

Attention and Action

James J. Clark¹, Ziad M. Hafed² and Li Jie¹

¹Centre for Intelligent Machines, McGill University

²Sloan-Swartz Center for Theoretical Neurobiology, The Salk Institute for Biological Studies

1 Introduction

This article provides a historical overview of the work carried out by the first author and his co-workers in the area of visual attention and its connection to active vision. It is written, in parts, in the first person, from the perspective of the first author. This is not intended, however, to in any way minimize the contributions of the second and third authors, which are significant and extensive, particularly in the latter sections.

Attention has been studied in great detail by researchers since the beginnings of the scientific approach to cognitive psychology and the founding works of Hermann Helmholtz (Helmholtz 1890) and William James (James 1890). To this day, however, there is no universally agreed-upon theory of how attention functions in humans, and there is even still controversy as to what attention actually is. Attention remains, therefore, a very active area of research. Most would agree that attention is a *selective* process, which acts to focus sparse computational resources onto relevant aspects of the sensory input. My particular interest originated in the application of attention processes to machine vision. In particular, I was interested in so-called *active* vision systems, which deal with the problems and opportunities that arise when various aspects of the sensing process, such as the position and orientation of cameras, can be actively controlled. My interest in machine vision systems has since been extended to biological vision systems. This paper recounts the line of research that has grown out of this interest.

2 Active Vision and Attention in Robots

The line of research outlined in this article began when I was a neophyte professor working in the Harvard Robotics Lab, under the direction of Roger Brockett. Along with my Ph.D. student, Nicola Ferrier (now a professor at the University of Wisconsin), I was interested in building an artificial active vision system, capable of moving a pair of video cameras under computer servo control so as to orient their gaze to a desired location in space. The resulting piece of hardware, one of the first such to be built, was the so-called Harvard Head (Ferrier and Clark, 1993). Nowadays the Harvard head seems primitive and ungraceful, but it did produce at least one interesting piece of research, namely the demonstration of how selective attention could be linked to the generation of

saccadic camera movements. We implemented the Koch and Ullman (1985) saliency-map model of attention on a combination Datacube/Sun-3 image processing platform (details can be found in (Clark and Ferrier 1992)), and used the peak of the saliency map to define the saccade target (Clark and Ferrier 1988, 1989). This saccade target was input as a command to a modal-control system developed by Brockett (Brockett 1988), which controlled the position and velocity of the camera pan, tilt, and vergence motors. The Harvard Head project was one of many in the field of active vision. The late 80's and early 90's was a productive era in the development of active vision hardware. From the Canadian perspective, one of the most important contemporary active vision hardware systems was the University of Toronto "Trish" system (Miliot *et al* (1990)). This system makes use of attentional mechanisms, and served as the development platform for much of John Tsotsos' influential work on attention modeling. An excellent overview of the state-of-the-art in the early 90's in active vision is the report produced by an NSF panel of experts (Swain and Stricker, 1991). This report also has a good presentation of what was considered to be the important research issues in the area, and these issues are still relevant today.

In the mid-90's I left Harvard University and joined the fledgling Nissan Cambridge Basic Research laboratory. This lab was located near MIT in Cambridge, Massachusetts, and was under the scientific direction of Whitman Richards, Ken Nakayama, and Warren Seering. It was focused on understanding human cognition while driving, for the purposes of making driving safer and more enjoyable. Nissan was interested in using eye movement measurements as a way of tapping into the driver's attentive state. Perhaps the most significant piece of research to come out of the CBR was the work on change-blindness due to disruption of attention (Rensink *et al* 1995, 1997, O'Regan *et al* 1999). Attentional change blindness refers to the phenomena of extreme difficulty in noticing image changes if these changes are masked by global transients in image salience. Normally, one would think of change-blindness as a bad thing, but the phenomena reflects the active nature of information acquisition in biological systems. The basic idea follows O'Regan's (1992) view of the "world as an outside memory", where detailed internal representations of the world are not needed, since the world can form the representation itself. Information need only be acquired by the visual system when changes in the world are detected. In this view, attention serves to localize the location of scene change, which is then followed by detailed analysis of the nature of the change. If nothing changes, then there is no need to re-acquire the information.

After two years working at Nissan, I joined the Centre for Intelligent Machines at McGill University in Montreal. At McGill, my lab has continued applying attentional processes to the construction of active vision systems. In Clark (1998b) I describe how an attentive active vision system can exhibit change-blindness.

3 Attention and Saccade Generation in Biological Systems

While at the Nissan CBR I learned a great deal about biological attention mechanisms and, in particular, the link between attention and eye movements. Along the way, I learned something about doing psychophysical experiments. My subsequent work at McGill has focussed more on modeling of biological vision systems than on implementing machine vision systems. The remainder of this article will therefore concentrate on my work related to modeling and studying the characteristics of biological attention mechanisms, and their link to the generation of eye movements.

3.1 Models of Saccade Targeting and Triggering

In my view, the first step in understanding the link between attention and eye movements is to have a working model of how eye movements are generated, and then to fuse this with a model of how attention is controlled. Modeling of the processes and neural circuitry underlying the generation of saccadic eye movements has occupied many researchers since the time of Helmholtz and James. Attention is to be found in these models with varying degrees of involvement. The oldest form of model considers the processes of saccadic eye movement generation and attention control to be completely independent. Examples of this type of model can be found in the works of Becker and Jurgens (1979), Reulen (1984a,1984b), and Deubel *et al* (1984). Posner (1980) accepted the notion that there is an apparent link between the activity of covert attention and the generation of saccades, but argued that this link was coincidental. He held that the two systems are only functionally related and suggested that covert attention and eye movements are both drawn to exogenous stimuli, but not to endogenous stimuli, and that the mechanism controlling these shifts are completely independent.

Some models do propose a connection between the generation of saccades and visual attention, but only require that attention be disengaged from the current locus for a saccade to occur. A representative of such a model is that of Fischer (1993). In Fischer's model the target of a saccadic eye movement is computed during the disengaged-attention phase by a localization system.

Finally, there are many models that infer a very strong coupling between attention and saccade generation. One common form of such models require that attention be engaged at a target location before a saccade can be made to that location. An early theory along this line was proposed by Wurtz and Mohler (1976). They suggested that attention shifts were *programs* for saccadic eye movements, to be subsequently executed by the oculomotor system. Klein's (1980) *oculomotor readiness theory* provides what is perhaps the most influential of the early attention/saccade models. This theory links attention and saccade generation by supposing that an attention shift to a particular spatial location results from a preparation of, or an oculomotor readiness to generate, a saccade to that location. This oculomotor readiness has the effect of enhancing information processing at the target location. Klein later *et al* 1992) qualified his theory to restrict the linkage between attention and saccade generation to situations where the saccade target is defined by an exogenous (image-based) stimulus, and not to situations involving *endogenous* (internally defined) saccade targets.

