomy entails the definition of *basic features* that can support parallel visual search. Although initial approaches considered the slopes in RT experiments, it became apparent that such definition failed in conjunction searches. In FIT, the human vision system processes in parallel each individual basic feature and encodes them into feature maps. Such maps register the activity in response to given features, whereas information

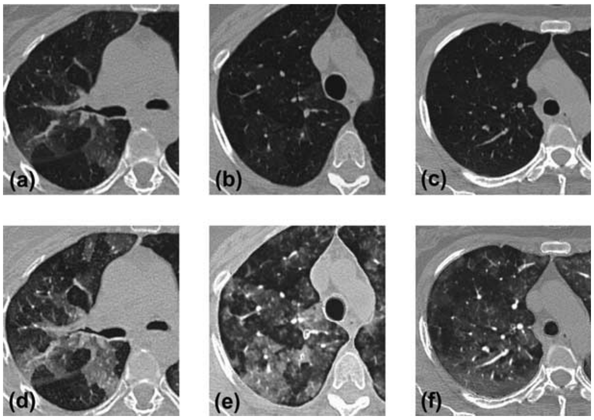

Fig. 4. (a–c) A set of CT images of patients with varying degrees of diffuse lung disease ranging from severe to mild, and (d–f) their corresponding image enhanced results to highlight affected lung parenchyma shown as dark patches. The enhanced results are not normally available in clinical environment and radiologists have to rely on their experience to differentiate the subtle intensity differences in the lung parenchyma.

about the location, spatial distribution within a given map, or relationships between each other is provided by a master map of location. A serial scan of visual attention on the master map is required when the information about the relationship between features is necessary (e.g. in conjunction search). According to this model, if there is a single basic feature involved in the search, it is possible to rely on the feature map alone to perform detection. The FIT presents, however, major drawbacks and has been criticised for its strict parallel/serial dichotomy. This is because a large number of RT tests have produced RT/set time plots that occur in a continuity of different gradients [12,13,18].

Without completely abandoning the FIT, Wolfe et al. later proposed a Guided Search Model, which aimed to explain the continuum, rather than the strict serial/parallel dichotomy, that was observed in the data [19,20]. It envisages that there exists an early parallel stage, which closely collaborates with later serial mechanisms [18]. The model suggests that an activation map is initially pre-attentively created to direct attention to the subsequent locations of interest in the visual field either through *bottom-up* (stimulus driven) or *top-down* (user-driven) processes. The bottom-up processing is highly effective when the target is considerably distinct from the distractors [18], and therefore, in such situations it leads to high activation values. On the other hand, the top-down processing predisposes the attentive resources to stimuli that attract attention. In the Guided Search Model, a ranking of stimuli is calculated by combining the information derived from the bottom-up and top-down processes. Thus, attention is thought to be directed by the visual system to highly prioritised regions.

The analysis discussed earlier is fundamental to the understanding of basic search mechanisms, its application to domain specific image understanding tasks, however, requires further considerations. Fig. 4 shows a set of computed tomographic (CT) images of three patients with different degrees of diffuse lung diseases. The essential function of the lung is the exchange of oxygen and carbon dioxide between the blood and the atmosphere by a process of molecular diffusion across an alveolar membrane. The manifestation of diffuse lung diseases on CT images, described in a simple term, is dark patches of the lung parenchyma during expiration due to air trapping. The simple method of RT might be suitable for Fig. 4(a), where the extent of the disease is severe. The manner in which an observer assimilates information for understanding Fig. 4(b) and (c) is, however, much more complicated. Initial glance can no longer provide sufficient information for reaching a decisive conclusion; close scrutiny relying on foveal vision for different areas of the image must follow. To understand this style of search, the scan path of foveal vision must be used to unlock the inherent characteristics of how information is assimilated.

It has now been widely accepted that fixations are governed by intention, and therefore areas containing more information are preferred. In practice, however, it is also common that some areas are completely ignored during a visual search. Kahneman [21] classified eye movements into three general types depending on the situation when they occur:

- *Spontaneous looking* occurs when the subject observes any scene without any specific task in mind.
- *Task-relevant looking* is guided by a particular question or task during the observation.
- *Orientation of thought looking* arises as a consequence of the observer not paying attention to the actual visual field under consideration but rather the attention is concentrated in an inner thought.

In a typical scan path experiment the subject is presented with a search space. As the subject searches over this space, his eyes are tracked and recorded. This data is then analysed to determine the fixations and scan-paths in the image. The search task can be looking for a pre-defined target, which can be similar to the search spaces used in RT experiments [5,9,22,23]. The alternative is to present a subject with a scene and observe how the eyes move over the image [24,25]. The main benefit of this style of testing is that a far more detailed understanding about the actual process of visual search can be obtained. Furthermore, it permits experiments with visual search spaces that are larger than the foveal field of view. McPeek et al. [7] suggests that attention and saccadic eye movement are linked. The RT test, which does not monitor eye movements, will not have information about how attention changes as the search progressed. The studies of McPeek et al. [7] and Yuille and Coughlan [26] all indicate that there is more information relating to the search available when the eye movement is considered.

To address the practical issues of image understanding based on visual search, extensive experiments have been

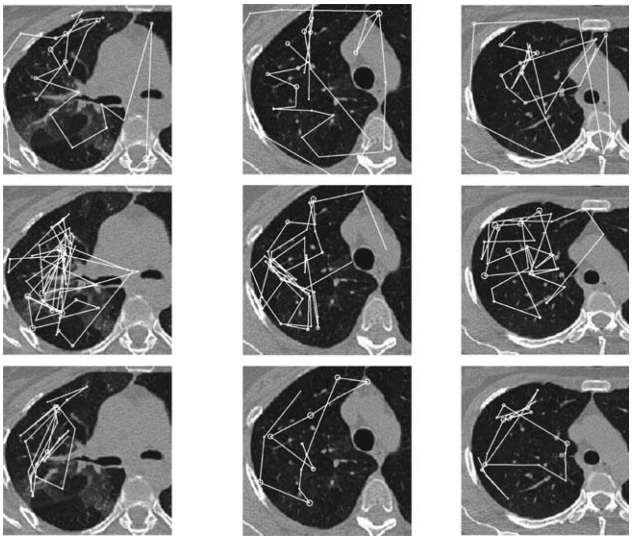

Fig. 5. The scan paths of two different observers, when reading the set of CT images shown in Fig. 4. Top row: result for a consultant radiologist who specialises in lung imaging. Middle and bottom rows: results for a novice observer before and after training.

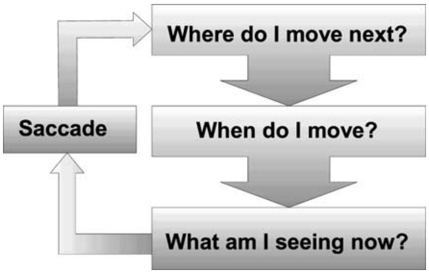

Fig. 6. A functional model describing the process of making a saccade.

conducted by Noton and Stark based on the scan path theory [27]. Their approach was based on tracking the eyes of a subject when observing an image for the first time. This was then followed by a secondary test to determine how the eyes moved when the subject was presented with an image that they had seen before. The initial findings of the scan path theory were that there was often a regular path that the eyes follow across the image. This would often be repeated, and the path was generally cyclic, returning to the start before repeating the same path again. This led to the proposal of a feature ring as an internal representation of how the eye observes, stores and recognises an image. The feature ring is based on remembering parts of the image that contain the most amount of information and their displacement to the next feature. Recognition occurs when this path can be repeated in a new image.

