Spatial Attention and Latencies of Saccadic Eye Movements

James J. Clark

Department of Electrical Engineering and Centre for Intelligent Machines, McGill University, Montréal, Quebec, Canada

Abstract

Recent theories of visual attention, such as the oculomotor readiness theory of Klein (1980), the premotor theory of Rizzolatti (1983), and the sequential attention theory of Henderson (1992), propose a link between shifts in spatial attention and the generation of saccadic eye movements. In this paper we show that a Winner-Take-All model of spatial attention, combined with a simple model for the link between attention and eye movements, can account for the variation in saccadic latency observed in many oculomotor phenomena. These phenomena include the gap effect (Saslow 1967), the effect of target jumps on saccadic latency (Becker and Jurgens 1979), the increase of saccadic latency as target eccentricity drops (Kalesnykas and Hallett 1994), and the modulation of saccadic accuracy using target predictability and saccadic latency (Coëffé and O'Regan (1987)).

Key words: attention, dynamics, saccade, latency

1 Introduction

A growing number of experiments support the view that there is a direct connection between saccadic eye movements and spatial attention (see, for example, Klein (1980), Rizzolatti *et al* (1983), Henderson (1992), Kustov and Robinson (1996)). In particular, it appears that the target location and the timing of saccadic eye movements are affected by visual attention. This apparent causal link between visual attention and saccadic eye movements suggests that many phenomena involving saccadic eye movements can best be understood as arising from the activity of visual attention. A detailed model of visual attention should, therefore, be able to predict, or explain, phenomena involving the timing of saccadic eye movements.

In this paper we examine, via computational modeling, a possible role for attention in affecting the timing and targeting of saccadic eye movements. We present a winner-take-all based model of spatial attention, similar to that proposed by Koch and Ullman (1985). We show that this model, when incorporating the spatio-temporal characteristics of low level feature detectors, can account for a wide range of oculomotor phenomena. In particular, we will use the model to account for the *gap effect* (Saslow 1967), the effect of target jumps on saccadic latency (Becker and Jurgens 1979), the increase of saccadic latency as target eccentricity drops (Kalesnykas and Hallett (1994), and the modulation of saccadic accuracy using target predictability and saccadic latency (Coëffé and O'Regan 1987).

2 Models of Visual Attention and the Link to Saccadic Eye Movement Generation

Many models have been proposed to explain various observed properties of saccadic eye movements. With regard to the connection between saccades and visual attention these models can be grouped into three types. The first type of model treats the generation of saccadic eye movements as independent of attention (see Becker and Jurgens (1979), Reulen (1984a,1984b), and Deubel (1984) for examples).

The second type of model posits a connection between saccades and visual attention, but requires only that attention be disengaged (or not directed to any particular focus) for a saccade to occur. The target of such a saccade is based on visual input. Perhaps the best example of such a model is that of Fischer (1992,1993). In Fischer's model the target of a saccadic eye movement is computed during the disengaged-attention phase by a localization system.

The third type of model assumes an even stronger involvement of attention, by requiring that attention be engaged at a target location before a saccade can be made to that location. This type of model has accrued much experimental support and there have been many different versions proposed.

One of the first theories of this sort was put forward by Wurtz and Mohler (1976), who proposed that attention shifts were programs for saccadic eye movements.

A detailed study of the link between spatial attention and saccadic eye movements was performed by Klein (1980), who proposed what he referred to as the *oculomotor readiness theory*. This theory posits that when attention to a particular spatial location is desired the observer prepares an eye

movement to that location, and this preparation, or oculomotor readiness, in turn enhances information processing at the target location. This oculomotor readiness hypothesis was later modified somewhat in light of subsequent experiments (Klein *et al* 1992) to state that attention is tightly linked to exogenously directed eye movement preparation, but is not linked to *endogenously* directed eye movement preparation. Indeed, the mechanisms underlying endogenous and exogenous orienting systems are not well known, and require further study. In the present paper, we are only modeling exogenous stimuli and hence feel safe in assuming that the oculomotor readiness theory is valid.

Posner (1980) provides a more relaxed view of the link between attention and eye movements. He treats the two orienting systems as being functionally related only. As in Klein's modification of his oculomotor readiness theory, Posner postulates a weaker connection between eye movements and covert attention mainly on the observations that endogenous attentional shifts can be made in ways that do not appear to affect eye movement preparations. Posner suggests that covert attention and eye movements are both drawn to exogenous (peripheral) stimuli, with covert attention moving more rapidly towards the stimulus. His model does not suggest a causal connection between eye movements and covert attention, however.

Perhaps the most extreme view of the connection between covert attention and saccadic eye movements is the *premotor theory* of Rizzolatti (1983). This theory, which actually includes other body movements in addition to eye movements, holds that the system that controls action is the same as that which controls spatial attention. In particular, one of the main claims of the premotor theory of attention is that "the mechanism responsible for spatial attention are localised in the spatial pragmatic maps. There are no such things as selective attention circuits defined as anatomical entities separated from the spatial maps" (Rizzolatti *et al* 1994). A pragmatic map is a neural representation of space that is used for carrying out some action. A particular example of such a pragmatic map would be the motor map present in the intermediate layers of the superior colliculus (Wurtz 1996). The premotor theory would then say that the activity of the superior colliculus directly affects the allocation of spatial attention. The premotor theory is rather controversial, and, in particular has been attacked on the grounds that it does not explain the results of the experiments done by Klein *et al* (1992). Rizzolatti has counterarguments, however, and the issue of the validity of the premotor theory is far from settled.

A more recent theory of the link between attention and eye movements is due to Henderson (1992), who proposed a *sequential attention* model wherein "programming" of a saccade begins when attention shifts once processing of the foveal input has been completed. The target of the eye movement is taken

to be the new locus of attention. In cases where the foveal processing is relatively light, the actual eye movement is modeled as occurring 80msec after this shift in attention. Henderson modified this sequential model somewhat, to account for the observation that increased foveal load (difficulty in processing foveal information) can reduce peripheral preview benefit, suggesting that eye movements begin to be planned before attention shifts, at some preset processing deadline time. This does not affect the conclusion that saccades are made to the locus of attention, however, but only modifies the mechanism which determines *when* eye movements are to be triggered.

The previous theories were based mainly on psychophysical evidence. There is a growing body of neuro-physiological evidence as well. Desimone *et al* (1989) found that local deactivation of small zones in the superior colliculus impaired an animal's ability to attend to a target in the presence of a distractor. Desimone (1990) points out that the oculomotor system and the covert attention system both involve the targeting of stimuli and could usefully share some common neural hardware. He also points out that the effects of a shift of gaze and a shift of covert attention are nearly identical on the visual system. Kustov and Robinson (1996) generated saccades in monkeys by electrical stimulation of the superior colliculus motor map. They found that both exogenous and endogenous attentional shifts caused deviations in the direction of the electrically evoked saccades. These deviations even occured when the monkey makes hand movements in response to the cue. Thus these deviations cannot be ascribed to conscious preparation of an eye movement. These findings are in accord with our view of the premotor theory of attention, in that preparation of a movement directly affects the allocation of attention, and vice-versa.

3 A Computational Model

In this paper, we assume a specific computational model linking spatial attention and eye movements that is very much along the lines of Rizzolatti's premotor theory. We will use simulations of this model to show that the premotor or ocular readiness approaches can account for a number of different oculomotor phenomena related to saccadic latency.

Our model brings together a number of ideas found in the literature. The key features of this model are:

- Attention is associated with spatial *pragmatic* maps (e.g. neural maps serving the execution of actions) (Rizzolatti *et al* 1987).
- Spatial attention is driven by a winner-take-all interaction between elements

of a spatial saliency map (Koch and Ullman 1985).
- Spatio-temporal integration for target specification (Deubel 1984).
- Distributed representation of target location (McIlwain 1975, and Lee, Rohrer and Sparks 1988).
- Triggering of reflexive saccades based on transitions of the winner-take-all network.

3.1 Attentional Dynamics

The purpose of this paper is to show that saccadic latency effects can be explained in a premotor theory by considering 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. In order to test this idea we need to have a model for the dynamical behavior of the attention shift process.

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 know as "Winner-Take-All". 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.

[[FIGURES 1 AND 2 GO HERE]]

We assume a rather simple form of a winner-take-all system. This model may differ in detail from the precise neural implementation to be found in the brain, but it is our belief that the qualitative nature of the dynamics of winner-take-all networks are sufficiently generic that the precise form of the implementation is irrelevant. The specific model of attention that we used in our simulations is depicted in schematic form in figure 1. In this model, feature maps of various kinds are computed and combined into a "saliency" map. Different features can be weighted by different amounts in producing this saliency map (as in Koch and Ullman 1985). There are two types of feature detectors posited 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. The raw saliency map values are then modulated via a shunting inhibition by an attentional signal. Shunting inhibition, which can be modeled by a division of the input by the inhibiting signal, has been observed in the retina (Amthor and Grzywacz 1991) and in cortex (Coombs *et al* 1955). The attention signal

is produced by a process that is depicted in figure 2. This inhibitory signal arises from a saturating, leaky, integrator. The integrator temporally accumulates the difference between a spatial average of the modulated saliency value at that location and a local estimate of the maximum modulated saliency value. If the spatial average modulated saliency at the location is greater than the local maximum then the integrator will discharge, reducing the inhibition. If, on the other hand, the average modulated saliency is less than the local maximum value, the integrator will begin to charge, increasing the inhibition. This positive feedback results in a winner-take-all process, wherein the location with the locally greatest saliency will inhibit its neighbours, reducing their activity even further and therefore strengthening its hold. When the input feature activity changes, the winner-take-all network will take on a new equilibrium, with new locations being inhibited and a new winning location established. It is the pattern of inhibition that corresponds to "attention" in our model. As this pattern changes due to changes in the input, so too does the allocation of spatial "attention". The spatial average operation spreads out the area of the winning location. If there is no spatial averaging then the winning location is condensed to a single spatial unit.