Henderson (1992) introduced a *sequential attention* model that invoked a *saccadic programming* process. In this model, the saccadic programming process is executed after attention shifts to the peripheral stimulus only after processing of the foveal visual input has been completed. The amount of foveal processing (or foveal load) affects the latency of the eye movement (the time from the appearance of a peripheral stimulus to the start of the saccade). For light foveal loads the latency is modeled as 80msec beyond the time taken for the attention shift. A modified version of his theory allowed for the saccadic program to start before the attention shift, which will reduce the latency, but still keeps the principle that saccades are ultimately directed to the locus of attention.

A rather extreme, but influential, conception of the link between attention and saccades is the *premotor theory* of Rizzolati (1983). The main tenet of this theory is that the system that controls action is the same as that which controls spatial attention. One module performs both functions,

leading to an economical architecture. One implication is that there can be many forms of attention, each associated with a different motor system (Rizzolati *et al* 1994). Each motor modality employs its own neural representation of space, referred to as a *pragmatic map*. One such pragmatic map is the Superior Colliculus motor map (Wurtz 1996). Saccade targeting activity involving the Superior Colliculus motor map would then affect the allocation of spatial attention.

The models mentioned above that invoke a strong link between attention and saccades were motivated primarily by the results of psychophysical studies. Support for the idea of a strong link between attention and saccades has also been provided by a number of neuro-physiological studies. Desimone (1990) made the general observation that the oculomotor system and the covert attention system both involve the targeting of stimuli and could usefully share some common neural hardware. Desimone *et al* (1989) observed an impairment of an animal's ability to attend to a target in the presence of a distractor when small areas of the superior colliculus were rendered non-functional. Conversely, Kustov and Robinson (1996) forced saccades to be generated in monkeys by injecting electrical current into various areas of the superior colliculus and observed that the trajectories of these induced saccades were altered by both exogenous and endogenous attentional shifts.

3.2 Saccadic Latency Phenomena

What are the implications of the premise that saccadic eye movements are linked to attention shifts? One is that manipulation of the attention process should result in measurable changes in eye movement characteristics, such as the amplitude, direction, or timing. In fact, each of these characteristics are indeed subject to measurable changes. In my work I have concentrated mainly on looking at temporal modulation effects, as these are often quite strong as compared with spatial effects.

Saccadic eye movements to suddenly appearing peripheral targets are never instantaneous. There is always some time delay, referred to as the *saccadic latency*, before the saccade begins. This is not surprising, as clearly there must be some chain of processing that takes place between the arrival of the photons on the retina and the activation of the eye muscles. If one looks, however, at the shortest neural pathways from retina to eye muscle, the path propagation time is rather short, on the order of 50 msec (c.f. Fischer 1993). Observed saccadic latencies, on the other hand, can be quite long, typically ranging from about 150 to 300 msec. So what is the additional delay due to? There are many theories as to the source of this delay, but most posit some time-consuming mechanism for preparing, or programming, the saccade in response to the onset of a visual target.

To judge the merits of the various proposals for saccade generation models, we can look at a number of well-known phenomena associated with conditions that can modulate the saccadic latency. Some of these phenomena are described in the following sections.

3.2.1 The Gap Effect

The *Gap Effect* was first reported by Saslow (1967), who observed that saccadic latencies were reduced when the temporal gap between the offset of the fixation stimulus and the onset of the target stimulus was increased. The latencies were seen to increase when there was a temporal overlap between the disappearance of the fixation mark and the appearance of the peripheral target. Reulen (1984a) performed a detail study of the gap effect, and measured latencies as a function of the asynchrony between fixation offset and target onset in seven subjects. He found that the data could be fit by a simple piecewise linear model, consisting of three sections: a constant latency for negative asynchronies (overlap), a (lower) constant latency for large positive asynchronies (gap), and a linear transition region between these two constant regions, located at small positive asynchronies. Kingstone and Klein (1993), and Walker *et al* (1995) observed that giving instructions to direct attention to a target location did not lead to any decrease in the magnitude of the gap effect although there was an overall reduction in latency. A similar result was observed experimentally by Reuter-Lorenz (Reuter-Lorenz *et al* 1991), who found that the gap effect is unaffected by the luminance of the target.

3.2.2 The Global Effect

When a saccade target is embedded in a field of distracting objects, the actual saccade often does not land right on the target, but instead tends to land on the “centre-of-gravity” of the target and distractor group Coren and Hoenig (1972). Findlay (1982) called this phenomenon the *global effect*. It was later shown by Coëffé and O’Regan (1987) that the global effect is strongest when the saccadic latency is short. In their experiments subjects made saccadic eye movements to a cued letter in a string of 10 letters presented in the periphery of the visual field. When subjects made saccades with very short latencies, the landing position of the eye was seen to overshoot the cued location for targets on the end of the string nearest to the fixation point and undershoot the cued location for targets on the end of the string furthest from the fixation point. As latencies were increased, the amount of over- or under-shoot was decreased. No target location under- or over-shoot was observed when only single letters were present, indicating that it was the presence of the other, non-cued, letters in the string that were giving rise to the under- and over-shoots.

3.2.3 Retinal Eccentricity Effects

Wyman and Steinman (1973) observed that saccadic latency increases rapidly as the target gets very close to the fovea. Kalesnykas and Hallett (1994) examined in detail saccadic latency for a wide range of retinal eccentricities and several different stimulus conditions. They found that latencies increase sharply for very small eccentricities and increase slowly at high eccentricities. The peak at small eccentricities is broader for less salient stimuli. For example, for target stimuli near detection threshold, the peak is about 4 degrees wide, while for target stimuli 1000 times foveal detection threshold the peak is only about 1.5 degrees wide. Target colour did not seem to affect the peak, ruling out effects due to wavelength dependent absorption of light by macular pigments. They also found that the presence of the central latency peak did not depend on head or eye position, and the peak appeared even when latency was plotted against saccadic amplitude rather than eccentricity.

3.3 Saccadic Programming or Attention Shifting?

So, how can these various saccadic latency phenomena be accounted for? Perhaps the most frequently proposed explanation is in terms of a *saccadic programming* process (see, for example, Abrams (1992), Abrams and Jonides (1988), Findlay (1992), He and Kowler (1988), and Sereno (1992)). The saccadic programming process is usually decomposed into two components - amplitude programming and direction programming. These components are usually thought of as computational processes, each running in their own architecturally distinct modules, that are initiated, run for a while, and then provide a result. In the saccadic programming view saccadic latencies reflect the time taken by these processes to produce the required amplitude and direction parameters for the saccade.

The saccadic programming view is quite natural for an engineer, and a modular approach to camera movement control is frequently seen in active machine vision systems. But I would argue that the saccadic programming approach is unnecessarily complicated. Instead, I suggest that the Premotor approach, in which attention and eye movement control mechanisms are shared, is a better way to go. My view of the saccade generation process makes three basic assumptions: (1) The target of a saccade is given by the current locus of attention, which itself is determined by a saliency map. (2) A command to make a saccade is generated every time attention shifts. (3) The execution of saccade commands can be suppressed.