The concept of feature ring indicates the influence of prototypic impression on subsequent visual search tasks. It is, however, largely due to the work of Nodine and Kundel, that a more systematic framework has been proposed for modelling skilled search behaviours in image understanding. The *global-focal model* advocated by Nodine and Kundel [28] suggests the following main stages of the process involved in image understanding: *global impression*, *discovery search*, *reflective search*, and *post-search recall*. The global impression is concerned with the initial coarse screening by identifying gross departures from the observer's normal prototypic impression. In this stage of the process, the peripheral vision plays a dominant role. The discovery search is triggered by the input from the global impression and involves a systematic scrutiny of target features, whereas the reflective search is the process of building up evidence through cross-referencing image features for reaching a decision about a complicated or ambiguous target. If the search time is limited and a decision is not reached before the display is finished, post-search recall based on the memory imprint of preceding search results has to be applied. Although the global-focal model is presented in a radiographical image understanding context, it is expected to be applicable to most situations where domain specific knowledge plays a key role. One important aspect of the model is its emphasis on the cooperative perceptual-cognitive strategy behind skilled search. It has been found that, when image features are not conspicuous, the perceptive aspect of the survey shifts emphasis towards the cognitive side, which very much depends on training and experience. Fig. 5 illustrates an example search path for the images shown in Fig. 4, and those of a novice before and after training.

Existing research has also found that there are neural links in the brain to co-ordinate the 'what' and 'where' behaviours of the search [7,29,30]. These links form two pathways in the brain. The 'where' pathway goes to the parietal cortex; this is the area of the brain that is used to determine spatial locations and relationships. The 'what' pathway leads to the inferior temporal cortex. This is believed to be that part of the brain that recognises the objects. Combining these pathways with the results of a pre-attentive search and in the case of the 'what' pathway, a foveal windowing leads naturally to a search system that can recognise objects in the foveal field of view and uses the peripheral vision to determine where to move to next, as indicated in Fig. 6. The next salient feature in the visual field is the location that the saccade moves to. This is achieved by processing the result of the pre-attentive search.

Since the eyes perform a problem-oriented selection of information under the control of visual attention, the behavioural and cognitive aspects of such a task should be carefully considered. It is important to observe that the prior knowledge about the scene being observed largely affects the way that the image is scanned. In an earlier paper, Noton and Stark [27] state that "Recordings of the points inspected in the scanning of a picture and of the path the eye follow in the inspection provide clues to the process whereby the brain perceives and recognises objects".

The prior knowledge described earlier can be analysed from another point of view. Suggestions or hints can be introduced to change the manner with which an observer

Fig. 7. Examples of eye trackers manufactured by Applied Science Laboratories (ASL): (a) model ASL-501 (accuracy 0.5–1.0°, visual range 50° (H), and 40° (V), sampling rate 50–60 Hz with optional increase to 120–240 Hz) and (b) ASL-504 with remote optics (accuracy 0.5°, visual range 50° (H), and 40° (V), sampling rate 50–60 Hz with optional increase to 240 Hz). Image courtesy of Applied Science Laboratories.

reads an image. Experiments have shown that despite the difference in scan paths of different observers, the search patterns are rarely random. Suggestions or hints can have a profound effect on the way we observe a scene. Yarbus [31] in his famous experiment in the 1960s recorded the eye movements of a subject while he was observing a painting by Repin named ‘An Unexpected Visitor’. During the experiment, the pattern of eye movements changed substantially as the viewer was given a number of prompts. Such prompts turned out to guide the visual search undertaken. It is important to note that other aspects of visual search related to learning [3,32,33] and behavioural models [29,30] have also been established. Despite major advances in our understanding of visual search, the quest for elucidating its intrinsic mechanism and constructing practical vision systems that are inspired by their findings continue.

3. Eye-tracking hardware

Hitherto, most experiments involving eye tracking are conducted in Clinical Psychology/Psychiatry, with schizophrenia being one of the most frequently considered topics. On-line processing of eye-tracking data was initially developed by military research but it had been expanded to other domains. A substantial increase in the research interests in this area is mainly due to the availability of cheaper, more accurate, and easier to use eye-tracking devices in the last few years. Although the techniques aimed at tracking the gaze of a person have existed for many years, many experiments so far are still conducted in laboratories and involve obtrusive actions.

The objective of eye-tracking systems is to determine when and where fixations occur. The time order of the list of fixations represents the actual visual search that takes place. There are various methods of tracking the eye position relative to the head. Assuming the head position is known, then, this can be used to follow the gaze of a subject. Fig. 7 gives an example of two different types of eye-tracking systems; the one shown on the left has to be mounted on the head of the subject whereas the other uses remote optics, and therefore offers a much more natural working environment. There are currently three major categories of techniques for measuring eye movements: electro-oculography (EOG), scleral contact lens/search coils, and techniques based on reflected light such as video-oculography (VOG). A survey of eye movement recording methods is provided by Young and Sheena [34].

EOG is based on the measurement of skin electric potential difference around the eye which varies as the eye rotates. Electrodes can be placed above, below and to the sides of the eyes, with the potential difference across the eye indicating the eye position. This method can achieve an accuracy of ±1.5°. Since it effectively measures the position of the eye relative to the head, head tracking must be used if the gaze position is to be determined.

Scleral contact lens and search coils rely on either mechanical or optical reference object mounted on contact lens worn directly on the eyes. The basic principle of the technique dates back to over a century ago, although at that time the reference object was placed directly onto the

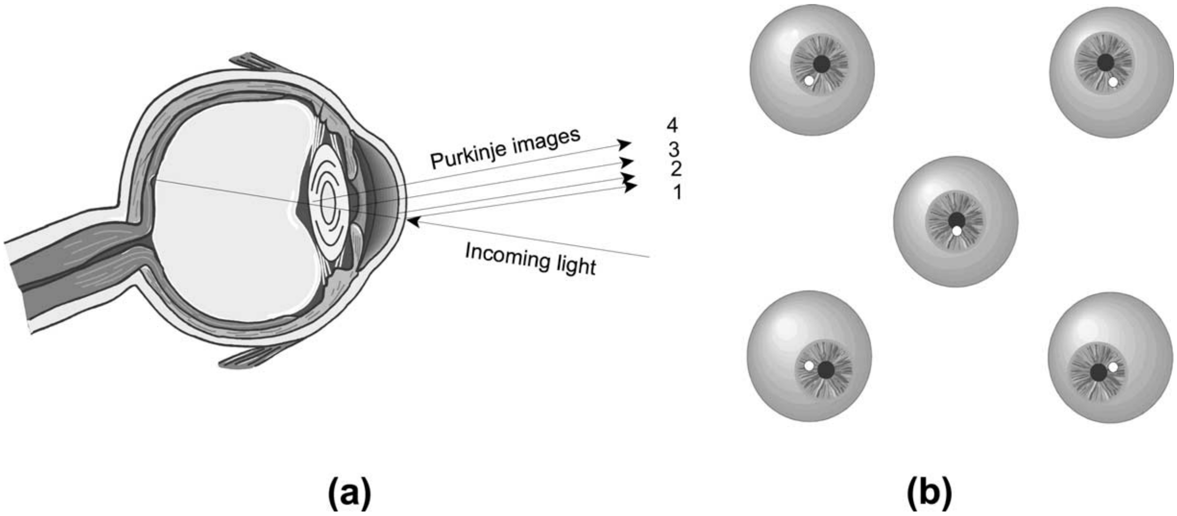

Fig. 8. (a) The basic principle of how the four Purkinje images are formed. (b) The relative positions of the pupil and the first Purkinje reflection during visual search as seen from the camera when the observer fixates to the centre and the four corners of the screen.

cornea. The current implementation of the system commonly involves the measurement of the response of wire coils mounted on the contact lens moving through an electromagnetic field. Despite the invasive nature of the method, it is one of the most accurate measurement techniques for determining relative position of the eye and the head.