The question arises as to where in the brain this winner-take-all dynamical process takes place. If one takes the premotor theory in its strict form, this process would take place in the motor maps that control eye movements. The obvious candidate for this would be the intermediate layers of the superior colliculus. The colliculus model of Arai *et al* (1994) includes inhibitory lateral connections between neurons in the motor layer, which could give rise to winner-take-all behavior. The experiments of Desimone *et al* (1989) that were described earlier also implicate the superior colliculus in the control of attention. Other candidates for the locus of the winner-take-all process are cortical regions such as the parietal lobe. The only firm conclusion that can be made at this time is that the location of the attentional winner-take-all system is still very much an open research question.

3.2 Targeting and Triggering of Saccades

There are two major aspects to the generation of a saccadic eye movement. The first is specification of the target of the saccade, and the second is the specification of the time at which the saccade is to be executed. Other important factors include the control of the movement once it is underway, in order to ensure that the eye reaches its target. We will not consider these factors in this paper, but will instead concentrate only on the determination of the saccade target and the timing of the start of the saccadic eye movement.

A recurring theme in both the eye movement literature as well as in the more general neurophysiological community is that temporal and spatial factors are often separately programmed (see for example, the discussion of the "WHEN" and "WHERE" neural subsystems by Van Gisbergen *et al.* 1981). Thus it seems likely that the mechanism responsible for determining when a saccade is to be made is separate from the mechanism which determines the target of the saccade. This is reflected in our model, and we will consider these two aspects separately in what follows.

3.2.1 Saccade Triggering

Recent studies (Munoz and Wurtz 1993) have demonstrated the presence of a separate neural system dedicated to maintaining ocular fixation. Current models of the functioning of the superior colliculus (Wurtz 1996) emphasize the importance of the so-called "fixation cells" (Munoz and Wurtz 1993) in the rostral pole of the superior colliculus in triggering saccades. Activity in these cells suppresses saccadic eye movements. When these cells are inhibited, a saccade is generated. Premotor theory suggests that every attentional shift is associated with planning of an eye movement to the new location. A fixation system such as that found in the superior colliculus would act to prevent eye movements from occuring every time attention shifts.

In the above view of fixation, a saccade is triggered when ever fixation is released, through inhibition of the fixation cells. It is our view that one way in which this inhibition can be imposed is through the activity of the other cells in the motor layers of the superior colliculus. These cells are normally inhibited by the fixation cells, but they may also exert a reciprocal inhibitory effect on the fixation cells. It is conceivable, then, that a shift in attention from one location to another (or a transition in the winner-take-all network) may cause enough transient activity in the superior colliculus motor map to sufficiently inhibit the fixation cells, thereby triggering a saccade. The fixation cells could also be inhibited by cortical input, such as from the frontal eye fields. In this way volitional saccadic eye movements could be triggered.

In our model, we take the view that saccades can be triggered by transitions of the attentional winner-take-all system. That is, a reflexive or exogenous saccade is triggered when the level of attentional inhibition at any location drops from its maximum level (set by the saturation of the winner-take-all integrator) to zero. The dynamics of the winner-take-all are such that the value of the shunting inhibition signal at any location will, in steady state, be either at its minimum or maximum values. The requirement that the inhibition signal drop from its maximum to zero at a given location before a saccade is made to that location, provides a form of "inhibition-of-return". That is, a

location that was previously fixated must become inhibited before a saccade can be made to it. This should not be taken as a full explanation of the classical inhibition of return phenomena, however, as it does not require that a previously attended location become inhibited relative to other, unattended, locations. Note that the inhibition dropping to zero at a location only triggers the saccade. Depending on the feature activity elsewhere the target of the saccade may not be that particular location. In single target cases the saccade will usually be made to the location which triggers the saccade, however. Using the terminology of Posner (Posner *et al* 1982) our model states that a saccadic eye movement is triggered when attention is "engaged" at a new location. This is in contrast to models, such as that proposed by Fischer (1992), in which eye movements are triggered when attention is "disengaged" from its current location.

The experiments of Henderson (1992) suggest that eye movement planning can start before attention shifts when the foveal processing load is high. He suggests that there is a temporal deadline measured from the start of fixation at which the eye must move. This could be handled in our model by positing a signal which inhibits the fixation cells in the superior colliculus after a certain length of time since the previous saccade. This could be simply implemented with a temporal integrator that is reset after each saccade, and whose output inhibits the fixation cells.

3.2.2 Saccade Targeting

The premotor theories, and similar theories, do not say very much about how the *target* of the saccadic eye movement is determined. We propose the simple idea that the target of the saccadic eye movement is taken to be the centre of mass of the modulated feature activity, as reflected by the visual input to the superior colliculus. This centre of mass need not be computed explicitly, as the command for the eye movement can be represented in distributed form using a population coding (Lee, Rohrer, and Sparks 1988) of the motor command for the saccade, as is the case in the superior colliculus (Wurtz 1996).

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 (or shutting down of the fixation cell activity).

3.3 Endogenous vs. Exogenous Orienting

Most models of attentional orienting assume that there are two parallel systems (see e.g. Sereno (1992)), one handling endogenous (or volitional) orienting, the other exogenous (or reflexive, image-based) orienting. Our model follows this view, and assumes, furthermore, that both systems act on the same substrate, the facilitation or inhibition of low level feature detectors. In this paper we have explicitly modeled the dynamics of only the "exogenous" attentional activity. Although the dynamics of "endogenous" attention are not modeled, the effect of endogenous attention on eye movements can be straightforwardly included in our model by proposing that volitional eye movements occur through high-level modulation of the low-level substrate that underlies reflexive saccadic eye movements. The effect of endogenous inputs can be implemented by adding in a new input component to the saliency map that enhances the salience at the location to be volitionally attended to. Positional priming can also be handled in this fashion. Sustained or repetitive feature detector activity at a given location may build up, via a neural temporal integration mechanism, a short term increase in salience at that location.

There are, however, a number of experiments that suggest that this simple view needs to be elaborated somewhat. Rafal *et al* (1989) showed that inhibition of return does not occur following attentional shifts driven by endogenous cues, and Briand and Klein (1987) showed that exogenous and endogenous cues operated differently in feature integration. Likewise, the experiments performed by Klein (1980) showed that endogenous shifts in attention did not appear to facilitate saccadic eye movements to the attended locations. This suggests that the link between saccades and endogenous attention is not as direct as that between saccades and exogenous attention.

In light of these apparent differences between exogenously and endogenously directed attention, it is evident that the model we are describing in this paper should only be taken to refer to exogenous processes. The oculomotor phenomena that we describe and explain with our model in the following section all are concerned with exogenous stimuli only.

4 Phenomena Related to the Timing of Saccadic Eye Movements

In this section we describe a number of oculomotor phenomena that have been observed in humans. To our knowledge, none of the current models of saccadic eye movement generation can account for all of these oculomotor phenomena. We show, via computer simulations, that in each of these cases,

our model of attention and eye movement generation is able to account for the phenomenon.

4.1 Saccadic Latency and the Dynamics of Attention

Of interest is the time interval predicted by our model between the appearance of a target stimulus and the triggering of a saccade towards that target. The length of this time interval is often referred to as the *s*accadic latency. In the following sections we describe a number of oculomotor phenomena that are mainly concerned with saccadic latency. We will show that these phenomena are readily explained by our model.

In our model, saccadic latency depends on the dynamics of the attentional system. These dynamics are set primarily by two factors. The first is the temporal response of the feature detectors that feed into the winner-take-all network, and the second is the dynamics of the winner-take-all network itself.

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.

A detailed analysis of the dynamics of the winner-take-all is given in the Appendix. There it is shown that, if the initial feature value at a given location is too small compared with the value at the currently winning location, the network will not switch. If the feature value is high enough, switching will take place. Most importantly for understanding of saccadic latency phenomena, the switching time is seen to be proportional to the integrator time constant $(1/k)$ and the salience winning feature. The salience at a given location is defined here as the sum of the attentionally modulated feature outputs at that location.

4.2 The Gap Effect

Saslow (1967) 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. This phenomenon has come to be known as the "gap effect". Furthermore, it was observed that saccadic latencies increase when there is a temporal overlap between the fixation offset and the target onset.

Reulen (1984a) measured saccadic latencies as a function of the asynchrony between fixation offset and target onset in seven subjects. He found that the data fit a simple schematic model. This function, shown in figure 3, consists of three parts: 1) an overlap asymptote, 2) a transition region, and 3) a gap asymptote.

[[FIGURE 3 GOES HERE]]

The gap/overlap effect is readily apparent in simulations of our model. The time units for the simulation are arbitrary, however, and so no absolute comparisons of the simulation results with observed data can be made. In figures 4 and 5 we show the saccadic latencies predicted by our model for a range of stimulus asynchronies. In figure 4 we vary the saliency of the target stimulus, while in figure 5 we vary the saliency of the fixation stimulus. Note that the shape of the curves follow the form observed by Saslow (1967) and Reulen (1984a) and depicted in schematic form in figure 3. Note also the relative insensitivity of the value of the gap asymptotic latency to the fixation saliency compared with the value of the overlap asymptotic latency.