In this premotor view, saccades are programmed, if they can be said to be programmed at all, by the attention mechanisms. Conversely, one could just as rightly say that attention shifts are programmed by the saccade generation mechanisms. In the premotor view it is meaningless to separate these as the mechanisms are the same. No saccadic “programming” need take place as the amplitude and direction of saccades are implicit in the attentionally defined target locations, and are always available. There is no programming process that needs to be initiated or that needs to be reset, modified, or restarted in response to a change in target position (Clark 1999).

In Clark (1998a) I showed that all of the saccadic latency effects described above can be explained by a premotor theory. The basic idea is that the main variable component of saccadic latency is the time needed for spatial attention to shift to the location of the target for the saccade. To motivate my conclusions I developed a simple dynamical model of the attention shift process and its link to saccade generation, and the various latency phenomena were replicated in computer simulations of the simple attentional model.

There are a number of computational models that describe the dynamical mechanisms underlying attention shifts (e.g. Koch and Ullman 1985, Tsotsos 1990). These differ greatly in their details, but generally the type of behavior known as “Winner-Take-All” (or WTA). A winner-take-all system is one in which elements compete against each other using mutual inhibition. The positive feedback inherent in such a system results in a stable state wherein one of the elements (the “winner”) is maximally enhanced and all the other elements are maximally inhibited. I used a uncomplicated version of these WTA models using a simple saliency map, which weights various visual features (as in Koch and Ullman 1985). Following models of the human visual cortex, I posited two types of feature detectors in the model, transient and sustained. The transient feature detectors are fast responding but have relatively low spatial resolution. The sustained feature detectors are slower to respond but have higher spatial resolution. A number of researchers (Breitmeyer and Ganz (1976),

Lennie (1980), Yantis and Jonides (1984)) have suggested that the transient effects observed in tasks requiring visual attention (e.g. those reported by Nakayama and Mackeben (1989), and by Posner, Cohen, and Rafal (1982)) may be due to the transient responses of low level feature detectors. The dynamics of the feature detectors arise from the temporal properties of their constituent neurons.

As for the generation of the saccadic eye movements, following the proposal of Lee, Rohrer, and Sparks (1988), my model assumes that the command for the saccadic eye movement is coded in a distributed fashion by a population of neurons, whose activity is attentionally modulated, such as in the superior colliculus (Wurtz 1996). Thus the target will be specified by the centre of mass of this pattern of activity, and is continually available. It should be noted that in taking this approach, the saccade target is *always* defined. There is no distinct *saccadic programming* module which computes the saccade target in response to some trigger stimulus. The target is always defined, and the saccade target is that which is defined at the moment of triggering.

3.4 Explanation of the Saccadic Latency Phenomena

In my view, the explanation of the phenomena described above lies in the dynamics of the WTA process. One of the principal features of a WTA process is that it multi-stable. The output of a WTA system is stable until the input is perturbed sufficiently, at which time it can switch rapidly to another stable stage. Based on this aspect of WTA function, the saccadic latency, in our model, depends on the time taken for the WTA attention process to *switch* from one location to another when the visual input changes (for example, because of the appearance of a salient target).

3.4.1 Gap Effect

In short, the explanation of the Gap Effect is that the time taken for a WTA system to switch from one location to another depends on the relative salience of the two locations. In the overlap condition the feature detector activity at the target location and that at the fixation location compete against each other in the winner-take-all competition. If the fixation point and the target have similar salience, this competition may take a long time to resolve, resulting in a long latency. In the gap condition, the target location is unopposed in this competition and thus wins it quickly, with a speed dependent on the target salience.

3.4.2 Global Effect

The explanation of the Global Effect is linked to two factors in my model. The first is the assumption of both transient and sustained features. For short latencies the feature detector response is dominated by the transient component, which has a low spatial frequency cutoff, effectively blurring the target and distractors together. At longer latencies the sustained component dominates, which has a higher spatial frequency cutoff, and hence creates less blurring of the target and distractors. The second factor contributing to the dependence of feature centroid on latency is the action of the winner-take-all network. Increasing the latency allows more time for the ultimate winning feature location to suppress its neighbouring distractors, thus reducing the effect of the spatial blurring.

3.4.3 Eccentricity Effects

The attentional model also provides an explanation for the increase in saccadic latency for targets with small retinal eccentricities. Once again, the explanation lies partly in the assumption that saccadic latencies reflect the time taken for the WTA network to shift from one stable state to another. This time is a function of the difference between the target saliency and the value of the local maximum function times some weight less than one. The other important piece of the puzzle comes from an assumption that the influence of a salient location in the WTA is not limited to its immediate neighborhood, nor is it constant spatially, but falls off rather slowly with distance. This implies a *local* implementation of the WTA, which is more plausible to be implemented under the constraints of biological architectures.

If we assume a local WTA, when the target is far from the fixation feature, the target will have little competition in the WTA process and latencies should be relatively short. If the target is near to the fixation, and if the salience at the fixation location is greater than the target salience, then the local maximum value may be larger than the target value, hence the winner-take-all transition time will be longer than when the target is far from fixation. As the target salience increases the distance at which the local maximum value becomes equal to the target salience becomes smaller. Thus the eccentricity at which the saccadic latency begins to increase should decrease as the target saliency increases. Conversely, it is seen in the simulations of the model that the drop-off in latency with eccentricity is slower for low saliency targets than for highly salient targets, in accordance with the results of the experiments of Kalesnykas and Hallett (1994).

4 Covert Attention Tracking and Microsaccades

So far, we have seen that there is an apparent link between the processes of selective visual attention and the generation of saccades. Various characteristics of saccadic eye movements can be influenced by attentional manipulations. This implies that observations and measurements of the saccade characteristics can provide information on the attention process, beyond what is provided by the position of the eye alone. This being said, the eye position remains the most-widely used direction measure of attentional allocation. But what about the situations in which the eye is held fixed, with saccades volitionally suppressed? In such a case there can be a significant disconnect between the position of the eye and the locus of spatial attention. This led me to consider more closely what information could be acquired regarding the allocation of attention during ocular fixation.

In 1995, while working at the Nissan CBR, I had a lunch-time discussion with co-workers Ron Rensink and Kevin O'Regan, where Ron pondered whether it would be possible to track the locus of covert attention, in much the same way that an eye-tracker can track the locus of overt attention. Ron Rensink thought that it would be difficult to do so, if only because covert attention has many non-spatial aspects (such as featural selection), and may have multiple spatial loci. Kevin O'Regan, playing the role of contrarian as usual, claimed (in jest?) that he did not believe in covert attention at all and stated that all observed effects attributed to attention could be explained by changes in acuity induced by small eye movements, or microsaccades. While I did not really give this idea much credence, I turned it backwards in my mind and thought that maybe microsaccades could be explained by attention processes. In particular, I had the thought that microsaccades could be

residual motions resulting from imperfectly suppressed normal saccades linked to attention shifts. In this view the only difference between covert and overt saccades is the presence of a suppression in the former case, perhaps mediated by the fixation cells in the superior colliculus (Munoz and Wurtz 1993). If microsaccades are simply suppressed targeting saccades then they should be associated with a covert attention shift just as are regular targeting saccades. If this is true, then microsaccades could provide a potentially useful probe into the state of covert attention.