The third category, which is based on reflected light, is perhaps the most commonly used eye-tracking technique. It involves the measurement of distinguishable features of the eyes during visual search. These features include limbus, pupil, and corneal reflection. Anatomically, the front view of the eye can be divided into three parts: the sclera, the iris and the pupil. The interface of the sclera and the iris is called the limbus. The limbus moves proportionally to the rotation of the eye and its position can be tracked optically with the use of photodiodes. This technique, however, is only effective for horizontal tracking. For vertical tracking the eyelid blocks the limbus and thus prevents its true position from being obtained. The technique can nevertheless be extended to tracking the pupil position as the pupil/iris interface is better defined and the eyelid does not obscure the pupil. Both of these techniques require the head to be fixed in position or the use of head mounted optical sensors.

In visual search, it is important to determine the point in space that is being looked at (the point-of-regard) rather than the position of the eye relative to the head. To accomplish this, it is necessary to have a combined analysis of the above features. Most current techniques rely on the measurement of both pupil and corneal reflection. The geometrical structure of the eye is formed in such a way that the corneal reflection of a fixed light source, known as Purkinje reflection, is relatively stable while the eyeball, and hence the pupil, rotates in its orbit. Since the relative position between the centre of the pupil and the corneal reflection remains constant with minor head movements, it is therefore possible to use the corneal reflection as the reference point to determine the point-of-regard without physically fixing the position of the head. To avoid the problems of eyes being out-of-focus or even falling out of the viewing angle of the camera due to head movement, systems based on the pupil/corneal reflection method usually require automatic eye tracking combined with auto-focusing features. The range with which such systems can operate is normally between ± 12–15° as further eye movements will cause the corneal reflection to fall outside the spherical part of the cornea, thus requiring more complicated geometrical correction steps. Furthermore, if a larger amount of head movement is to be accommodated for, an electro-magnetic tracking device will need to be incorporated into the system.

Given a fixed light source, typically infra-red, there are in fact four Purkinje reflections associated with it. Among these four reflections, two of them are caused by the reflection of the front and rear surfaces of the cornea, and the other two from the reflection of the lens. By measuring the position of the first and fourth Purkinje reflections, it is possible to separate translational and rotational eye movements. Although this technique offers more accurate measurement results, the fourth-Purkinje image is normally very weak due to a change of the refraction index at the rear of the lens. Therefore, the surrounding lighting must be strictly controlled in such experiments. Fig. 8 describes how the four Purkinje reflections are formed as well as the relative positions of the pupil and the first Purkinje reflection when the eye fixates to the centre and the four corners of a computer screen.

There are currently many different types of eye movement tracking devices available commercially. A comprehensive list can be found at the Eye Movement Equipment Database [35]. Table 1 summarises the main features of these systems in terms of subject constraints, accuracy, sampling rate, range and resolution.

Table 1
A comparison of the invasiveness, resolution, range and sampling frequency of different approaches of eye tracking

	Electro-oculography	Limbus tracking	Search coils	Video-oculography	Dual Purkinje
Invasiveness	No, but needs electrodes	No, but needs head stabilisation	Yes	No	No
Resolution	1°	0.1°	0.017°	0.25°	0.017°
Range	± 70°	± 20°	± 25°	± 40°	± 20°
Sampling frequency	40 Hz	0.2–4 kHz	4 kHz	50–240 Hz	> 1 kHz

4. Applications of eye tracking

With the advent of accurate, compact, and affordable eye-tracking systems, the practical applications of visual search have flourished in recent years. Eye-tracking systems may be thought as falling into the following two categories [36]: (a) real-time interactive systems, and (b) off-line diagnostic systems. The first category accounts for those applications where the user's gaze has to be known in order to enable the interaction (e.g. interactive environments guided by the user's gaze), whereas the second category relies on the recording of data for subsequent analysis (e.g. learning from scan paths). The previous classification may also divide systems into *active mode* systems in the case of real-time interactive applications and *passive mode* systems for off-line diagnostic applications. In this framework, eye-tracking devices operating in passive mode are aimed at monitoring the eye activities. The data are then registered and stored for further analysis. Contrarily, those devices operating in active mode make use of the movement of the eyes to undertake any specific actions. It is important to note that the applications of eye tracking are diversified and in this section we will only illustrate some of the key examples in psychology, psychiatry, cognitive science, behaviour analysis, medicine, and human–computer interaction.

4.1. Psychology/psychiatry and cognitive science

So far, psychology, psychiatry, and cognitive science are the most active research areas that involve eye tracking. The main objectives of these experiments are to obtain a detailed understanding about the cognitive process of visual search [13,37] and examine the developmental differences of oculomotor functioning at different stages of our lives. Particular interest is being paid to cognitive processes in infants since even those a few months old, can use eye gaze to note what is of interest to them, either by directional looking or pupil enlargement. In parallel to this, experiments have also been undertaken to analyse visual perception in animals, mainly monkeys, so as to improve our understanding of visual recognition under different stimuli [38,39], as well as gaining further insight into the neural code that relate to visual perception [40] and neural mechanisms causing strabismus.

It has been found that eye-tracking techniques are particularly useful in diagnosing reading disabilities. Research has shown that there exist characteristic patterns of eye movements which are symptomatic of particular reading disabilities [41–43]. Among them, people who suffer from dyslexia show an abnormal pattern of the saccadic eye movements when they are reading [44,45]. Particularly, when the number and duration of fixations during reading are compared between dyslexics and normal readers, it indicates that the number of forward saccades is higher and their size smaller for dyslexics. Furthermore, their fixation durations are longer than those from normal readers [46]. Eye-tracking techniques can also be used to investigate psycholinguistic issues in cognitive processes (e.g. contextual representation in interpretation, models of idiom processing, spoken word recognition and ambiguity resolution) [47–52].

Eye movement abnormalities have been actively used to investigate psychiatric disorders. This line of research was triggered by Philip Holzman in the mid-1970s after noticing the pursuit eye movement disturbances not only in schizophrenic patients, but also in their unaffected relatives. In recent years, much work has been done to investigate the neurobiological basis of affective disorders, obsessive compulsive disorder, dementias, attention deficit disorder and autism [53]. For patients suffering from mental diseases, eye-tracking experiments have been widely used in assessing schizophrenia [54–59], Alzheimer [60,61], and bipolarity [62].

4.2. Behaviour analysis

Since the early days of visual tracking research, it has been recognised that the knowledge of visual search patterns of experienced observers can potentially be useful for improving the learning process of trainees by adopting a learning-by-imitation procedure. Such an approach has been found to be particularly effective in medical imaging where the interpretation of image details is largely subjective. Eye tracking provides an effective means of gaining information about how experts assimilate different visual clues to reach a confident diagnosis [28,63–68]. It also offers an objective way of assessing the competence of trainees in fulfilling the recognition tasks. Furthermore, gaze-contingent processing also enables the exploitation of implicit knowledge for teaching, joint problem solving, or in a global framework, the communication of practical knowledge and skills. Due to the wide spread use of on-line tutorials, eye-tracking techniques can also play an important role in assessing how subjects interact with the tutorial in distance learning systems.

Capturing visual attention has been found to be particularly useful for reducing errors in critical tasks such as those undertaken by control operators and clinicians [69,70]. Stresses related to fatigue and noise are major factors involved in making poor judgements. Eye tracking is becoming an integral part of driving-simulators as it provides vital information about the reasoning, prediction, and level of alert of the driver [71]. For investigating the safety implications of using mobile phones while driving, as there exists a shift of attentional resources [72], measures derived from eye tracking, such as the number of fixations and duration of fixations, have been used together with RT measures to assess the effect of a cellular phone conversation on the performance of driving tasks such as searching for traffic signals.

Another important area that involves visual tracking is usability studies. The term usability refers to the extent to which the users can operate a product to achieve specific goals [73]. For example, eye-tracking techniques enable the characterisation of the user's behaviour in using a particular design of computer interface. By analysing parameters such as the time spent looking at each region, the dwell time on certain icons, and the latency in finding a specific function, it is possible to derive the general friendliness and intuitiveness of the interface [74]. The value of this approach for optimising the designs of human–computer interface has received extensive interests and the SIGCHI '99 conference contained a workshop dedicated to this issue [75].