In figure 4 we see that varying the target salience essentially shifts the latency curve up or down, and has only a slight effect on the magnitude of the gap effect. This was observed by Kingstone and Klein (1993), and Walker *et al* (1995) who showed 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. In our model the role of the instructions given would be to increase salience at the target location, which would result in the shift of the latency curve shown in figure 4. A functionally similar result was observed experimentally by Reuter-Lorenz and coworkers (Reuter-Lorenz 1991), who showed that the gap effect is unaffected by the luminance of the target.

In both the gap and overlap conditions, our model predicts that the *relative* saliency of the target and fixation stimuli affects the saccadic latency. This effect has been observed in human subjects in many studies (e.g. Wheeless *et al.* 1967, Unema 1995). These studies have shown that saccadic latencies increase when the salience (e.g. luminance) of the target stimulus is reduced.

[[FIGURES 4 AND 5 GO HERE]]

The shape of the latency/asynchrony curves can be straightforwardly understood in terms of our dynamical model of attention. For example, the gap asymptote is approached when the fixation stimulus and the local maximum network have both decayed to zero at the time when the new target appears. In this case the new target stimulus is unimpeded in switching the winner-

take-all. The latency will then consist of only the time required for the target stimulus to discharge the shunting inhibition integrator at its location. Thus the latency at the gap asymptote will depend only on the time constant of the shunting inhibition integrator, the target saliency and the time constant of the transient channel. The switching time in this condition does not depend at all on the fixation stimulus salience. Likewise, the overlap asymptote is approached when the target stimulus is able to win the winner-take-all competition away from a continually present fixation stimulus (i.e. infinite overlap). As the saccade trigger is based on the state of the winner-take-all competition, any increase in overlap time beyond the overlap asymptotic latency merely results in the fixation stimulus persisting *after* the saccade has already been made, and so can have no effect on the saccadic latency. This is the reason for the asymptotic behavior.

The precise value of the overlap asymptotic latency depends on the switching time of the winner-take-all competition. As seen in the appendix the switching time is proportional to the shunting inhibition integrator time constant. If the target stimulus has a feature value that is much larger than the fixation stimulus value, the switching time is inversely proportional to the target stimulus value. For smaller feature values of the target stimulus relative to the fixation stimulus, the switching time is inversely proportional to both the target stimulus value and the ratio of the target to fixation stimulus values. For ratios of the target stimulus value to the fixation stimulus value that are close to the minimum required for switching, the switching time is inversely proportional to the difference between this ratio and the minimum ratio. Hence, in this case, switching times can be very large. If the ratio is too small, no switching of the winner-take-all will occur and no saccade will be generated.

4.3 Modulation of the Global Effect with Saccadic Latency

Coren and Hoenig (1972) observed that the amplitudes of saccades to point targets can be systematically affected by the presence of distractors. Saccades tend to bring the eye to the "centre-of-gravity" of the target+distractor complex. This phenomenon, called the *global effect* by Findlay (1982), has been observed in many other experiments and with various stimulus configurations.

Coëffé and O'Regan (1987) presented experimental results which point at two ways in which the effect of distractors on the landing position of the eye can be reduced. These are: increasing latencies and increasing the predictability of the target location. Their experimental paradigm was to have the subject make 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 overshot the cued location for targets on the end of the string nearest to the fixation point and undershot 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 overshoots. They also found that, when the trials were arranged in blocks wherein the target location was the same, the over- and under-shoots were reduced relative to cases in which the target locations were randomised from trial to trial. It was noted by He and Kowler (1989) that the influence of target probability does not necessarily improve the accuracy of saccades, as the actual target may lie in a location of low target probability. They found that saccades were biased towards locations that have a high probability of the target appearing. This suggests, as does the target predictability experiment of Coëffé and O'Regan, that visual memory or other higher level processes can affect the endpoint of a saccadic eye movement.

[[FIGURE 6 GOES HERE]]

We use the arrangement depicted in figure 6 as the input for our simulation. We use a grid of 15 points. The fixation stimulus is located at the 5th point and the distractors are located at positions 8 through 12. Initially the fixation saliency is set to a small, non-zero value (0.01), and the target and distractor saliencies are set to zero. All inhibition values are initialised to 1. After 3×10^6 time steps the fixation saliency is set to zero and the distractor saliencies are set to a value of 0.5 units. The target saliency is set to twice the distractor saliency at this time. The target location is varied from run to run of the simulation. The saccade triggering portion of the model is ignored for the purposes of these runs. We measure the centroid of the modulated saliency values every 50×10^3 time steps after the onset of the target. There is no overlap or gap between the onset of the target and distractors and the offset of the fixation. These centroids reflect, in our model, the landing position of the saccade (ignoring motor effects) assuming that a saccade is generated with the appropriate latency.

[[FIGURE 7 GOES HERE]]

As can be seen in figure 7, our simulations show the same behavior as observed by Coëffé and O'Regan. The over-shoot of near targets and the under-shoot of far targets is seen to decrease as latencies increase. This is due to two effects in our model. The first is the transient and sustained components of the feature detectors. For short latencies the feature detector response is dom-

inated 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. In our simulation, the first factor dominates at longer latencies (after the peak of the transient response, which occurs at around 500×10^3 time steps in our simulation), and the second factor dominates at short latencies.

Our simulations also showed that, when there are no distractors present, the centroid of the modulated feature activity is close to the target location for all latency values. This shows that the under- and over-shoots seen in figure 7 are due to the distractors.

The results of Coëffé and O'Regan for the case of target plus distractors exhibit a systematic shift in the eye-landing-position/target-position curves towards the fixation point. This cannot be explained by the effect of the salience of the fixation point, as the shift is observed at all latencies, and does not appear in the case of no distractors. Coëffé and O'Regan suggest that this effect is due to cortical magnification. Points in space closer to the fixation point will have more photoreceptors and hence more cortical neurons associated with it, and so will be weighted more heavily in a centroid calculation. As our simulation does not model this cortical magnification, it should not produce any shift in the centroids towards the fixation point.

We also simulated conditions in which the salience of the target relative to the distractors was varied. The results were that the over- and under-shoots are reduced when the salience of the target is increased relative to the distractors. It is our view that this provides an explanation of the target location predictability effect observed by Coëffé and O'Regan, as well as the results of the experiments done by He and Kowler (1989). In this view spatial priming due to repetitive target presentation at a specific location results in an enhanced salience for features at that location (by an endogenous process which is not modeled by us). This enhancement of the target location has the same effect as the increase of target salience used in our simulation.

Our model predicts that targets defined by equiluminant colour changes should not exhibit the global effect, as this effect relies on the transient component of the saliency map. It is well known that the response characteristics of colour opponent cells in the retina are primarily sustained (Gouras (1968)), with little transient response. Thus saccadic latencies to targets defined by

equiluminant colour changes only should be long, and the influence of distractors should be minimal. Supporting this view is an experiment described by Theeuwes (1995) which shows that equiluminant colour changes do not capture attention.

4.4 Retinal Eccentricity and Saccadic Latency

Wyman and Steinman (1973) noted a small narrow central peak in saccadic latency as a function of retinal eccentricity. 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 retinal eccentricity.

Our model provides an explanation for the increase in saccadic latency for targets with small retinal eccentricities. In the appendix it is shown that saccadic latencies reflect the time taken for the winner-take-all 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 response of the local maximum network to an impulse (e.g. from a point stimulus at the fixation location) decays exponentially with the distance away from the impulse location. If the target is far from the fixation, the local maximum value will be that of the target salience. 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.

[[FIGURE 8 GOES HERE]]

In figure 8 we show the results of a simulation in which the location of the target stimulus relative to the fixation stimulus was varied. The onset of

the target stimulus coincided with the offset of the fixation stimulus in this simulation. Two different target stimulus values were used, one just above the fixation stimulus value and one 20 times this level. We see that our model produces a rise in saccadic latency as target eccentricity decreases similar to that observed in humans. In addition, it is seen 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).

Our model does not take into account the dependence of photo-receptor density on retinal eccentricity found in the human retina, and so may not reproduce all aspects of the variation in saccadic latency with eccentricity observed in experiments on humans.

4.5 Saccadic Programming and Saccades to Stepped Targets

The process of specifying the position of the target of a saccadic eye movement has often been referred to in the eye movement literature as "saccadic programming" (see, for example, Abrams (1992), Abrams and Jonides (1988), Findlay (1992), He and Kowler (1988), and Sereno (1992)). This programming process has typically been viewed as consisting of two components, amplitude programming and direction programming. In addition, these computational components or modules, are usually thought of as distinct processes that are initiated, run for a while, and then provide a result. The underlying idea is that saccadic latencies reflect the time taken by these processes to produce the required amplitude and direction parameters for the saccade. The point of view inherent in the idea of saccadic programming can be seen clearly in the following quote (from Abrams (1992)):

> "Table 5.3 shows the probability that subjects cancelled or modified the the initial motor program and looked directly to the final target location. ...as more and more motor programming has been completed, it becomes more difficult to halt the execution of the program."

Note the key phrases, "initial motor program", "motor programming has been completed", and "halt the execution of the program". These phrases all imply the existence of a distinct motor programming process or computational module.