When I came to McGill University I presented my thinking to Ziad Hafed and suggested it as the basis of a PhD thesis topic. Ziad had previously done a Master thesis developing a neural network model of the superior colliculus (Hafed and Clark 2000), and he thought that the idea of microsaccades as suppressed regular saccades was at least conceivable given the current understanding of colliculus function. So Ziad and I looked into the microsaccade literature, such as it is, and planned out an investigation of the hypothesis. The study of microsaccades has never been a glamorous affair, and the the history of microsaccade thinking has been quite spotty, with the occasional theory put forward now and then. In perhaps the most influential of the early works, Cornsweet (1956) theorized that microsaccades provide the function of correcting the intersaccadic drifts of the eye. It was subsequently proposed that microsaccades (also) serve to prevent vision from fading during fixation due to a static retinal image (Carpenter 1988). Other researchers go so far as to say that there is no function associated with microsaccades, or that they just reflect noise in the oculomotor plant (Kowler and Steinman 1980). Recently, however, microsaccades have drawn closer scrutiny and some other views of microsaccades have been proposed. For example, neurophysiological evidence has suggested that microsaccades might help maintain perception by modulating neural responses in the visual cortex (Martinez-Conde *et al* 2000, Leopold and Logothetis 1998). Nobody had suggested a direct causal link between covert attention shifts and the generation of microsaccades, however, so Ziad set out to test this idea.

4.1 Microsaccades During Fixation

The experiments carried out by Ziad basically involved requiring the subject to maintain fixation and presenting the subject with attention-grabbing peripheral stimuli. The subjects were given a (color) discrimination task which required them to attend to both the fixation target and the peripheral stimulus. We recorded eye movement information (using an ISCAN head-mounted eye-tracker) while the subjects carried out the task, and did an off-line analysis of the eye track data. The analysis detected the occurrence of microsaccades (defined as movements with a velocity above a set threshold, and with amplitudes less than 1 degree). Details on the experiments can be found in Hafed (2003), and Hafed and Clark (2001, 2002, 2003).

The key to unlocking the secrets of the microsaccade data lay in looking at the *time course* of the microsaccade occurrences, relative to the onset of the peripheral visual stimulus. When we do this, as shown in figure 1, a clearly visible peak in the microsaccade frequency is evident with a latency of around 250 msec. These latencies are very similar to those observed for regular saccades, suggesting that these microsaccades are directly related to the peripheral stimulus onsets. Furthermore, the horizontal *direction* of the microsaccades are biased in the horizontal direction of the peripheral stimuli, as shown in figure 2. From the data depicted in this figure we can see that the majority of microsaccades were first directed towards the peripheral stimulus, at a latency of about 200 msec. At a later time, around 400 msec after stimulus onset, another peak in microsaccade frequency

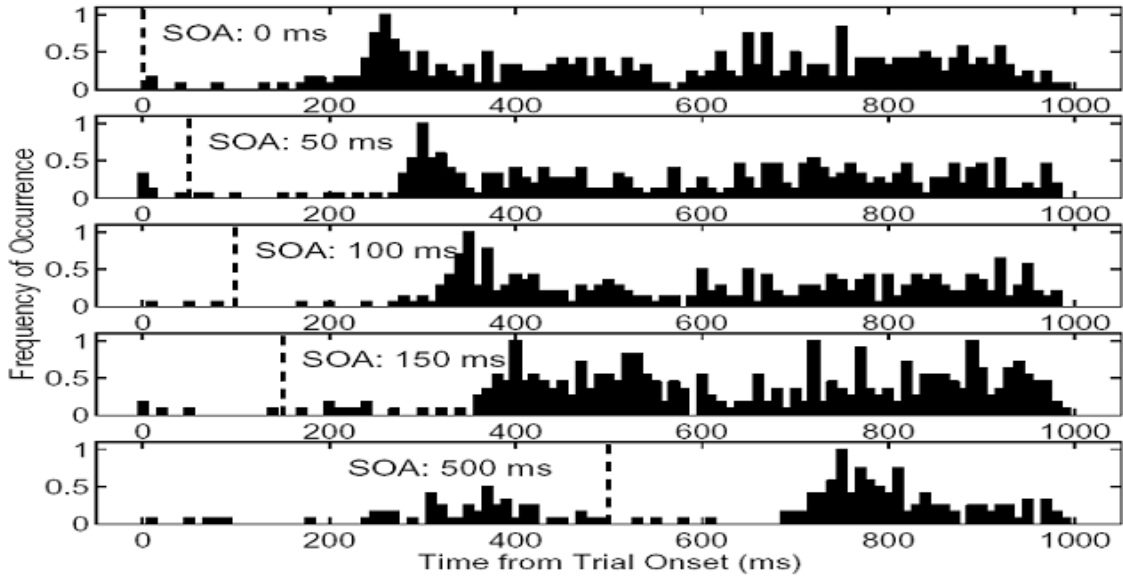


Figure 1: The frequency histogram of microsaccade occurrence aligned with the start of the trial. The dashed line indicates the onset of the peripheral stimulus, which occurs a variable length of time (the SOA) after the start of the trial.

appears, in which the majority of the microsaccades were made in the other direction, suggesting an attention shift back to fixation. Keep in mind that the subjects were instructed to maintain fixation, and that information about the objects at fixation was needed to accomplish the task. Therefore the fixation area was highly salient to the subject, and hence it should not be surprising that attention would shift back to the fixation area. In figure 1 many peaks are visible in the microsaccade time course plots, suggesting many shifts of attention during the performance of a task.

About the same time that we were carrying out our experiments, a group in Postdam, Germany were working along a similar track, and published their work very soon after we did (Engbert and Kliegl 2003). They also looked at the time course of microsaccade occurrence relative to a visual stimulus onset. Their results show a significant modulation of microsaccade activity related to the peripheral visual stimulus. As in our experiments, the direction of the microsaccades in the Engbert and Kliegl study appear to be biased in the direction of the stimulus. The work of Engbert and Kliegl provided, if not independent confirmation of, then at least additional independent evidence for, the link between covert attention and microsaccade occurrence.