Based on the same principles, it is also possible to use eye-tracking systems to determine what attracts people's attention. For advertising, it is particularly useful for understanding where people devote their visual interest. The material involved can either be static images or video sequences [76–82]. A study on consumers' eye movement patterns on yellow pages advertising was recently undertaken by Lohse [83], which revealed several interesting behaviours of a general population. Among them it has been found that consumers noticed over 93% of the quarter page display advertisements whereas almost 75% of the plain listings were completely ignored [83]. Some other interesting results arise from on-line advertising and the realisation that Internet users do not tend to fixate web advertisements but rather seem to avoid them. Furthermore, eye-tracking techniques can be useful for assessing advertisement strategies. In particular, they have been considered to assess how people react to 'explicit' advertisements (i.e. depicts the target product together with a related headline in a semantically straightforward way) and 'implicit' advertisements (i.e. the images and text depicted are not directly related to the product) [84].

4.3. Medicine

Since the discovery of X-ray by Wilhelm Roentgen more than a century ago, medical imaging has advanced to a stage that was inconceivable 100 years ago, with X-rays, fluoroscopy, ultrasound, CT, magnetic resonance imaging (MRI), and positron emission tomography (PET) scans becoming an integral part of modern healthcare. The technology has now advanced to a stage that the time taken to acquire the image has reduced so much that the longest part of radiology is often the analysis and reporting procedures [85]. The key problems in medical image understanding include errors due to missing or not recognising important diagnostic features during visual search, and the variation of interpretation between different observers. The study reviews the sources of the errors that occur in the interpretation of images in radiology as well as the factors influencing the magnitude of observer variation. Inadequate search time, incomplete coverage of the image by the visual field, or breaking off a search before all possible locations have been explored are common causes of errors.

It has been over 20 years since the initial use of eye tracking for analysing problems associated with medical image understanding. An early study in scan-path analysis of radiographs by Kundel et al. [65] provided information on the nature of errors in detecting pulmonary nodules in chest X-ray images. It was found that 30% of false-negative errors were due to incomplete scanning, 25% were due to failure of recognising visual features, and 45% due to the wrong decisions. Under the current climate of clinical governance in medicine, the potential implication of using visual tracking for quality assurance and training is extensive, as it provides an objective way of addressing some of the problems mentioned earlier. For example, the issue of incomplete scanning can be avoided by using eye tracking to highlight areas that are missed during visual examination. For training, the scan paths of consultant radiologists can be replayed to trainees to understand the intrinsic relationships between different feature points that attract fixations.

In addition to the global-focal model, Nodine and Kundel also proposed a visual dwell time algorithm [86], which has shown to be of value for avoiding the failure of recognising abnormal visual features. This approach used the scan path co-ordinates and the associated dwell time to identify hot spots in the image. These hot spots signify areas towards which attention has been drawn during the visual search and are subsequently prompted to the interpreter at the end of examination to reconfirm their visual significance. This dwell time analysis was subsequently improved by Nodine et al. [87], who provided a more in-depth analysis of eye position recordings based on fixation clustering. They have also investigated the effect of training on the visual search performance [88]. The analysis of time-to-hit data (i.e. time needed for the observers to fixate on to the target region) indicated that the importance of search was not to sample the image exhaustively, nor to find a given target, but rather to find something peculiar or quaint about the image. It is through experience that trained observers become more aware of the presence of such peculiarities in the images.

Such awareness brings visual attention to this quaint region for interpretation and, subsequently, relies on a knowledge base to recognise atypical variants for diagnostic purposes.

The issues arising from these studies indicate the significance of gaining more experience in interpreting visual input once the eyes fixate onto a target candidate. Covering a specific region of interest does not by itself guarantee detection. It is possible to detect features of interest on images with peripheral vision. However, the likelihood of detection decreases with the eccentricity of the feature of interest with regard to the centre of the foveal attention. Overall, it can be concluded that skilled visual search comes under more cognitive control with training. Furthermore, visual attention becomes more perceptually attuned to the detection and decision-making requirements of the task.

Beyond radiology, eye tracking has also been used to investigate pharmacological effects on patients, assist surgical operations, and map human brain activity with functional MRI (fMRI). Pokorny et al. [89] presented an early study on the influence of drugs on eye-tracking tasks. Fant et al. [90] later found that both acute and residual subjective, physiologic, and performance effects of consuming any drug can be assessed by eye-tracking tests. In laser refractive surgery, patient eye and head movements are major sources of error due to the change in relative position of the surgical beam and the target [91]. Active eye-tracking systems have been incorporated into the laser delivery systems, which are used to cancel out the effect of patient eye movement [92–94], and therefore helps to minimise post-operative glare. In fMRI, eye-movement monitoring has been integrated into the system to examine the relationship between visual perception and regional brain activities. The use of eye-tracking techniques is also useful in the study of diseases such as vertigo symptoms [95], multiple-sclerosis [96], and insulin-dependent diabetes-mellitus [97], through visuomotor responses of the patients.

4.4. Human–computer interaction

It is expected that the major impact of accurate, compact, and easy to use eye-tracking systems is likely to be in the design of human–computer interaction. Military research sponsored some of the early work in eye-tracking systems. Such research was aimed at enabling pilots to guide weapons by selecting targets with their eyes [98–100]. Existing research has found that eye-tracking is invaluable in designing environments such as aircraft cockpits, air traffic control stations, and robot tele-operation stations where the operator has to remain in a confined position to control a device, often with a high level of stress. The optimum design of these environments is considered to be a 'black art' and requires the designers to be familiar with a '40 different specialties, ranging from anthropometry to optics to research psychology' [101]. This has motivated the development of the man–machine integration design and analysis system (MIDAS). MIDAS is a simulator that replicates the user in the system by using a number of agents. Each of these is an autonomous system that acts in a specific behaviour of the user. One of these agents is a behavioural model of visual search for perception and attention. So far, MIDAS has been used for several different purposes such as the design of the next generation nuclear power plant consoles, or optimising commercial transport.

In conventional applications, one of the most common uses of eye-tracking techniques is for assisting disabled people. To this end, the gaze may be used to select a given action from a computer screen, enabling them to interact and control their environment. In these systems, careful considerations must be made so as to avoid undesired activations [102–104]. A usual way to type text using the gaze as input is to display a keyboard on the computer screen and use dwell time as a selection technique. It is unlikely that text input with eye gaze is to become a substitute to keyboard typing for general use except for certain special areas such as mobility enhancement for disabled people. By incorporating eye tracking, Hyrskyhari et al. [105] describe a translation aid, iDict, to improve the reading of foreign language documents with the assistance of gaze path information. The system contains built-in information that includes general knowledge of the reading process, lexical and syntactical language analysis of the text that helps in the identification of potentially unknown words or syntactically complex structures, and user's profile. The operation of iDict is based on tracking the reading process of an on-screen document and detecting where the user seems to have troubles. The system infers what help is needed and delivers it to the user. This application is targeted to a wide audience whereas most current applications are designed for people with disabilities.

In the last few years, eye-tracking techniques are attracting increasing interests from the computer graphics community due to the rich sets of visual stimuli, ranging from 2D imagery to 3D immersive virtual scenarios that they generate. Visual attention and perception can influence the way scenes are constructed. One of the possibilities opened to the community is the use of eye tracking for gaze-contingent applications for which the resolution and quality of an image or animation is adjusted according to the viewer's saccades and fixations. Another line of research not much studied yet is the analysis of the perception of synthesised images and animations with the objectives of optimising the realism of such perception and improving the efficiency of the graphics algorithms [106].