This view of saccade generation, however, is by no means the only view that can be used to explain the observed data. It is our contention that explicit saccadic programming of either saccade amplitude or direction is overly complicated and unneccesary. Rather, we propose that the pattern of the low

level feature detector activity, as modulated by visual attention, is used to determine both the timing and target of the saccadic eye movement. Following the proposal of Lee, Rohrer, and Sparks (1988), our 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. Thus the target will be specified by the centre of mass of this pattern of activity, and is continually available. The execution of the saccade is *triggered* by the shift in attention to the new location. The actual saccade target will then be the attentional centre-of-mass at this moment in time. No saccadic "programming" need take place. The amplitude and direction of saccades are implicit in the 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.

The oculomotor phenomenon most often used as support for the saccadic programming theory is the behavior observed in "double-step" experiments, such as those performed by Becker and Jurgens (1979). In this type of experiment the target initially jumps to a position P1 and then subsequently jumps to a position P2 before the saccade is made. The subjects are instructed to move their eyes to the target as soon as it appears. In these experiments subjects typically move their eyes either to the first target location, followed by a second saccade to the second target location, or directly to the second target location. In many cases, however, the eye lands in a location somewhere between the two target locations. The amplitude of the first saccade of a double step response appears to be determined primarily by the *secondary latency*, which is the time delay between the second target step and the onset of the response (i.e. the saccadic latency minus the interstep interval). For short delay times the eye moves to the first target location while for longer delay times the eye moves to the second target location. For intermediate delay times there is a transition region where the eye moves to a location somewhere between the two target locations. The minimum secondary latency for which the eye moves away from the first target location is called the *modification time* by Becker and Jurgens. This terminology arises from the observation that, from the saccadic programming viewpoint, the modification time is the minimum delay that must elapse if the second step is to modify the amplitude of the saccade that is being prepared in response to the first step. It was found by Becker and Jurgens that this modification time depended greatly on the nature of the second step as compared with the first. The modification time was longest, on the average (203 msec), when the second step was in the same direction as the first (the "Lengthen" case). A somewhat shorter modification time (172 msec) occured for the case that the second step was in the opposite direction as the first and crossed over the original fixation point (the "Change-direction" case). Shorter still (81 msec) was the case where the second step was back towards the original fixation location (the "Shorten" case).

We ran three different simulations to see whether our model could replicate these observations. In each of these simulations, the fixation mark is located at point 7 (counting from zero) of a 15 point discrete grid, and the first target location is at point 10. In the first simulation, the second target location is further away from the fixation, at point 12. In the second simulation, the second target location is nearer to the fixation, at point 8. In the third simulation the second target location is on the other side of the fixation, at point 4 (which is at the same distance from the fixation location as the first target location). For each simulation we varied the time interval between movements of the target. We summarize the results of these three simulations in figure 9. This figure shows the centroid of the salience map as a function of the secondary latency, or delay time. This corresponds to the "amplitude transition functions" of Becker and Jurgens, from which we can estimate the modification times.

[[FIGURE 9 GOES HERE]]

4.5.1 Discussion

In our simulations we observe that the secondary latency decreases as the time between target steps increases. Another way of interpreting this is to say that secondary latencies are smaller when the first saccade is to the initial target and are larger when the first saccade is to the final target. This is in accord with the Becker and Jurgens experiments where it was found that for long reaction times (primary latencies) the response was directed to the final target location, while for short reaction times the response was directed to the initial target location. In these experiments the inter-step time was fixed, and reaction times varied randomly, with some response occurring quickly and some more slowly. In our simulation no random elements were introduced and thus fixing the inter-step interval also fixed the reaction time. Thus we had to vary the inter-step interval to obtain the dependence of saccade target location to reaction time. It is expected that adding in random variation to our model will result in the same sort of variation of saccade targeting as a function of reaction time for fixed inter-step interval as in the Becker and Jurgens experiment.

The shape of the curves shown in figure 9, are seen to be comparable to the amplitude transition functions observed by Becker and Jurgens. From these curves we can determine roughly the modification times for each case, by estimating the secondary latency at which the tangent to the curves in the transition between initial and final target responses intersects the low secondary latency asymptote. The modification time so determined is seen to be longest for the "Lengthen" case (about 125 time steps) and shortest for the "Shorten" case (about 50 time steps), as in Becker and Jurgen's experiments.

In our simulations saccadic latencies were longest for the "Shorten case", at least for final target responses. This is explained by the inhibitory effect of the decaying trace of the fixation stimulus. Presumably this effect is only significant if the second target location is very close to the fixation stimulus, that is, within two degrees of visual angle (Kalesnykas and Hallett (1994)). In the experiments of Becker and Jurgens, the location of the second target in the shorten case was about 15 degrees away from the fixation point, and so should not exhibit any increase of saccadic latency due to low eccentricity.

Becker and Jurgens claim that the amplitude of the first saccade of a double step response is determined primarily by the time delay between the second target step and the onset of the response (i.e. the saccadic latency minus the interstep interval). The model that they propose to explain these results invokes a bilateral (directional) decision mechanism combined with a time averaged amplitude computation. The decision mechanism takes in a retinal error signal and compares it to a pair of thresholds (one for left-ward errors, one for right-ward errors). If one of these thresholds is exceeded, a decision signal for the appropriate direction is generated after some delay. The threshold signal immediately inhibits the other direction decision signal. Thus, if the retinal error changes sign (as in a crossed double step) the original direction decision signal will be blocked, and the new direction decision signal will be generated, after a fresh delay period. Thus, the decision time for a crossed double step will be longer than for an uncrossed double step. Once a decision signal has been generated the average of the retinal error over a time window (of 110 msec) is computed and used to specify the amplitude of the saccade. The actual saccadic motion is triggered at the end of the time window. Note that, in this model, the directional programming of the saccade occurs before the computation of the saccade amplitude. Hence, this model predicts that a change in target direction will lead to a greater latency than a change in target displacement.

However, as Sereno (1992) points out, targets lying in a different direction (but at the same distance from fixation) are typically far apart. Hence saccadic latencies, which in our model depend on the time required for attention to shift, should be longer in moving to the new target. When only amplitude is changed, the distance from the first target to the new target is smaller, and hence latencies should be shorter, as compared with the case where direction is changed. Sereno refers to this problem with the interpretation of the double-step results as an "attentional confound". In our view, there is no "attentional confound"; rather it is the pattern of attentional activity that determines the parameters of the saccade. Our simulations with stepped target stimuli support this view. There is no need to assume separate direction and amplitude saccadic programs, and, indeed, no need for any saccadic program at all. The view that attentional activity increases at the first target location, tipping

the scales of the winner-take-all network towards that location, is sufficient to explain the double step phenomena. The longer the time period between the target steps, the greater the buildup of attentional activity at the first target location, and hence the greater the time needed to shift to the second location, and the greater the probability that the saccade will be made to the first target location.

It should be noted that our simulations of the double step paradigm have only looked at the first saccade. The timing of the second saccade has not been investigated. The reason for this is that the model as it stands is a retinotopic model. This means that the circuitry for generating the attentional inhibition signal is fixed relative to retinal coordinates. As the eye moves in a saccade, the location of the target on the retina will shift relative to the winning location of the winner-take-all, and will therefore cause another attentional shift (and another saccade). A more realistic model would perhaps place the attentional signal generation in a spatiotopic map, such as one using a head-centred coordinate system. In this case the location of the winner-take-all would shift along with the eye movement and so no extra shift in attention would be needed to be made after the saccade. The timing of the second saccade in a double step response will therefore depend on whether we assume a retinotopic or spatiotopic substrate for generation of the attention signals. We are currently investigating the reformulation of our model in spatiotopic coordinates, and will revisit the double-step experiments when this is complete.

5 Summary

Motivated by premotor theories of attention, this paper has shown that a low-level, winner-take-all based model of attention, combined with a simple approach to the implicit specification of saccade parameters can account for a wide range of phenomena related to saccadic eye movements.

The viewpoint espoused in our model differs fundamentally from most existing models of saccadic eye movement generation. Its principal aspects are that:

- The targets of saccades are not "programmed" by any modular process, but are continuously defined by the pattern of activity of the attentionally modulated feature values.
- Saccades are triggered when attention is "engaged" at a new location (unless suppressed by a volitional fixation or gating signal).

We showed, via computer simulations, that our model can replicate a wide

range of oculomotor behavior, such as the gap effect, the global effect, the effect of target eccentricity on saccadic latency, and the temporal characteristics of the initial saccade in the response to stepped target motion. In replicating these phenomena with our model, we are able to provide insights as to their underlying mechanisms. For example, from the viewpoint of our model, the reasons for the gap effect become apparent. 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. 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. The increase in saccadic latency as the eccentricity of the target decreases is seen to arise due to the greater strength of the decaying trace of the fixation stimulus in the winner-take-all competition as the target gets closer to the fixation location. Likewise, the global effect is easily understood as resulting from the spatiotemporal characteristics of the feature detectors, and the specification of the saccade target in terms of the overall spatial pattern of the attentionally modulated activity of the feature detectors. Phenomena which have been heretofore used as a primary justification of the "saccadic programming" theory, such as the oculomotor response to stepping targets, can be accounted for by our model in a way which obviates the need for explicit programming of saccades. In our model, changing the location of a target before a saccade has been generated to that target merely changes the landscape of attentional modulation activity, and delays the time needed for inhibition at a given location to be sufficiently reduced to trigger a new saccade to that location.