4.2 Microsaccades Related to non-Visual Events

A closer examination of figure 1 shows significant tails of long latency microsaccades occurring well after the stimulus onsets (and, indeed, even before the stimulus onset). What do these long latency microsaccades correspond to? One might argue, as did Engbert and Kliegl (2003), that these long tail microsaccades simply correspond to a baseline, or steady state, population, the rate of which is merely *modulated* for a short while by attention shifts (revealing the low latency peaks). If this

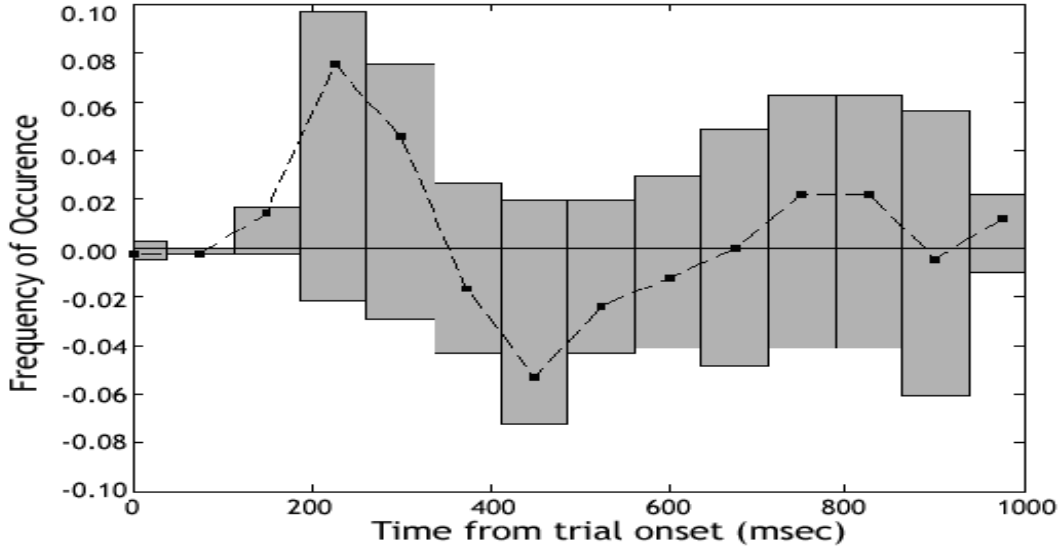


Figure 2: Microsaccades exhibited patterns of shifts towards the peripheral stimulus and back. The figure shows latency histograms of microsaccades as a function of peripheral stimulus location: positive bars indicate movements in its direction and negative bars indicate movements in the opposite direction. Bars are normalized by the total number of microsaccades shown in the figure. The thin line plots the difference between positive and negative bars for each latency bin.

is in fact the case then microsaccades may not be as useful a tool for indexing covert attention shifts as we are suggesting in the present work. This is so because with such a baseline rate, the observation of a microsaccadic occurrence at a particular time would only reflect an attention shift in some cases close to visual onsets but not others.

In neurobiology and psychophysics, baseline rates are often viewed with suspicion, and the suspicion is usually that a spread out population of events is a sign that the data is not aligned correctly. So Ziad decided to try aligning the microsaccade data to various significant task times. In doing so, we hit pay-dirt and found a rather remarkable result. If we align the microsaccade data with the time of the subject's response, the so-called baseline population resolves into a well-defined peak of microsaccade activity. This is shown in figure 3. The microsaccades that occurred after subject responses were concentrated around a fairly constant latency of 200 ms regardless of SOA. These microsaccades formed the diffuse long latency baseline population observed in Fig. 1. The implication is that these microsaccades that they were elicited by an event occurring around subject response time. This suggests that there is really no baseline rate for microsaccades but, more importantly, it implies that there are detectable attention shifts that are time-locked to response execution in psychophysical tasks. In particular, if one considers the 250 ms lag of microsaccades that we observed here as an estimate for the time difference between an attention shift and a microsaccade, then the data in figure 3 suggests that the response-related attention shifts in our task may have occurred approximately 50 ms before the execution of responses. The response-related microsaccades also show a significant modulation of microsaccade direction by peripheral stimulus location, generally in the direction of the peripheral stimulus. We hypothesize, based on this direction and on the relatively short time before the actual response, that the implied attention shift is to perhaps aid in a final check on the subject's decision before responding.

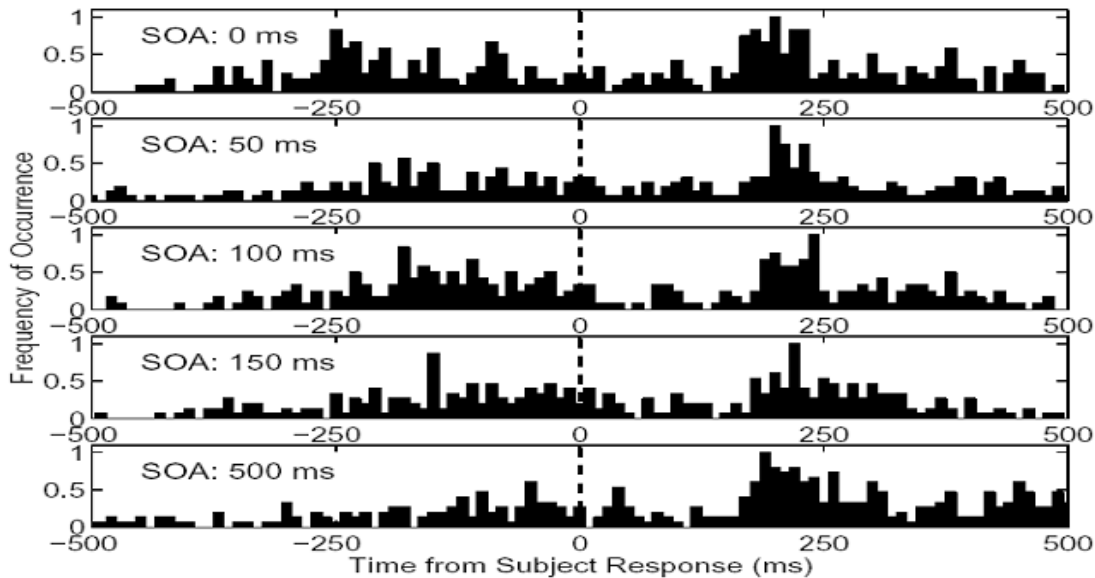


Figure 3: The frequency histogram of microsaccade occurrence aligned with the subjects' manual responses.

4.3 Microsaccades During Pursuit

The initial studies done on the link between microsaccades and attention, those by our group, and those of Engbert and Kliegl, only looked at situations in which the eye is fixed. But humans are not always fixating. In most natural viewing conditions, in addition to fixation and normal saccades, large human eye movements consist of smooth pursuit of moving targets, vestibular-ocular reflex motions in response to head motion, vergence motions shifting the plane of gaze in depth, and optokinetic-nystagmus in response to large-scale rapid image flow. Thus, if the use of microsaccades as a probe of attentional state is to have wide applicability, it is important that the characteristics of microsaccades be studied in the non-fixational viewing situations. Recently, our lab has been investigating (Jie and Clark 2005) the issue of whether microsaccades occur during pursuit and, if so, whether, and in what way, these microsaccades are related to covert attention shifts.