Gaze contingent displays rely on the observation that most of the resources that are used to produce large and high-resolution displays are actually wasted, particularly when dealing with single-user displays. This is due to the physiology of the human visual system that can only resolve detailed information within a very small area at the fovea. Therefore, by knowing where the gaze of the observer is directed at a given instant, it would be convenient to display high-resolution information in that region. This requires a dynamic updating of the image in real-time based on eye-gaze direction data. The production of a display completely indistinguishable of a full-resolution display, however, is a very challenging task since the update rate of the display needs to be extremely high and the degradation in the peripheral fields has to be kept minimal. The effect of varying the spatio-temporal parameters that affect the performance of eye-contingent multi-resolutional displays has been discussed by Loschky et al. and Parkhurst et al. [107,108].

In tele-communication and tele-work tasks, it is important to know what is the state of attention of the participants. In this context, visual attention indicated by gaze direction plays an important role [109–113]. An example of such an application is the GAZE Groupware system. This is an eye-controlled audio-visual conferencing which, based upon eye-tracking techniques, allows users to see who is talking to whom about what in a 3D meeting room on the Internet [114]. Another interesting application of eye tracking is the so-called interest and emotion sensitive media [115,116], which enables audience-directed script navigation by considering the users' interests and emotional states. The potential of this technology has motivated an intensive research interest [117]. Both IBM and MIT have proposed prototype systems that allow the viewer to interact with a multi-threaded documentary using a multi-modal interface [118,119].

In the field of affective computing, Partala et al. [120] have investigated pupillary responses to emotionally provocative sound stimuli. The results suggest that the pupil size discriminates among different kinds of emotional stimuli. Furthermore, eye-tracking techniques may be used to improve a human interface. Chino et al. [121] simulated a '*meta*-communication' by requiring from the user intending to input voice commands to a given system, to previously gaze the agent in order to request to talk. This allows the system to accept only intended voice input. It is interesting that eye tracking is also finding its way to consumer products. Canon has devices (35 mm camera and video camera) which use this technology to focus in the direction of the user's gaze rather than in the direction in which the camera is pointed. Until recently, the greatest number of focusing points available on a 35 mm camera was five. The Canon EOS-3 35 mm camera enables photographers to enjoy 45 focusing points. All focusing points are linked directly to Canon's eye-controlled autofocus system [122].

5. Future trends and discussions

Looking ahead, both research and practical applications of visual search are likely to experience a steady increase in the next few years. Its use in psychology, psychophysics, and cognitive science will continue to provide researchers with further insight into and fundamental understanding of how visual understanding takes place. With the increasing sophistication and versatility of functional imaging techniques, eye tracking is valuable in neuroscience for investigating active functional cortical regions of the brain and identifying cognitive loads during various visual search tasks.

In ergonomics and human–computer interaction, eye-tracking techniques are vital in the instrument design and the development of next generation virtual environments. It is foreseeable that current research in binocular eye-tracking system for determining the viewer's gaze in 3D space will become an integral part of future immersive media systems. This will be attractive in virtual reality based training environment, particularly in medicine, in that over the last 10 years there has been a strong movement away from open surgery towards improved techniques of minimal access surgery. Endoscopy, including bronchoscopy and laparoscopy is the most common procedure in minimal access surgery, which is carried out through natural body openings or small artificial incisions. If handled properly, endoscopes are completely harmless to patients. Minimal access surgery and diagnostic endoscopy can achieve their clinical goals with minimal inconvenience to patients. Compared with conventional techniques, patient trauma and hospitalisation can be greatly reduced, and diagnostic accuracy and therapeutic success increased. However, the complexity of the instrument controls, restricted vision and mobility, difficult hand–eye co-ordination, and the lack of tactile perception require a high degree of manual dexterity of the operator. Computer simulation provides an attractive possibility for certain aspects of this training, particularly the hand–eye co-ordination aspects. Eye-tracking techniques, in this case, can be used as a means of instrumental control and navigation. In system design, it will also offer the possibility of using gaze-contingent strategies with foveo-peripheral rendering so that the level of details of a particular region of the scene is controlled by the direction of the gaze.

For surgical robot, the eye-gaze contingent paradigm may be used to facilitate the transmission of images and enhance the performance of the procedure, enabling a much more accurate navigation of the instruments. Current applications of robotics in surgical assistance include dexterity enhancement, systems networking, image-guided therapy, as well as remote operations using real time tele-observation and monitoring. The recent FDA approval of the daVinci surgical system, which is a completely robotic surgery device, allows the surgeons to perform procedures while seated at a computer console and 3D video imaging system across the room from the patient. The surgeons operate controls with their hands and feet to direct a robotically controlled laparoscope. The use of eye-tracking technology can be extremely useful in providing navigation at high-resolution through the anatomy as required from the surgeon's gaze. This enables the positioning of the micro-cameras to be done in a more natural fashion as they would cover in real-time the visual demands of the surgeon rather than requiring explicit mechanical manoeuvres.

Among the host of applications mentioned earlier, the application of visual search in human–computer interaction can become the centre of focus for the graphics and vision research community. Both research and development are underway for providing gaze enabled interactive environment and collaborative systems for multi-party communication. One application that might experience an increasing attention in the future is the *interest and emotion sensitive media*. Although it faces the challenging problem of inferring the *interests* of a subject, the consideration of eye-tracking data within a more general framework based on the integration of other devices (e.g. cameras to analyse face expression among others) might open a new world of entertainment and other fields beyond it.

It is always difficult to foresee what the future holds for a new technique. For the image and vision computing community, visual search is set to offer a range of opportunities for future research and development. This is mainly due to the following two motivating factors. First, there is still a range of system level problems that need to be addressed. The existing eye-tracking systems still require a cumbersome calibration process, and issues related to accuracy, reliability, and user friendliness can potentially become problems when eye tracking is used for general mass-market applications. The potential use of eye-tracking systems for web-based applications can become the driving force of developing a low-cost, versatile, and automatically calibrated eye-tracking system that allows the user to interact with the digital environment in a natural, unrestricted manner. The use of a multiple camera configuration for real-time tracking of both head and eye movement as well as gesture information can be a viable option, and this is the area that the vision community has a lot to offer.

The second, and perhaps, a more important area for the image and vision computing community is that the information on visual search can provide an important link between cognitive science, artificial intelligence (AI), and computational vision. One example of area of development is decision support in image understanding. Decision support for image understanding was one of the first fields to which AI techniques were applied in the 1950s. Explicit domain knowledge, in this case represented as low-level visual features, is integrated into the system along with ad hoc feature extractors and decision rules. Although there has been substantial development in the handling of uncertainties in such systems, this approach has inherent drawbacks. It is important to note that human visual pattern recognition is different from general reasoning. Explicit domain knowledge representation often overlooks those factors that are subconsciously applied during visual recognition.

In radiology, for instance, the manner in which an expert radiologist assimilates information from a single image (for example, a chest radiograph) has been studied at a basic level, and results are available on visual tracking patterns and the dwell-times of certain features of the image. Experienced radiologists do not appear to scrutinise an image thoroughly, but rather linger on key areas, leaving large areas of the image unexamined. It has been shown that experts examine fewer fixation points compared to trainees or unskilled observers. After an area of interest or abnormality is fixed upon, there follows a rapid checking of the validity of the observation to exclude technical causes or artefacts and, more importantly, the search for ancillary and corroborative signs to support the initial observation. This process is perhaps the most difficult part of diagnostic information gathering, because many of the actions and filtering that occur are rapid and happen at an unconscious level [65]. Furthermore, the ad hoc nature of grouping of low-level visual features means that there is no coherent approach in the development of decision support systems. It can be envisaged that information on visual search can provide a general framework for knowledge gathering in decision support in image understanding. In medicine, this can be used not only for diagnosis, but also for training in visual recognition. The traditional training of radiologists is by a process of 'directed search' initially for obvious localised abnormalities that can be detected by examining all areas of the image systematically. Providing information on which image features are most thoroughly studied by experts, and what cross-checking is carried out, should improve the speed and efficacy of training. This idea is supported by the observation that decision support systems that provide a second opinion and draw the attention of radiologists to areas of potential abnormalities, have been shown to improve diagnostic accuracy and reproducibility.