The results of the simulation of our model also provide strong support for theories which propose a strong link between the operation of spatial attention and saccadic eye movement systems, as they show that these theories can account for a wide range of oculomotor phenomena.

6 Acknowledgements

The bulk of this work was performed at the Nissan Cambridge Basic Research laboratory. The research was partially funded by NSERC grant number 229-66. The author would like to thank Jack Beusmans, Ken Nakayama, and Whitman Richards for their comments on this manuscript, as well as the reviewers for their valuable suggestions.

References

[1] Abrams, R.A. (1992), "Planning and producing saccadic eye movements", in **Eye Movements and Visual Cognition**, Rayner, K. ed., Springer-Verlag, pp 66-88

[2] Abrams, R.A. and Jonides, J. (1988), "Programming saccadic eye movements", *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 14, pp 428-443

[3] Amthor, F.R. and Grzywacz, N.M. (1991), "Nonlinearity of the inhibition underlying retinal directional selectivity", *Visual Neuroscience*, Vol. 6, No. 197-206

[4] Arai, K., Keller, E.L. and Edelman, J.A. (1994), "Two-dimensional neural network model of the primate saccadic system", *Neural Networks*, Vol. 7, No. 6/7, pp 1115-1135

[5] Becker, W. and Jurgens, R. (1979) "An analysis of the saccadic system by means of double step stimuli", *Vision Research*, Vol. 19, pp 967-983

[6] Breitmeyer, B. and 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

[7] Coëffé, C. and O'Regan, J.K. (1987), "Reducing the influence of non-target stimuli on saccade accuracy: predictability and latency effects", *Vision Research*, Vol. 27, No. 2, pp 227-240

[8] Coombs, J.S., Eccles, J.C., and Fatt, P. (1955), "The inhibitory suppression of reflex discharges from motoneurones", *Journal of Physiology*, Vol. 130, pp396-413

[9] Coren, S. and Hoenig, P. (1972), "Effect of non-target stimuli upon length of voluntary saccades", *Perceptual Motor Skills*, Vol. 34, pp 499-508

[10] Deubel, H., Wolf, W., and 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

[11] Desimone, R., Wessinger, M., Thomas, L. and 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*, Vol. 15, p 162.

[12] 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.

[13] Findlay, J.M. (1982), "Global visual processing for saccadic eye movements", *Vision Research*, Vol. 22, pp 1033-1045

[14] Findlay, J.M. (1992), "Programming of stimulus-elicited saccadic eye movements", in **Eye Movements and Visual Cognition**, Rayner, K. ed., Springer-Verlag, pp 31-45

[15] Fischer, B. (1992), "Saccadic reaction time: Implications for reading, dyslexia, and visual cognition", in **Eye Movements and Visual Cognition**, Rayner, K. ed., Springer-Verlag, pp 31-45

[16] Fischer, B. (1993), "Express saccades and visual attention", *Behavioral and Brain Sciences*, Vol. 16, pp 553-610

[17] He, P. and 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

[18] Henderson, J.M. (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

[19] Kalesnykas, R.P. and Hallett, P.E. (1994), "Retinal eccentricity and the latency of eye saccades", *Vision Research*, Vol. 34 (4), pp 517-531

[20] Kingstone, A. and Klein, R.M. (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

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

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

[23] Koch, C. and Ullman, S. (1985), "Shifts in selective visual attention: Towards the underlying neural circuitry", *Human Neurobiology*, Vol. 4, pp 219-227

[24] Kustov, A.A. and Robinson, D.L. (1996), "Shared neural control of attentional shifts and eye movements", *Nature*, Vol. 384, pp 74-77

[25] Lee, C., Rohrer, W.H., and Sparks, D.L. (1988), "Population coding of saccadic eye movements by the superior colliculus", *Nature*, Vol. 332, pp 357-359

[26] Lennie, P. (1980), "Parallel visual pathways: A review", *Vision Research*, Vol. 20, pp 561-594

[27] McIlwain, J.T. (1975), "Visual receptive fields and their images in the superior colliculus of the cat", *Journal of Neurophysiology*, Vol. 38, pp 219-230

[28] Munoz, D.P. and Wurtz, R.H. (1993), " Fixation cells in monkey superior colliculus. I. Characteristics of cell discharge", *Journal of Neurophysiology*, Vol. 70, No. 2 pp 559-575

[29] Posner, M.I. (1980), "Orienting of attention", *Quarterly Journal of Experimental Psychology.* Vol. , No. 1, pp 3-25

[30] Posner, M.I., Cohen, Y., and Rafal, R.D. (1982), "Neural systems control of spatial orienting", *Philosophical Transactions of the Royal Society, London B*, Vol. 298, pp 187-198

[31] Reulen, J.P.H. (1984a), "Latency of visually evoked saccadic eye movements. I. Saccadic latency and the facilitation model", *Biological Cybernetics*, Vol. 50, pp 251-262

[32] Reulen, J.P.H. (1984b), "Latency of visually evoked saccadic eye movements. II. Temporal properties of the facilitation mechanism.", *Biological Cybernetics*, Vol. 50, pp 263-271

[33] Reuter-Lorenz, P., Hughes, H.C., and 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

[34] 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

[35] Rizzolati, G., Riggio, L., Dascolo, I. and Umilta, C. (1987), "Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention", *Nueropsychologia*, Vol. 25, pp 31-40.

[36] Rizzolati, G., Riggio, L., and Sheliga, B.M., (1994), "Space and selective attention", in **Attention and Performance XV**, Umilta, C. and Moscovitch, M., eds., MIT Press, Cambridge, MA, pp 231-265

[37] Saslow, M.G. (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

[38] Sereno, A.B. (1992), "Programming saccades: The role of attention", in **Eye Movements and Visual Cognition**, Rayner, K. ed., Springer-Verlag, pp 89-107

[39] Shepherd, M, Findlay, J.M., and Hockey, R.J. (1986), "The relationship between eye movements and spatial attention", *Quarterly Journal of Experimental Psychology*, Vol. 38A, pp475-491

[40] Theeuwes, J, (1995) "Abrupt luminance change pops out; abrupt color change does not", *Perception and Psychophysics*, Vol. 57(5), pp 637-644

[41] Unema, P.J.A. (1995), **Eye Movements and Mental Effort**, Dr. Phil. Dissertation, Technischen Unversitat Berlin, Verlag Shaker Aachen

[42] Tsotsos, J.K. (1990), "Analyzing vision at the complexity level", *Behavioral and Brain Sciences*, Vol. 13, pp 423-469

[43] Van Gisbergen, J.A.M., Gielen, S., Cox, H., Bruijns, J., and Kleine Schaars, H. (1981), "Relation between metrics of saccades and stimulus trajectory in target tracking: implications for models of the saccadic system.", **Progress in oculomotor research**, Fuchs, A.F., and Becker, W., (eds.), Amsterdam, Elsevier, pp 17-27

[44] Walker R., Kentridge, R.W., Findlay, J.M., "Independent contributions of the orienting of attention, fixation offset and bilateral stimulation on human saccadic latencies", Experimental Brain Research, Vol. 103, pp 294-310

[45] Wheeless, L.L., Cohen, G.H., and Boynton, R.M. (1967), "Luminance as a parameter of the eye movement control system", *Journal of the Optical Society of America*, Vol. 52, pp 210-213

[46] Wurtz, R.H. (1996), "Vision for the control of movement", *Investigative Ophthalmology and Visual Science*, Vol. 11, pp 2130-2145

[47] Wurtz, R.H. and Mohler, C.W. (1976), "Organization of monkey superior colliculus: Enhanced visual response of superficial layer cells", *Journal of Neurophysiology*, Vol. 39, pp 745-765

[48] Wyman, D. and Steinman, R.M. (1973), "Latency characteristics of small saccades", *Vision Research*, Vol. 13, pp 2173-2175

[49] Yantis, S. and 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

Appendix A. Simulation Details

In each of the simulations shown in this paper we used a discrete one-dimensional grid of 15 elements. The integrators were implemented with Euler's method using a small time step size compared with the time constants of the systems involved (typically around 5×10^5 time steps).

The feature detectors were modeled as simple step functions passed through either a sustained or transient temporal filter. The sustained temporal filter was modeled as an exponential averager implemented with the following difference equation:

$$y_t = A_S y_{t-1} + (1 - A_S) x_t$$

where x_t is the input to the filter at time step t and y_t is the filter output at time step t. A_S is the sustained filter coefficient. In the simulation we set $A_S = 0.9999984$, which corresponds to a time constant of $\tau_S = -1/\ln(A_S) = 6.25 \times 10^5$ time steps. The transient temporal filter was modeled as a simple differentiator followed by a difference of two exponential averaging filters, having coefficients $A_{T1} = 0.9999978$, and $A_{T2} = 0.9999982$. This yields a bandpass filter whose impulse response reaches a peak at $\tau_T = (\ln\ln A_{T2} - \ln\ln A_{T1})/(\ln A_{T1} - \ln A_{T2}) = 5.02 \times 10^5$ time steps.

The shunting inhibition factor α was set to 5 in the simulations, and the inhibition integrator constant k was set to 0.0005.