To determine whether humans engaged in pursuit generate microsaccades, current PhD student Li Jie created a simple psychophysical task in which subjects pursued a moving object, and measured their eye movements. To study possible links to covert attention shifts, such shifts were induced while the subjects maintained pursuit by the abrupt onset of a peripheral square, something that is known to exogenously capture attention (Yantis and Jonides 1984). The subjects were instructed to maintain pursuit, so the resulting attention shifts should be entirely covert. The details of the experiments are as follows. At the onset of every trial, a square cue appeared either on the left or right side edge of the display and remained visible for a variable period between 700 and 1100 msec in duration. The pursuit target, consisting of a cross shape appeared at the location of the cue and began to move horizontally as soon as the cue disappeared. The color of the cross shape changed periodically during the trial. The total duration for each trial after the cue's onset was 3.5 sec. At some random time during the trial, a square object briefly appeared at a distance of 11 degrees

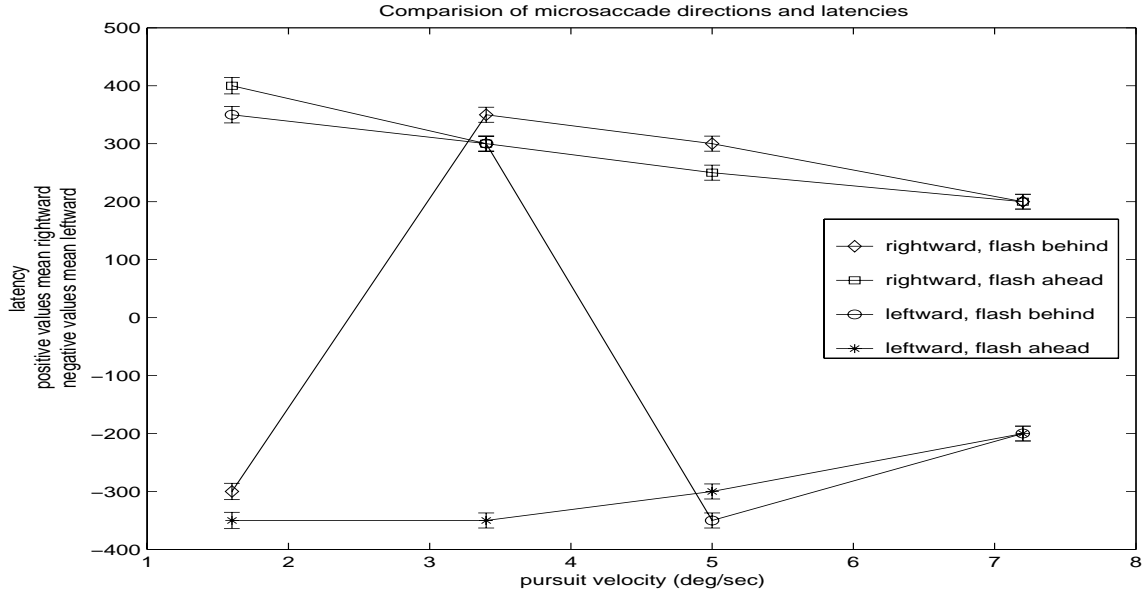


Figure 4: Comparison of the latencies and directions corresponding to the peak of the first population of microsaccades after stimulus onset, as a function of the pursuit velocity.

either to the left or to the right of the pursuit target. Subjects were instructed to maintain pursuit at all times on the moving cross and report as soon as possible after the onset of the distractor whether the pursuit target had the same color as the initial cue square.

The results of these experiments showed that microsaccades do indeed occur during ocular pursuit and that they have similar characteristics to those occurring during fixation. We observed that microsaccades appear either singly or in pairs of opposing movements during pursuit, similar to the patterns of microsaccadic motion discovered in fixation. After detecting all of the microsaccades that occurred during pursuit, we aligned all microsaccades to the onset time of the square object.

Following the same reasoning that we used in interpreting the microsaccade data during fixation we associate the first peak in the microsaccade time histogram with the covert attention shift related to the onset of the flashed square. Microsaccades with longer latencies from the square onset are presumably, in this view, associated with other covert attention shifts, perhaps back to the pursuit target.

In repeating the experiment at faster pursuit target velocities up to 7 degrees/second, we found that the microsaccades contributing to the short latency peak are tend to be biased to be in the direction of pursuit, and that this bias increases with the pursuit velocity. This phenomenon suggests that attention tends to shift more readily in the direction of pursuit with higher velocities. We also noted that the response time decreases as the pursuit velocity increases. Figure 4 shows the microsaccade directions and latencies as a function of the velocity of the pursuit target. Positive values on the vertical axis correspond to rightward microsaccades, while negative values correspond to leftward microsaccades. Actual latencies are the absolute value of the vertical dimension. It is clear that the directions and latencies of microsaccades are correlated with both the flash directions and the pursuit velocities. Microsaccade directions are biased in the direction of pursuit and that this bias increases with increases in pursuit velocities and microsaccade latencies decrease with increases of

pursuit velocities. A similar result has been found for regular saccades during pursuit (Tanaka et al (1998)). In that work it was shown that that saccades in the same direction as the pursuit direction had shorter latencies than those in the opposite direction. This similarity supports the argument that saccades and microsaccades have the same dynamics and that they are generated by the same system responsible for saccade generation (Zuber *et al* 1965).

Our findings regarding attentional allocation (as indicated by the characteristics of microsaccades) during pursuit are supported by the studies reported by van Donkelaar (1999) and van Donkelaar and Drew (2002). They found that covert attention leads pursuit targets, with a lead amount that increases with the pursuit velocity. We can use the van Donkelaar and Drew result to construct a hypothesis regarding what is going on in our pursuit experiments. We assume that when a microsaccade occurs it reflects a covert attention shift to one of three locations - the moving object itself, the location required to effectively pursue the moving object, or to the distracting object. van Donkelaar's research suggests that when pursuing a fast-moving object, attention is normally allocated to a location ahead of the moving object, with a distance that increases with pursuit velocity. As pursuit velocity increases, the salience of the location ahead of the pursuit object increases, due to the increased difficulty of maintaining pursuit (i.e. the pursuit control system gain becomes higher). If attention is actually at this location when the distractor appears, then most of the time no attention shift will be generated, and a few times an attention shift will be made to the distractor. If attention is located on the moving object when the distractor appears, then most of the time there will be a covert attention shift to the optimal pursuit lead location, and a few times to the distractor. In both situations the percentage of shifts to the distractor will drop as the pursuit velocity increases. This, then, would predict the type of bias towards the pursuit directions that we observe in our experiments.