Visual search is a common task that people perform throughout their daily life. The number of applications that it can be applied to is potentially large. The examples described earlier by no means attempt to be a complete survey of the host of applications being carried out in this field. Nonetheless, it is hoped that they provide a flavour of what is the current state-of-the-art of this captivating area. The use of visual search models for objective assessment of image interpretation skills in engineering and medicine, designing improved human–computer interfacing for interacting with virtual environments, incorporating visual search modelling systems into the behaviour of foraging robots, and defining new frameworks for computed assisted image understanding is only the tip of an iceberg which is gradually gathering its momentum moving towards more widespread practical applications.

References

[1] A. Sturgis, Time to look, Telling Time, National Gallery, London, 2000.

[2] C. Buquet, J.R. Charlier, V. Paris, Museum application of an eye tracker, Medical and Biomedical Engineering and Computing 26 (1988) 277–281.

[3] A. Ellison, V. Walsh, Perceptual learning in visual search: some evidence of specificities, Vision Research 38 (3) (1998) 333–345.

[4] T.S. Horowitz, J.M. Wolfe, Visual search has no memory, Nature 394 (6693) (1998) 575–577.

[5] I.T.H.C. Hooge, C.J. Erkelens, Adjustment of fixation duration in visual search, Vision Research 38 (9) (1998) 1295–1302.

[6] U. Leonards, R. Rettenbach, R. Sireteanu, Parallel visual search is not always effortless, Cognitive Brain Research 7 (1998) 207–213.

[7] R.M. McPeek, V. Maljkovic, K. Nakayama, Saccades require focal attention and are facilitated by a short term memory system, Vision Research 39 (1999) 1555–1566.

[8] C.M. Moore, H. Egeth, How does feature-based attention affect visual processing, Journal of Experimental Psychology: Human Perception and Performance 24 (4) (1998) 1296–1310.

[9] B.C. Motter, E.J. Belky, The guidance of eye movements during active visual search, Vision Research 38 (1998) 1805–1815.

[10] H.C. Nothdurft, Focal attention in visual search, Vision Research 39 (1999) 2305–2310.

[11] J. Theeuwes, A.F. Kramer, P. Atchley, Attentional effects on preattentive vision: spatial precues affect the detection of simple features, Journal of Experimental Psychology: Human Perception and Performance 25 (2) (1999) 341–347.

[12] J.M. Wolfe, What can 1 million trials tell us about visual search? Psychological Science 9 (1) (1998) 33–39.

[13] J.M. Wolfe, G. Gancarz, Guided search 3.0, in: V. Lakshminaragaman (Ed.), Basic and Clinical Applications of Vision Science, Kluwer Academic Publishers, Dordrecht, 1996, pp. 189–192.

[14] J.A. Deutsch, D. Deutsch, Attention: some theoretical considerations, Psychological Review 70 (1963) 80–90.

[15] D.E. Broadbent, Perception and Communication, Pergamon Press, New York, 1958.

[16] A. Treisman, Contextual cues in selective listening, Quarterly Journal of Experimental Psychology 12 (1960) 242–248.

[17] A. Treisman, G. Gelade, A feature integration theory of attention, Cognitive Psychology 12 (1980) 97–136.

[18] J.M. Wolfe, in: H. Pashler (Ed.), Attention, University College London Press, London, UK, 1996.

[19] M.M. Chun, J.M. Wolfe, Visual attention, in: E.B. Goldstein (Ed.), Blackwell Handbook of Perception, 2000.

[20] J.M. Wolfe, K.R. Cave, S.L. Franzel, Guided search: an alternative to the feature integration model for visual search, Journal of Experimental Psychology: Human Perception and Performance 15 (1989) 419–433.

[21] D. Kahneman, Attention and Effort, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[22] G.J. Zelinsky, R.N.P. Rao, M.M. Hayhoe, D.H. Ballard, Eye movements reveal the spatiotemporal dynamics of visual search, Psychological Science 8 (6) (1997) 448–453.

[23] G.J. Zelinsky, D.L. Sheinberg, Eye movements during parallel–serial visual search, Journal of Experimental Psychology: Human Perception and Performance 23 (1) (1997) 244–262.

[24] J.R. Barrett, N.G. Johns, E.S. deParedes, T.E. Hutchinson, S.J. Dwyer III, Visual search patterns during diagnosis of repeatedly displayed mammograms, Engineering in medicine and biology society, Engineering advances: new opportunities for biomedical engineers, Proceedings of the 16th Annual International Conference of the IEEE, vol. 1, 1994, pp. 588–589.

[25] T.T. Blackmon, Y.F. Ho, K. Matsunaga, T. Yanagida, L.W. Stark, Eye movements while viewing dynamic and static stimuli, Engineering in medicine and biology society, Proceedings of the 19th Annual International Conference of the IEEE, vol. 6, 1997, pp. 2814–2819.

[26] A.L. Yuille, J. Coughlan, High-level and generic models for visual search: when does high level knowledge help? Computer vision and pattern recognition, IEEE Computer Society Conference, vol. 2, 1999, pp. 631–637.

[27] D. Noton, L. Stark, Eye movements and visual perception, Scientific American 224 (6) (1971) 35–43.

[28] C.F. Nodine, H.L. Kundel, The cognitive side of visual search in radiology, in: J.K. O'Regan, A. Levy-Schoen (Eds.), Eye Movements: From Physiology to Cognition, Elsevier, Amsterdam, 1986.

[29] G.J. Gieflin, H. Janssen, H. Mallot, Saccadic object recognition with an active vision system, Conference A: Computer Vision and Applications, The 11th IAPR International Conference on Pattern Recognition, vol. I, 1992, pp. 664–667.

[30] I.A. Rybak, V.I. Gusakova, A.V. Golovan, L.N. Podladchikova, N.A. Shevtsova, A model of attention-guided visual perception and recognition, Vision Research 38 (1998) 2387–2400.

[31] A.L. Yarbus, Eye movements and vision, Plenum Press, New York, 1967.

[32] M. Ahissar, S. Hochstein, Learning pop-out detection: specificities to stimulus characteristics, Vision Research 36 (1996) 3487–3500.

[33] R. Sireteanu, R. Rettenback, Perceptual learning in visual search: fast, enduring but non specific, Vision Research 35 (1995) 2037–2043.

[34] L.R. Young, D. Sheena, Methods and designs: survey of eye movement recording methods, Behaviour Research Methods and Instrumentation 7 (5) (1975) 397–429.

[35] Eye Movement Equipment Database, http://ibs.derby.ac.uk/emed/.

[36] Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA, November 6–8 2000, http:/www.vr.clemson.edu/eyetracking/et-conf/.

[37] M. Kutas, K.D. Federmeier, Minding the body, Psychophysiology 35 (2) (1998) 135–150.

[38] G. Schweigart, T. Mergner, G. Barnes, Eye movements during combined pursuit, optokinetic and vestibular stimulation in macaque monkey, Experimental Brain Research 127 (1) (1999) 54–66.

[39] S.G. Lisberger, Postsaccadic enhancement of initiation of smooth pursuit eye movements in monkeys, Journal of Neurophysiology 79 (4) (1998) 1918–1930.

[40] S. Martinez-Conde, S.L. Macknik, D.H. Hubel, Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys, Nature Neuroscience 3 (3) (2000) 251–258.

[41] K. Rayner, G.W. McConkie, What guides a reader's eye movements? Vision Research 16 (1976) 829–837.

[42] M.A. Just, P.A. Carpenter, Eye fixations and cognitive processes, Cognitive Psychology 8 (1976) 441–480.

[43] M.A. Just, P.A. Carpenter, A theory of reading: from eye fixations to comprehension, Psychological Review 87 (4) (1980) 329–354.

[44] http://www.brain.uni-freiburg.de/fischer/welcome.html.

[45] B. Fischer, K. Hartnegg, A. Mokler, Dynamic visual perception of dyslexic children, Perception 29 (5) (2000) 523–530.