[[FIGURE A-1 GOES HERE]]

We implemented the spatial averaging process of figure 1 by means of the resistive grid circuit shown in figure A-1a. Such circuits have often been used as models of electronic spread in dendritic trees and horizontal cells in the retina (Mead, 1989). Likewise, the local-maximum function was implemented, in our simulations, by a nonlinear difference equation modeling the resistive netowrk shown in figure A-1b.:

$$\frac{dM_i}{dt} = \frac{1}{C}\left(\frac{D(V_i - M_i)}{R_V} + \frac{(M_{i-1} - 2M_i + M_{i+1})}{R_H} - \frac{M_i}{R_L}\right)$$

The subscript i indexes the individual nodes in the network. In the continuous network limit this equation becomes:

$$\frac{dM_i}{dt} = \frac{1}{\tau}\frac{d^2M}{dx^2} - \frac{R_H}{\tau R_L}M + \frac{R_H}{\tau R_V}D(V - M)$$

where $\tau = R_H C$. This is a diffusion equation with a dissipative term and a forcing function $D(V-M)$, where $D(y)$ indicates a rectification operation (i.e. $D(y) = y$ if y < 0, and 0 otherwise). There is mounting evidence that neurons exhibit the type of rectifying behaviour (Nichols and Lopatin, 1997) that is required in this model. In addition, there have been a number of models of cortical function that invoke the rectification properties of neurons (Carandini *et al* (1996, and Nestares and Heeger, (1997)).

Diffusion serves to propagate the maximum input value across the network. The forcing function affects the diffusion only if the input value V is greater than the local maximum value M. The dissipative term forces the local maximum value M to zero when all inputs are zero. The steady state response of the network to an spatial impulse is an exponential decay away from the location of the impulse, of the form $M(x) = M_0 \exp\left(-\sqrt{\frac{R_H}{R_L}}x\right)$. The decay rate, and hence the "locality" of the local maximum computation, is set by the leakage resistance R_L. If R_L is large the spatial scale of the local maximum function is large. If the input at a particular location is less than the local maximum value, M, then the forcing function at that location will be zero, and the output there will follow the exponential decay away from the winning location.

Please note that the maximum network provides a localized measure only, but has the advatage of requiring only local interactions between neural elements.

The resistive grid spatial averaging and local maximum networks that we modeled in our simulations had no dynamics associated with them (i.e. they had no capacitive or inductive elements). For more complete modeling, dynamics should be added into these networks. It is our belief, however, that the time constants of these networks would be much lower than the time constants of the feature detectors and of the winner-take-all betwork, and would therefore have little effect on the qualitative details of our simulations.

We made no effort in our simulation development to relate the simulation time steps to time constants that would be observed in a biological system. Thus, the simulation results can be used only for qualitative comparisons and predictions of psychophysical observations.

Appendix B. Analysis of the Winner-Take-All Dynamics

In this appendix we examine the dynamics of a simplified version of the winner-take-all network used in our simulations.

We examine the case where we have two competing feature locations close enough so that we can ignore the spatial falloff in the local maximum network. We also assume that there is no spatial averaging. With these assumptions we can model the winner-take-all with the following equations:

$$y_1 = \frac{x_1}{1 + \alpha I_1}$$

$$y_2 = \frac{x_2}{1 + \alpha I_2}$$

$$\dot{I}_1 = -k(y_1 - \beta \max(y_1, y_2))$$

$$\dot{I}_2 = -k(y_2 - \beta \max(y_1, y_2))$$

where $\beta < 1$, x is the feature input and y is the modulated feature value after shunting inhibition is applied. The attentional modulation signal I is obtained as the solution to the above pair of first order nonlinear differential equations.

In addition, we apply the following inequality constraints to $\dot{I}$:

$$\text{if } I = 0, \dot{I} \geq 0$$

$$\text{if } I = 1, \dot{I} \leq 0$$

These constraints implement the saturation of the temporal integrators. They restrict the phase space of the system $(I_1, I_2, \dot{I}_1, \dot{I}_2)$ to the region defined by $0 \leq I_1 \leq 1, 0 \leq I_2 \leq 1$.

It can be easily seen that there are no fixed points of the system in the interior of this region. To see this observe that in each of the two possible cases $(y_1 < y_2)$ and $(y_1 > y_2)$ either $\dot{I}_1$ or $\dot{I}_2$ is non-zero:

$$y_1 < y_2 : \dot{I}_2 = -ky_2(1 - \beta) < 0$$

$$y_1 > y_2 : \dot{I}_1 = -ky_1(1 - \beta) < 0$$

Furthermore, it can be shown that there are only two fixed points on the boundary of the constrained phase space region. These occur at $I_1 = 0, I_2 = 1$

and $I_1 = 1, I_2 = 0$. At the point $(I_1, I_2) = (1, 0)$ we have that

$$\dot{I}_1 = -kx_2 \left(\frac{R}{1+\alpha} - \beta \max \left(\frac{R}{1+\alpha}, 1 \right) \right)$$

$$\dot{I}_2 = -kx_2 \left(1 - \beta \max \left(\frac{R}{1+\alpha}, 1 \right) \right)$$

where $R = x_1/x_2$. If $R > 1 + \alpha$ we have that $\dot{I}_1 < 0$, and hence no fixed point exists. If $R < 1 + \alpha$ then

$$\dot{I}_1 = -kx_2 \left(\frac{R}{1+\alpha} - \beta \right)$$

This will be greater than or equal to zero when $R \leq \beta(1 + \alpha)$. If $\dot{I}_1$ is greater than zero when $I_1 = 1$, the saturation will keep $I_1 = 1$. Thus $\dot{I}_1 = 0$ when $R \leq \beta(1 + \alpha)$.

$$\dot{I}_2 = -kx_2(1 - \beta) \leq 0$$

The saturation at $I_2 = 0$ will force $\dot{I}_2$ to zero in this case. Thus the point $(I_1, I_2) = (0, 1)$ is a fixed point whenever $R \leq \beta(1 + \alpha)$. By symmetry, there will be a fixed point at $(I_1, I_2) = (1, 0)$ whenever $R \geq \beta(1 + \alpha)$.

Thus the behavior of the system is to sit at one of the two fixed points until the value of R changes to a value sufficient to flip the fixed points. The system will then move to the new fixed point.

In our model of saccade generation, saccades are triggered by "engagement" of attention onto the target. This is defined in our model to occur when the level of shunting inhibition at the target location reaches zero. In terms of the above notation, this occurs when $I_1 \to 0$. Here we take I_1 to be the target shunting inhibition and I_2 to be the fixation shunting inhibition. The saccadic latency is then, in our model, the time taken for I_1 to move from one to zero, after R is set to a value sufficient to cause switching.

The motion of I_1 from one to zero will, in general, consist of three phases. In the first phase, the motion is along the $I_2 = 0$ constraint line (where I_1 is fixed at one). The first phase may be absent if R is large enough. In the second phase the motion is in the interior of the configuration space. If this motion terminates on the $I_2 = 1$ line for $I_1 > 0$ there will be a third phase. If the system reaches the $I_1 = 0$ line before I_2 reaches one, there will be no third phase. The third phase consists of motion along the $I_2 = 1$ constraint line. Each of these motions will have its own time scale.

Let us consider the situation where $R > (1+\alpha)$. In this case, when $I_2 = 0$, we have that

$$\dot{I}_1 = -\frac{kx_2}{(1+\alpha I_1)} R(1-\beta)$$

which is always negative, and

$$\dot{I}_2 = -\frac{kx_2}{(1+\alpha I_1)} \left(\frac{(1+\alpha I_1)}{(1+\alpha I_2)} - \beta R \right)$$

which will be non-positive for $I_2 = 0$ (and hence keep I_2 on the lower saturation limit) when $I_1 \geq (\beta R - 1)/\alpha$. Thus, phase 1 ends when I_1 has discharged to the level $min(1, (R\beta - 1)/\alpha)$. If $R \geq (1+\alpha)/\beta$ then there will be no phase 1. During phase 2 the system moves through the interior of the configuration space. Whether phase 3 is present or not is dictated by the particular constraint line that was arrived at during phase 2. This will depend on the relative (dis)charging rates of I_1 and I_2 during this time. It can be seen from the above equations for $\dot{I}_1$ and $\dot{I}_2$ that when $R < (1+\alpha)/(2\beta - 1)$ $\dot{I}_1$ has a greater magnitude than $\dot{I}_2$. Hence, in this case, the constraint line at $I_1 = 0$ will be reached first. If $R > (1+\alpha)/(2\beta - 1)$ then the constraint line at $I_2 = 1$ will be reached first.

We see that are four cases to consider, if switching is to occur. The first case arises when $(1+\alpha)\beta < R < (1+\alpha)$. In this case phase 1 exists but there is no phase 3. The second case occurs when $(1+\alpha) < R < (1+\alpha)/\beta$, in which there is a phase 1 but no phase 3. The third case occurs when $(1+\alpha)/\beta < R < (1+\alpha)/(2\beta - 1)$ in which there is no phase 1 but there may be a phase 3. The fourth case arises when $R > (1+\alpha)/(2\beta - 1)$ where there is no phase 1 but there is a phase 3.

CASE 1: In this case $\dot{I}_2$ is small compared with $\dot{I}_1$, thus both phase 1 and phase 2 can be thought of as a motion along the $I_2 = 0$ constraint line. Thus we can combine the times in phase 1 and phase 2 to give:

$$T_{1,2} = \frac{I_1(T_{1,2}) - I_1(0)}{\dot{I}_1} \leq \frac{1+\alpha}{kx_2(R - \beta(1+\alpha))}$$

This is an upper bound obtained by setting $1 + \alpha I_1 = 1 + \alpha$. Note that this time grows without bound as R approaches $\beta(1+\alpha)$ from above.