5 What is Next?

The research described in this article makes the case for a strong link between attention and the generation of eye movements. We think that we can make use of this link both to further our understanding of the human visual system, and to construct intelligent man-machine interfaces. Knowledge of where people are attending is invaluable in presenting information to them. For example, in automotive or aviation applications, information should be presented in ways that are at once non-intrusive or distracting, yet quickly assimilated. If we know when, and where, people are attending to a display, we can present information at that location. Presumably, if someone is attending to a location then they can access information at that location more quickly and accurately. Conversely, if we wish to make changes to a display that are not distracting, we could use the change blindness phenomenon to our advantage and make the changes at locations deemed to have low probability of being attended to currently. Such information would then have to be accessed by a slow search strategy, but in some applications, such as demanding search-and-rescue helicopter piloting, this may be a reasonable price to pay for the lack of distraction.

Currently in our lab, we are developing statistical covert attention tracking models, which fuse many sources of information regarding the allocation of covert attention. These sources include macro- and micro- eye movement measurements to keep track of gaze, and to detect microsaccades, as well as image information. The image data is used to compute image based salience measures. Pursuit

eye movements are also detected, in which case attention leading of the pursued object is modeled. One of our goals is to apply the covert attention tracking system to real-time modification of visual displays in a video game. We are also looking into applying our ideas to automotive displays.

Finally, attention processes in the brain are not restricted to the visual modality. They operate in all sensory modalities. We are planning to extend our work to multimodal (visual, audio, tactile) attention control and tracking, and apply the resulting models to practical user-interface design for products such as cellular telephones. Stay tuned for future reports from our lab!

Acknowledgements

Portions of the research described in this article were carried out in the Harvard Robotics Laboratory (HRL), Nissan Cambridge Basic Research (CBR), and in the Centre for Intelligent Machines (CIM) at McGill University. We would like to acknowledge the contributions of Roger Brockett and Nicola Ferrier (now at the University of Wisconsin) from the HRL, and Ron Rensink (now at the University of British Columbia) and Kevin O'Regan (now at the Université de Paris V) at Nissan CBR.

All psychophysical studies performed by our group at McGill have received the approval of the Ethics Review Committee of the Faculty of Education at McGill University.

References

- Abrams RA (1992). Planning and producing saccadic eye movements, in *Eye Movements and Visual Cognition*, Rayner, K. ed., Springer-Verlag, pp 66-88.
- Abrams RA, Jonides J (1988). Programming saccadic eye movements, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 14, pp 428-443.
- Becker W, Jurgens R (1979). An analysis of the saccadic system by means of double step stimuli, *Vision Research*, Vol. 19, pp 967-983.
- Breitmeyer B, Ganz L (1976). Implications of sustained and transient channel for theories of visual pattern masking, saccadic suppression and information processing, *Psychology Review*, Vol. 83, pp 1-36.
- Brockett R (1988). On the computer control of movement. *Proceedings of the 1988 International Conference on Robotics and Automation*, Philadelphia, 534-540.
- Carpenter R (1988). *Movements of the Eyes*. Pion, London.
- Clark JJ (1998a). Spatial Attention and Latencies of Saccadic Eye Movements, *Vision Research*, 39(3):583-600.

- Clark JJ (1998b). Spatial attention and saccadic camera motion, Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, pp 3247-3252.
- Clark JJ (1999). Linking covert and overt attention, commentary in Behavioural and Brain Sciences, 22(4), p. 676.
- Clark JJ, Ferrier NJ (1988). Modal Control of an Attentive Vision System", Proceedings of the Second International Conference on Computer Vision, Tarpon Springs Florida, pp 514-523.
- Clark JJ, Ferrier NJ (1989). Control of Visual Attention in Mobile Robots, Proceedings of the 1989 IEEE Conference on Robotics and Automation, Scottsdale Arizona, pp 826-831.
- Clark JJ, Ferrier NJ (1992). Attentive visual servoing, in An Introduction to Active Vision, A. Blake and A.L. Yuille, eds, MIT Press, Cambridge Massachusetts, pp 137-154.
- Coeffe C, O'Regan JK (1987). Reducing the influence of non-target stimuli on saccade accuracy: predictability and latency effects, Vision Research, Vol. 27, No. 2, pp 227-240.
- Coren S, Hoenig P (1972). Effect of non-target stimuli upon length of voluntary saccades, Perceptual Motor Skills, Vol. 34, pp 499-508.
- Cornsweet TN (1956). Determination of the stimuli for involuntary drifts and saccadic eye movements. Journal of the Optical Society of America, 46, 987-993.
- Desimone R, Wessinger M, Thomas L, Schneider W (1989). Effects of deactivation of lateral pulvinar or superior colliculus on the ability to selectively attend to a visual stimulus, Society for Neuroscience Abstracts, 15:162.
- Desimone R (1990). Complexity at the neuronal level (commentary on Vision and complexity, by J.K. Tsotsos), Behavioral and Brain Sciences, Vol. 13, No. 3, p. 446.
- Deubel H, Wolf W, Hauske G (1984). The evaluation of the oculomotor error signal, in Theoretical and Applied Aspects of Eye Movement Research, Gale, A.G. and Johnson, F. (eds.), Elsevier Science Publishers B.V., pp 55-62.
- Engbert R, Kliegl R (2003). Microsaccades uncover the orientation of covert attention. Vision Research 43, 1035-1045.
- Ferrier NJ, Clark JJ (1993). The Harvard Binocular Head, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 7, No. 1, pp 9-32.
- Findlay JM (1982). Global visual processing for saccadic eye movements, Vision Research, Vol. 22, pp 1033-1045.
- Findlay JM (1992). Programming of stimulus-elicited saccadic eye movements, in Eye Movements and Visual Cognition, Rayner, K. ed., Springer-Verlag, pp 31-45.
- Fischer B (1993). Express saccades and visual attention, Behavioral and Brain Sciences, Vol. 16,

pp 553-610.

Hafed ZM (2003). Motor theories of attention: how action serves perception in the visual system. Ph.D. thesis, McGill University.

Hafed ZM, Clark JJ (2000). Neural network model of saccade generation and control. *Investigative Ophthalmology and Visual Science*, 41(Suppl.), 315.

Hafed ZM, Clark JJ (2001). Microsaccades reflect covert attention shifts, European Conference on Visual Perception, Turkey.

Hafed ZM, Clark JJ (2002). Microsaccades as an overt measure of covert attention shifts, *Vision Research*, Vol. 42, No. 22, pp 2533-2545.

Hafed ZM, Clark JJ (2003). Detecting patterns of covert attention shifts in psychophysical tasks using microsaccades. In: *Vision Sciences Society Annual Meeting*.

He P, Kowler E (1989). The role of location probability in the programming of saccades: Implications for Center-of-Gravity tendencies, *Vision Research*, Vol. 29, No. 9, pp 1165-1181.

von Helmholtz H (1896). *Handbuch der Physiologische Optik*, 2nd ed. Leipzig and Hamburg: Leopold Voss.