[46] M. De Luca, E. Di Pace, A. Judica, D. Spinelli, P. Zoccolotti, Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia, Neuropsychologia 37 (12) (1999) 1407–1420.

[47] D.A. Titone, C.M. Connine, On the compositional and noncompositional nature of idiomatic expressions, Journal of Pragmatics 31 (12) (1999) 1655–1674.

[48] M.K. Tanenhaus, M.J. Spivey-Knowlton, Eye-tracking, Language and Cognitive Processes 11 (6) (1996) 583–588.

[49] M.J. Pickering, M.J. Traxler, M.W. Crocker, Ambiguity resolution in sentence processing: evidence against frequency-based accounts, Journal of Memory and Language 43 (3) (2000) 447–475.

[50] D. Dahan, D. Swingley, M.K. Tanenhaus, J.S. Magnuson, Linguistic gender and spoken-word recognition in French, Journal of Memory and Language 42 (4) (2000) 465–480.

[51] S. Garrod, M. Terras, The contribution of lexical and situational knowledge to resolving discourse roles: bonding and resolution, Journal of Memory and Language 42 (4) (2000) 526–544.

[52] S. Matsunaga, Universality in reading processes: evidence from an eye-tracking study, Psychologia 42 (4) (1999) 290–306.

[53] J. Sweeney, Eye movements and psychiatric disorders, Conference Program and Abstract from the 11th European Conference on Eye Movements, Turku, Finland, 2001, p. K2 (Abstract).

[54] H.R. Konrad, L.P. Rybak, D.E. Ramsey, D.R. Anderson, Eye tracking abnormalities in patients with cerebrovascular-disease, Laryngoscope 93 (9) (1983) 1171–1176.

[55] D.L. Levy, P.S. Holzman, S. Matthysse, N.R. Mendell, Eye tracking

and schizophrenia—a selective review, Schizophrenia Bulletin 20 (1) (1994) 47–62.

[56] D.L. Levy, C.M. Lajonchere, B. Dorogusker, D. Min, S. Lee, A. Tartaglini, J.A. Lieberman, N.R. Mendell, Quantitative characterization of eye tracking dysfunction in schizophrenia, Schizophrenia Research 42 (3) (2000) 171–185.

[57] C. Hooker, S. Park, Trajectory estimation in schizophrenia, Schizophrenia Research 45 (1–2) (2000) 83–92.

[58] R. Nicolson, M. Lenane, S. Singaracharlu, D. Malaspina, J.N. Giedd, S.D. Hamburger, P. Gochman, J. Bedwell, G.K. Thaker, T. Fernandez, M. Wudarsky, D.W. Hommer, J.L. Rapoport, Premorbid speech and language impairments in childhood-onset schizophrenia: association with risk factors, American Journal of Psychiatry 157 (5) (2000) 794–800.

[59] P.S. Holzman, Eye movements and the search for the essence of schizophrenia, Brain Research Reviews 31 (2–3) (2000) 350–356.

[60] G. Muller, S. Weisbrod, F. Klingberg, Test battery for the objectification and differential-diagnosis in the early-stage of presumed presenile-dementia of Alzheimer-type, Zeitschrift fur Gerontologie 26 (2) (1993) 70–80.

[61] G. Muller, R.A. Richter, S. Weisbrod, F. Klingberg, Impaired eye tracking performance in patients with presenile onset dementia, International Journal of Psychophysiology 11 (2) (1991) 167–177.

[62] P.S. Holzman, C. Obrian, C. Waternaux, Effects of lithium treatment on eye-movements, Biological Psychiatry 29 (10) (1991) 1001–1015.

[63] H.C. Cowley, A.G. Gale, A.R.M. Wilson, Mammographic training sets for improving breast cancer detection. In: H.L. Kundel (Ed.), Proceedings-SPIE The International Society for Optical Engineering, vol. 2712, 1996, pp. 102–113.

[64] H.C. Cowley, A.G. Gale, Non-radiologists and CAD systems in breast cancer screening, Computational Imaging and Vision (13) (1998) 371–374.

[65] H.L. Kundel, C.F. Nodine, D.P. Carmody, Visual scanning, pattern recognition and decision-making in pulmonary nodule detection, Investigative Radiology 13 (1978) 175–181.

[66] R. Niimi, K. Shimamoto, A. Sawaki, T. Ishigaki, Y. Takahashi, N. Sugiyama, E. Nishihara, Eye-tracking device comparisons of three methods of magnetic resonance image series displays, Journal of Digital Imaging 10 (4) (1997) 147–151.

[67] K.P. White, T.L. Hutson, T.E. Hutchinson, Modeling human eye behavior during mammographic scanning: preliminary results, IEEE Transactions on Systems Man and Cybernetics, Part A—Systems and Humans 27 (4) (1997) 494–505.

[68] E.A. Krupinski, Visual scanning patterns of radiologists searching mammograms, Academic Radiology 3 (1996) 137–144.

[69] S. Saito, Does fatigue exist in a quantitative measurement of eye-movements? Ergonomics 35 (5–6) (1992) 607–615.

[70] G.E. Larson, Z.A. Perry, Visual capture and human error, Applied Cognitive Psychology 13 (3) (1999) 227–236.

[71] D. Tock, I. Craw, Tracking and measuring drivers' eyes, Image and Vision Computing 14 (8) (1996) 541–547.

[72] C.T. Scialfa, L. McPhee, G. Ho, The effects of a simulated cellular phone conversation on search for traffic signs in an elderly sample, Proceedings of Eye Tracking Research and Applications Symposium, ACM, 2000, pp. 45–51.

[73] International Organization for Standardization (adapted from ISO 9241-11), http://www.iso.ch/.

[74] C.R. Benel, D. Ottens, R. Horst, Use of the eyetracking system in the usability laboratory, 35th Annual Meeting of the Human Factors Society, Proceedings of the Human Factors Society, New Orleans, LA, 1991, pp. 461–465.

[75] Advance Program SIGCHI '99. Table of Contents—Consortia, Tutorials and Workshops, http://www.acm.org/sigchi/chi99/cp/?ventoc = Workshop.

[76] J.H. Goldberg, C.K. Probart, R.E. Zak, Visual search of food nutrition labels, Human Factors 41 (3) (1999) 425–437.

[77] R. Pieters, E. Rosbergen, M. Wedel, Visual attention to repeated print advertising: a test of scanpath theory, Journal of Marketing Research 36 (4) (1999) 424–438.

[78] R. Pieters, L. Warlop, Visual attention during brand choice: the impact of time pressure and task motivation, International Journal of Research in Marketing 16 (1) (1999) 1–16.

[79] R.J. Fox, D.M. Krugman, J.E. Fletcher, P.M. Fischer, Adolescents' attention to beer and cigarette print ads and associated product warnings, Journal of Advertising 27 (3) (1998) 57–68.

[80] R. Pieters, L. Warlop, M. Hartog, The effect of time pressure and task motivation on visual attention to brands, Advances in Consumer Research 24 (1997) 281–287.

[81] R. Pieters, E. Rosbergen, M. Hartog, Visual attention to advertising: the impact of motivation and repetition, Advances in Consumer Research 23 (1996) 242–248.

[82] D.M. Krugman, R.J. Fox, J.E. Fletcher, P.M. Fischer, T.H. Rojas, Do adolescents attend to warnings in cigarette advertising—an eye-tracking approach, Journal of Advertising Research 34 (6) (1994) 39–52.

[83] G.L. Lohse, Consumer eye movement patterns on yellow pages advertising, Journal of Advertising 26 (1) (1997) 61–73.

[84] R. Radach, C. Vorstius. K. Radach, Eye movements in the processing of advertisements: effects of pragmatic complexity, Conference Program and Abstract from the 11th European Conference on Eye Movements, Turku, Finland, 2001, p. S49 (Abstract).