CASE 2: Case 2 is similar to case 1, save that the formula for $\dot{I}_1$ is different. As in case 1, we can get an upper bound for the time spent in phases 1 and 2

by assuming that $1 + \alpha I_1 = 1 + \alpha$ during this time. Then:

$$T_{1,2} \leq \frac{1+\alpha}{kx_2(1-\beta)}$$

CASE 4: In case four, there is no phase 1, and the phase 2 trajectory will end up on the $I_2 = 1$ constraint line. We can get an upper bound for the phase 2 time by assuming that the phase 2 trajectory is along the $I_1 = 1$ line. The time is then

$$T_2 \leq \frac{1+\alpha}{kx_2(R\beta(1+\alpha)-1)}$$

The time taken in phase 3 can be bounded above by letting the term $1 + \alpha I_1$ that appears in the formula for $\dot{I}_1$ be fixed at $1 + \alpha$ during this phase,

$$T_3 \leq \frac{1+\alpha}{kx_2R(1-\beta)} = \frac{1+\alpha}{kx_1(1-\beta)}$$

Note that $T_3 >> T_2$ for typical values of α and β and that T_3 is independent of the losing feature value, x_2. Thus, only the winning feature value significantly affects the switching time in this case.

CASE 3: In case three there is no phase 1 and the phase 2 trajectory can end up either on the $I_1 = 0$ constraint line or the $I_2 = 1$ constraint line. Depending on which constraint line the trajectory ends up one can apply the bounds of either case 2 or case 4.

In each of the above cases, it can be seen that the switching time is proportional to the time constant $(1/k)$ of the shunting inhibition integrator. For large R the switching time depends primarily on the inverse of the winning feature value. For values of R close to $\beta(1+\alpha)$ the switching time is inversely proportional to $R - \beta(1-\alpha)$ and hence can be very large. For intermediate values of R, the switching time will depend on the values of both the winning and losing feature values.

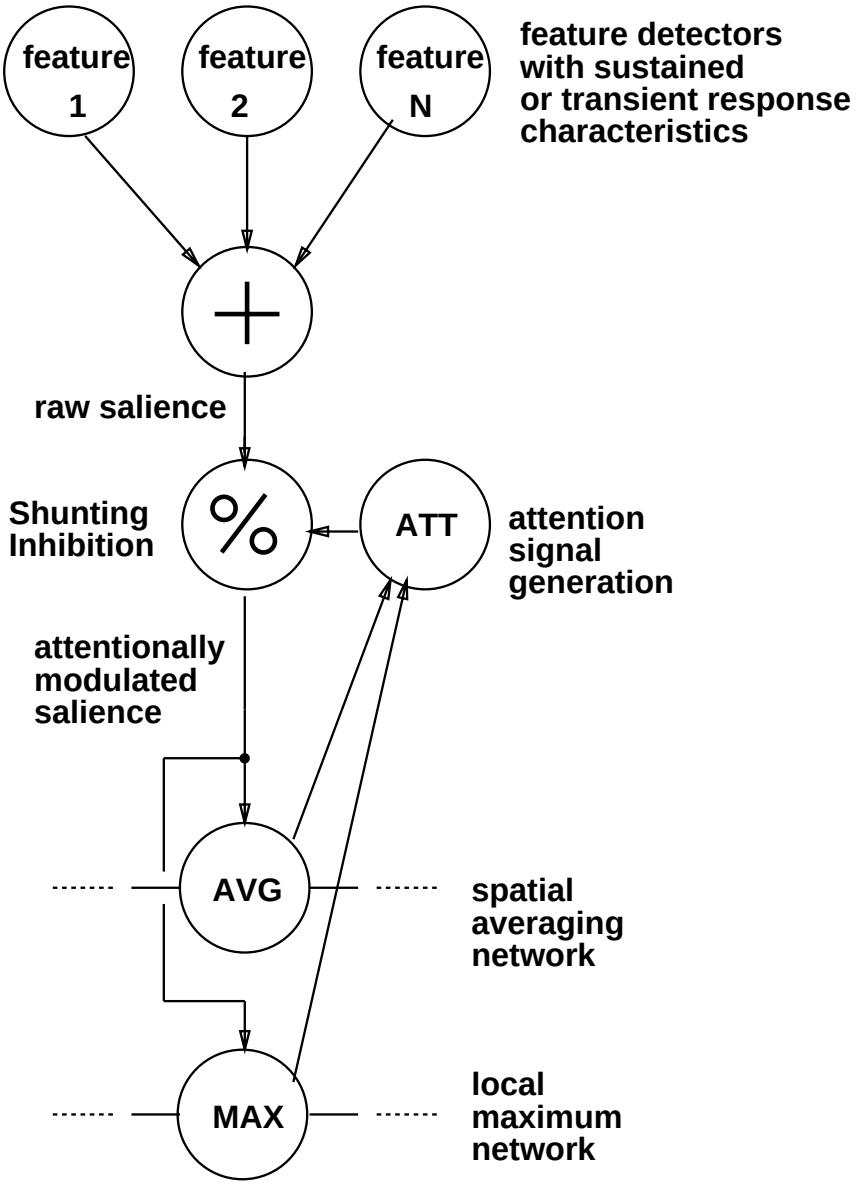

FIGURE 1. A schematic view of our neural network model of spatial visual attention.

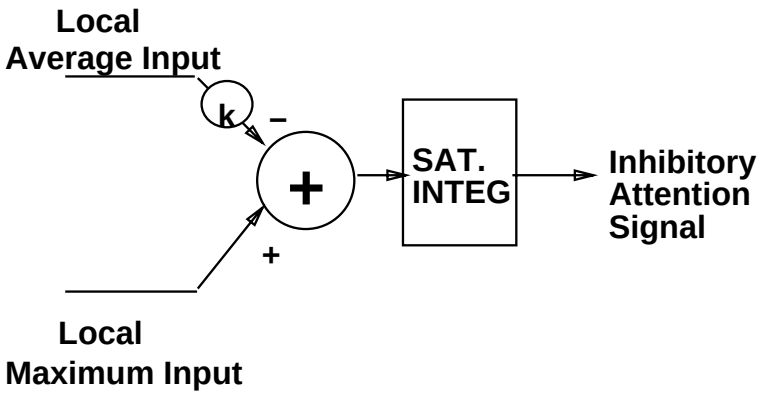

FIGURE 2. A model of the attentional modulation signal generation.

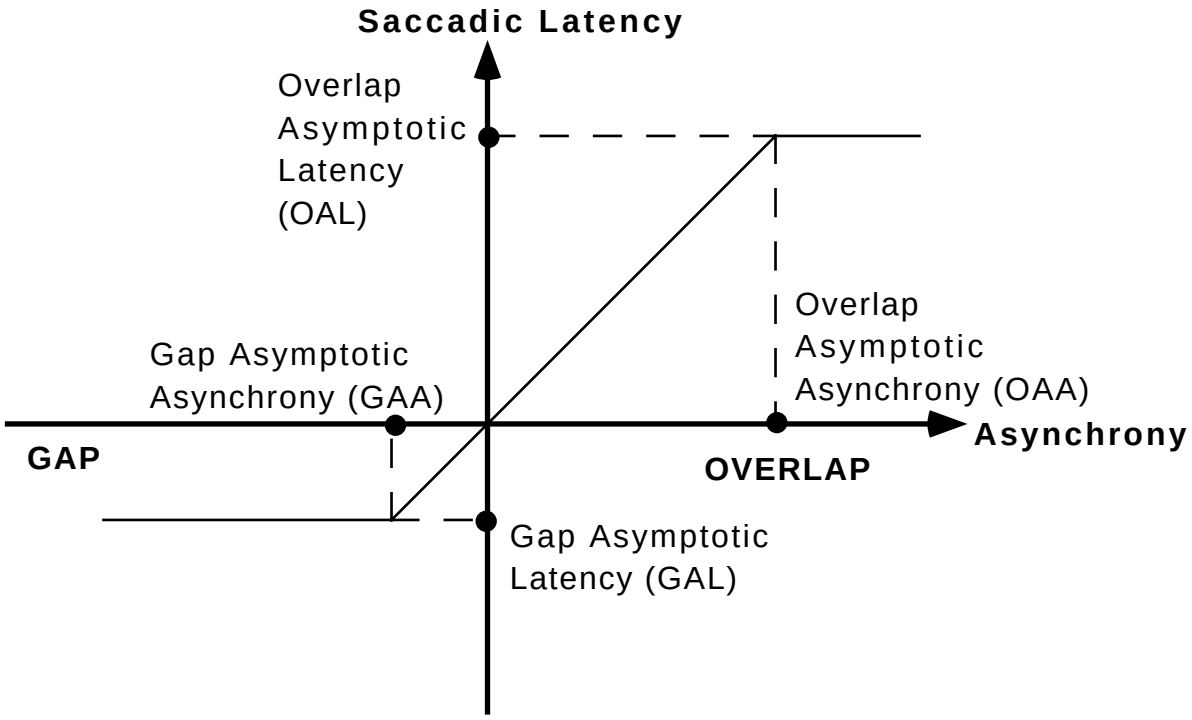

FIGURE 3. Reulen's piecewise linear model of the relationship between saccadic latency and stimulus asynchrony.