Henderson JM (1992). Visual attention and eye movement control during reading and scene perception, in *Eye Movements and Visual Cognition*, Rayner, K. ed., Springer-Verlag, pp 260-283.

Hikosaka O, Takikawa Y, Kawagoe R (2000). Role of the basal ganglia in the control of purposive saccadic eye movements, *Physiol Rev.* 2000 80(3), pp 953-78.

James W (1890). *The Principles of Psychology*, Volume 1, chapter 11, reprinted by Dover Publications, New York.

Jie L, Clark JJ (2005). Microsaccadic Eye Movements During Ocular Pursuit, 2005 Vision Sciences Society Annual Meeting (VSS), Sarasota, Florida.

Kalesnykas RP, Hallett PE (1994). Retinal eccentricity and the latency of eye saccades, *Vision Research*, Vol. 34 (4), pp 517-531.

Kingstone A, Klein RM (1993). Visual offsets facilitate saccadic latency: does predisengagement of visuospatial attention mediate this gap effect?, *Journal of Experimental Psychology, Human Perception and Performance*, Vol. 19, No. 6, pp 1251-265.

Klein RM (1980). Does oculomotor readiness mediate cognitive control of visual attention? in *Attention and Performance VIII*, Nickerson (ed.), Erlbaum, Hillsdale, N.J.

Klein R, Kingstone A, Pontefract A (1992). Orienting of visual attention, in *Eye Movements and Visual Cognition*, Rayner, K. ed., Springer-Verlag, pp 46-65.

- Koch C, Ullman S (1985). Shifts in selective visual attention: Towards the underlying neural circuitry, *Human Neurobiology*, Vol. 4, pp 219-227.
- Kowler E, Steinman RM (1980). Small saccades serve no useful purpose. *Vision Research*, 20, 273-276.
- Kustov AA, Robinson DL (1996). Shared neural control attentional shifts and eye movements. *Nature*, 384, 74-77.
- Leopold DA, Logothetis NK (1998). Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Experimental Brain Research*, 123, 341-345.
- Lee C, Rohrer WH, Sparks DL (1988). Population coding of saccadic eye movements by the superior colliculus, *Nature*, Vol. 332, pp 357-359.
- Lennie P (1980). Parallel visual pathways: A review, *Vision Research*, Vol. 20, pp 561-594.
- Martinez-Conde S, Macknik SL, Hubel DH (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nature Neuroscience*, 3, 251-258.
- Milios E, Jenkin M, Tsotsos JK (1990). Design and performance of TRISH, a binocular robot head with torsional eye movements. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 7, No. 1, pp 51-68.
- Munoz DP, Wurtz RH (1993). Fixation cells in monkey superior colliculus. I. Characteristics of cell discharge, *Journal of Neurophysiology*, Vol. 70, No. 2 pp 559-575 23
- Nakayama K, Mackeben M (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29, 1631-1647.
- O'Regan JK (1992). Solving the 'real' mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, 46, 461-488.
- O'Regan JK, Rensink RA, Clark JJ (1999). Change blindness as a result of mudsplashes, *Nature*, Vol. 398, p. 34
- Posner MI (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Posner MI, Cohen Y, Rafal RD (1982). Neural systems control of spatial orienting, *Philosophical Transactions of the Royal Society, London B*, Vol. 298, pp 187-198
- Rensink RA, O'Regan JK, Clark JJ (1995). Image flicker is as good as saccades in making large scene changes invisible, *Perception (suppl.)*, Vol. 24, pp 26-28, 1995 (abstract of the European Conference on Visual Perception, Tübingen Germany, August 1995).
- Rensink RA, O'Regan JK, Clark JJ (1997). To See or Not to See: The Need for Attention to Perceive Changes in Scenes. *Psychological Science*, Vol 8, pp 368-373.

- Reulen JPH (1984a). Latency of visually evoked saccadic eye movements. I. Saccadic latency and the facilitation model, *Biological Cybernetics*, Vol. 50, pp 251-262
- Reulen JPH (1984b). Latency of visually evoked saccadic eye movements. II. Temporal properties of the facilitation mechanism. *Biological Cybernetics*, Vol. 50, pp 263-271
- Reuter-Lorenz P, Hughes HC, Fendrich R (1991). The reduction of saccadic latency by prior offset of the fixation point: An analysis of the gap effect, *Perception and Psychophysics*, Vol. 49, pp 167-175
- Rizzolati G (1983). Mechanisms of selective attention in mammals, in *Advances in Vertebrate Neuroethology*, Ewart, J.P., Capranica, R.R., and Ingle, D.J., (eds.), Plenum, New York, pp 261-297.
- Rizzolati G, Riggio L, Sheliga BM (1994). Space and selective attention, in *Attention and Performance XV*, Umiltà, C. and Moscovitch, M., eds., MIT Press, Cambridge, MA, pp 231-265
- Saslow MG (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement, *Journal of the Optical Society of America*, Vol. 57, pp 1024-1029.
- Sereno AB (1992). Programming saccades: The role of attention, in *Eye Movements and Visual Cognition*, Rayner, K. ed., Springer-Verlag, pp 89- 107.
- Swain MJ, Stricker M (eds.) (1991). *Promising Directions in Active Vision*. University of Chicago Technical Report CS 91-27
- Tanaka M, Yoshida T, Fukushima K (1998). Latency of saccades during smooth-pursuit eye movement in man. *Exp Brain Res* 121, 92-98.
- Tsotsos JK (1990). Analyzing vision at the complexity level, *Behavioral and Brain Sciences*, Vol. 13, pp 423-469
- van Donkelaar P (1999). Spatiotemporal modulation of attention during smooth pursuit eye movements. *Neuroreport* 10, 2523-2526.
- van Donkelaar P, Drew A (2002). The allocation of attention during smooth pursuit eye movements. *Prog. Brain Res* 140, 267-277.
- Walker R, Kentridge RW, Findlay JM (1995). Independent contributions of the orienting of attention, fixation offset and bilateral stimulation on human saccadic latencies, *Experimental Brain Research*, Vol. 103, pp 294-310
- Wurtz RH (1996). Vision for the control of movement. *Investigative Ophthalmology and Visual Science*, 11, 2130-2145.
- Wurtz RH, Mohler CW (1976), Organization of monkey superior colliculus: Enhanced visual response of superficial layer cells, *Journal of Neurophysiology*, Vol. 39, pp 745-765

Wyman D, Steinman RM (1973). Latency characteristics of small saccades, *Vision Research*, Vol. 13, pp 2173-2175

Yantis S, Jonides J (1984). Abrupt visual onsets and selective attention: evidence from visual search, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 10, pp 601-621.

Zuber BL, Stark L, Cook G (1965). Microsaccades and the velocity-amplitude relationship for saccadic eye movements. *Science*, 150, 1459-1460.