[85] P.J.A. Robinson, Radiology's Achilles' heel: error and variation in the interpretation of the Röntgen image, The British Journal of Radiology 70 (1997) 1085–1098.

[86] C.F. Nodine, H.L. Kundel, A visual dwell algorithm can aid search and recognition of missed lung nodules in chest radiographs, in: D. Brogdan (Ed.), First International Conference on Visual Search, Taylor and Francis, London, 1990, pp. 399–406.

[87] C.F. Nodine, H.L. Kundel, L.C. Toto, E.A. Krupinski, Recording and analyzing eye-position data using a microcomputer workstation, Behavior Research Methods, Instruments and Computers 24 (3) (1992) 475–485.

[88] C.F. Nodine, H.L. Kundel, S.C. Lauver, L.G. Toto, Nature of expertise in searching mammograms for breast masses, Academic Radiology 3 (1996) 1000–1006.

[89] R. Pokorny, G. Ferber, M. Matejcek, P. Irvin, H. Klee, The influence of drugs on eye tracking tasks, Electroencephalography and Clinical Neurophysiology 64 (1) (1986) 12.

[90] R.V. Fant, S.J. Heishman, E.B. Bunker, W.B. Pickworth, Acute and residual effects of marijuana in humans, Pharmacology, Biochemistry and Behavior 60 (4) (1998) 777–784.

[91] N.M. Taylor, R.H. Eikelboom, P.P. van Sarloos, P.G. Reid, Determining the accuracy of an eye tracking system for laser refractive surgery, Journal of Refractive Surgery 16 (5) (2000) S643–S646.

[92] I.G. Pallikaris, K.I. Koufala, D.S. Siganos, T.G. Papadaki, V.J. Katsanevaki, V. Tourtsan, M.B. McDonald, Photorefractive keratectomy with a small spot laser and tracker, Journal of Refractive Surgery 15 (2) (1999) 137–144.

[93] R.R. Krueger, In perspective: eye tracking and autonomous laser radar, Journal of Refractive Surgery 15 (2) (1999) 145–149.

[94] Y.Y. Tsai, J.M. Lin, Ablation centration after active eye-tracker-assisted photorefractive keratectomy and laser in situ keratomileusis, Journal of Cataract and Refractive Surgery 26 (1) (2000) 28–34.

[95] M. Ura, C.R. Pfaltz, J.H.J. Allum, The effect of age on the visuoocular and vestibuloocular reflexes of elderly patients with vertigo, Acta Oto-Laryngologica 481 (Suppl.) (1991) 399–402.

[96] A.B.D. Howard, D.C. Good, H.R. Konrad, Eye tracking tests and evoked-responses in early multiple-sclerosis, Electroencephalography and Clinical Neurophysiology 58 (2) (1984) 40.

[97] J. Virtaniemi, M. Laakso, J. Nuutinen, S. Karjalainen, E. Vartiainen, Voluntary eye-movement tests in patients with insulin-dependent diabetes-mellitus, Acta Oto-Laryngologica 113 (2) (1993) 123–127.

[98] C. Adams, If looks could kill: the eyes have it, Military and Aerospace Electronics, March (1990) 35–37.

[99] C.C. Smyth, B.B. Bates, M.C. Lopez, N.R. Ware, A comparison of eye-gaze to touch panel and head-fixed reticle for helicopter display control and target acquisition during a simulated armed reconnaissance mission, Proceedings of the Second Mid-Atlantic Human Factors Conference, Washington, DC, Fairfax, VA, George Mason University, 1994, p. 49 (Abstract).

[100] Bae Systems: The Eurofighter, http://www.baesystems.com/.

[101] B.R. Smith, S.W. Tyler, The design and application of Midas: a constructive simulation for human-system analysis, 1997, http://www.maad.com/SAE/publicat.htm.

[102] T.E. Hutchinson, K.P. White, W.N. Martin, K.C. Reichert, L.A. Frey, Human–computer interaction using eye-gaze input, IEEE Transactions on Systems, Man, and Cybernetics 19 (6) (1989) 1527–1534.

[103] L.A. Frey, K.P. White Jr., T.E. Hutchinson, Eye-gaze word processing, IEEE Transactions on Systems, Man, and Cybernetics 20 (4) (1990) 944–950.

[104] C. Colombo, A. DelBimbo, Interacting through eyes, Robotics and Autonomous Systems 19 (3–4) (1997) 359–368.

[105] A. Hyrskyhari, P. Majaranta, A. Aaltonen, K. Räihä, Design issues of iDict: a gaze-assisted translation aid, Proceedings of the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA, November 6–8 2000, pp. 9–14.

[106] C. O'Sullivan, J. Dingliana, G. Bradshaw, A. McNamara, Eye-tracking for interactive computer graphics, Conference Program and Abstract from the 11th European Conference on Eye Movements, Turku, Finland, 2001, p. S45 (Abstract).

[107] L.C. Loschky, G.W. McConkie, User performance with gaze contingent multiresolutional displays, Proceedings of the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA, November 6–8 2000, pp. 97–103.

[108] D. Parkhurst, E. Culurciello, E. Niebur, Evaluating variable resolution displays with visual search: task performance and eye movements, Proceedings of the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA, November 6–8 2000, pp. 105–109.

[109] H. Fuchs, Beyond the desktop metaphor: toward more effective display, interaction, and telecollaboration in the office of the future via a multitude of sensors and displays, Advanced Multimedia Content Processing 1554 (1999) 30–43.

[110] B. Velichkovsky, Communicating attention: gaze position transfer in cooperative problem solving, Pragmatics and Cognition 3 (2) (1995) 199–222.

[111] B. Velichkovsky, J.P. Hansen, New technological windows into mind: there is more in eyes and brains for human–computer interaction, Proceedings of ACM CHI96 Conference, 1996, pp. 496–503.

[112] R. Vertegaal, R. Slagter, G.C. Van der Veer, A. Nijholt, Why conversational agents should catch the eye, Summary of ACM CHI 2000 Conference on Human Factors in Computing Systems, The Hague, The Netherlands, 2000, http://www.cs.queensu.ca/~roel/publications/2000/chi2000AgentsandGaze.pdf.

[113] R. Vertegaal, G.C. Van der Veer, H. Vons, Effects of gaze on multiparty mediated communication, Proceedings of Graphics Interface, Canadian Human–Computer Communications Society, Montreal, Canada, 2000, pp. 95–102.

[114] The GAZE Groupware at ACM '97 Expo, The Next 50 Years of Computing, San Jose, CA, March 1997, http://www.cs.queensu.ca/~roel/gaze/home.htm.

[115] A.J. Glenstrup, T. Engell-Nielsen, Eye Controlled Media: Present and Future State, University of Copenhagen, Bachelor's Degree Thesis, DIKU (Institute of Computer Science) Denmark, 1995.

[116] J.P. Hansen, A.W. Andersen, P. Roed, Eye-gaze control of multimedia systems, Proceedings of the Sixth International Conference on Human Computer Interaction, Tokyo, Japan, vol. 1, Elsevier, Amsterdam, 1995, pp. 37–42.

[117] I. Starker, R.A. Bolt, A gaze-responsive self-disclosing display, Proceedings of SIGCHI 90, ACM Press, New York, 1990 pp. 3–9.

[118] BlueEyes: creating computers that know how you feel, http://www.almaden.ibm.com/cs/blueeyes/.

[119] CyberBELT, http://www.media.mit.edu/people/davet/proj/cyberbelt.html.

[120] T. Partala, M. Jokiniemi, V. Surakka, Pupillary responses to emotionally provocative stimuli, Proceedings of the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA, November 6–8 2000, pp. 123–129.

[121] T. Chino, K. Fukui, K. Suzuki, 'GazeToTalk': a nonverbal interface with meta-communication facility, Proceedings of the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA, November 6–8 2000, p. 111.

[122] Cannon's eye-controlled autofocus system, http://www.usa.canon.com/.