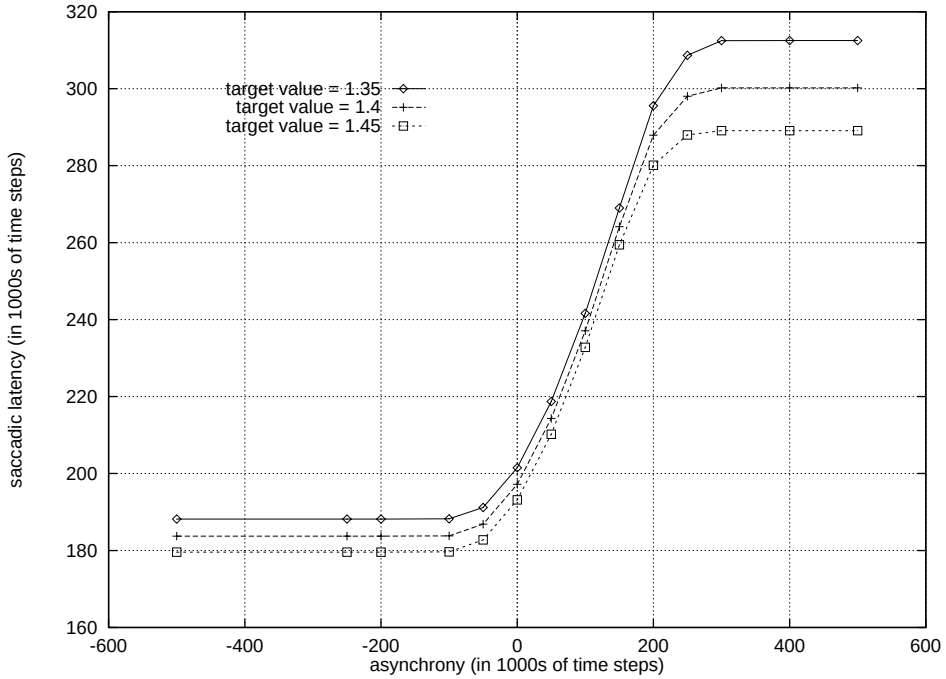

FIGURE 4. Simulations of the model exhibiting the gap/overlap effect for various values of the saliency of the target stimulus.

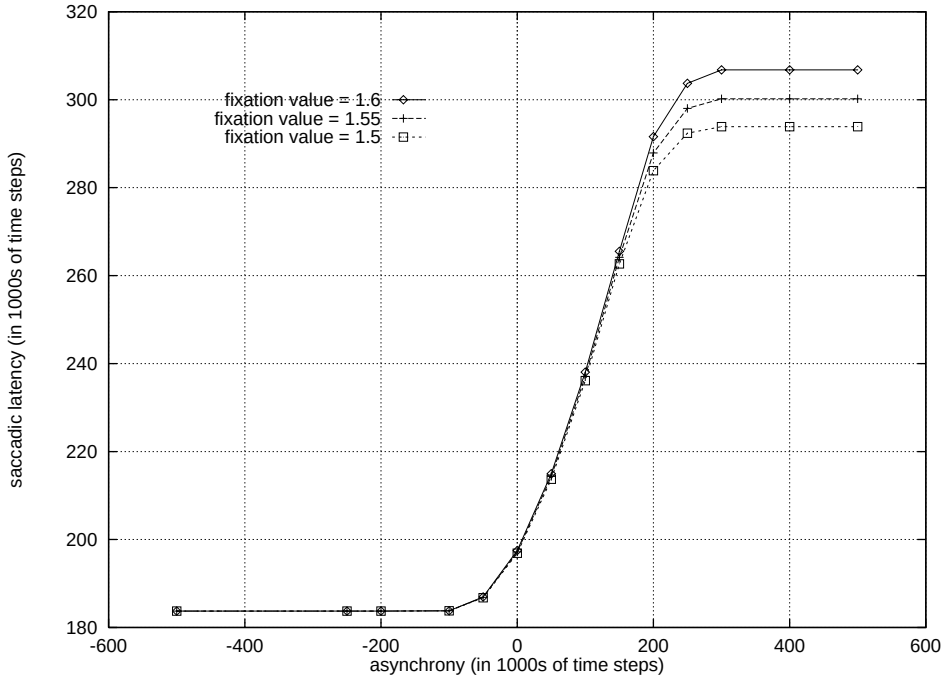

FIGURE 5. Simulations of the model exhibiting the gap/overlap effect for various values of the saliency of the fixation stimulus.

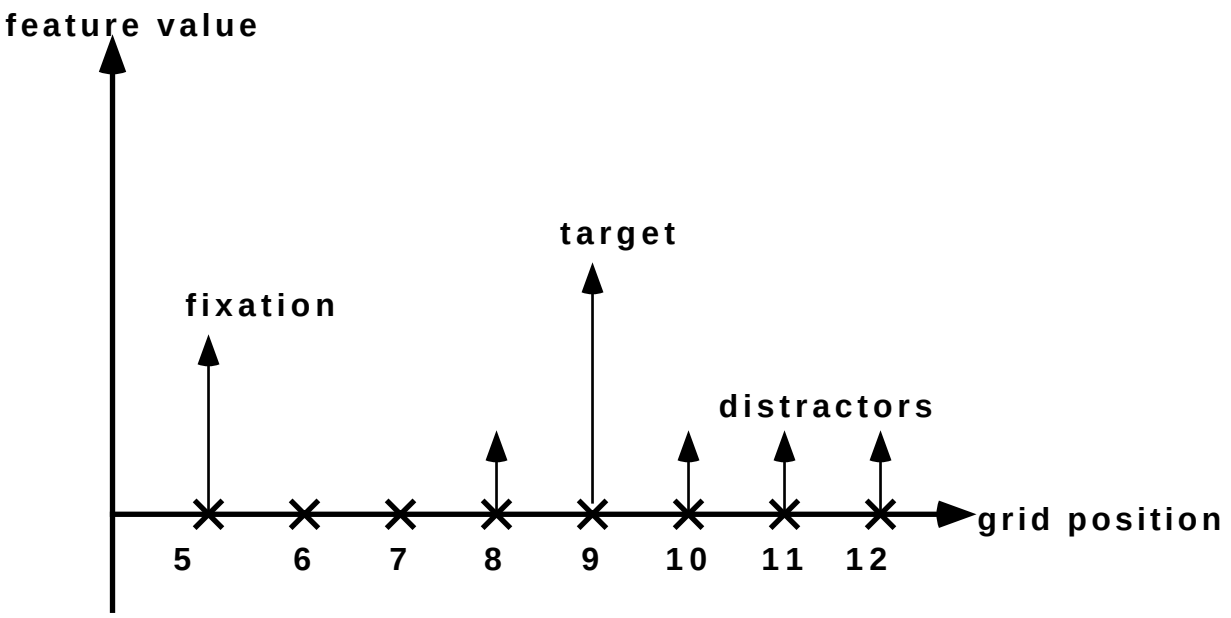

FIGURE 6. The configuration used in the simulation of the global effect.

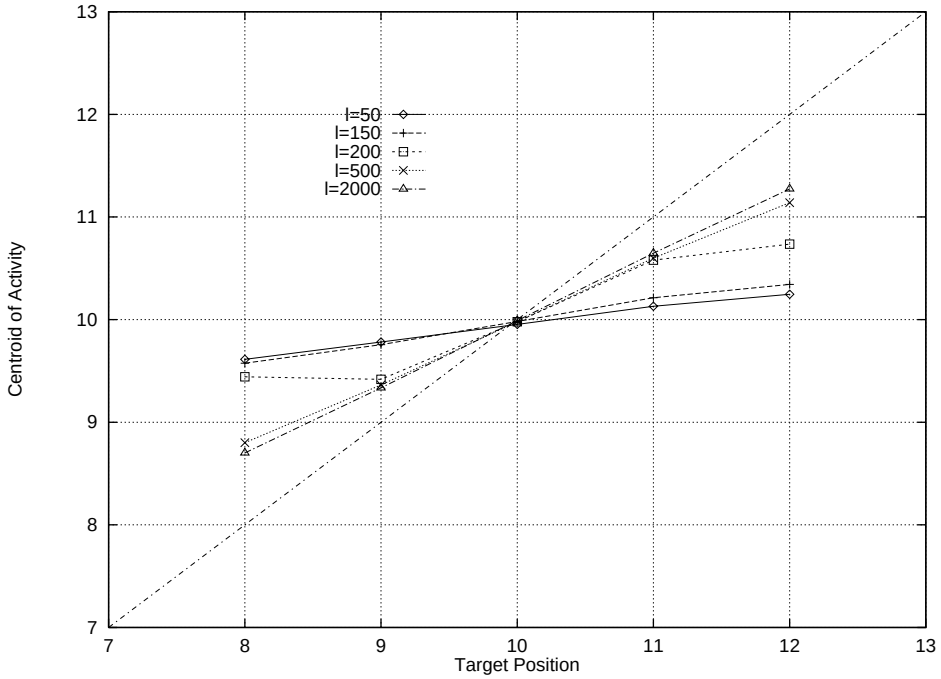

FIGURE 7. Simulations of the model responses to target and distractors. Each curve represents a different saccadic latency. Latency values are given in 1000's of time steps in the legend. The vertical axis is the centroid of the modulated feature activity, and is taken as the saccade command position. Increasing saccadic latency, l, is seen to reduce the global effect.

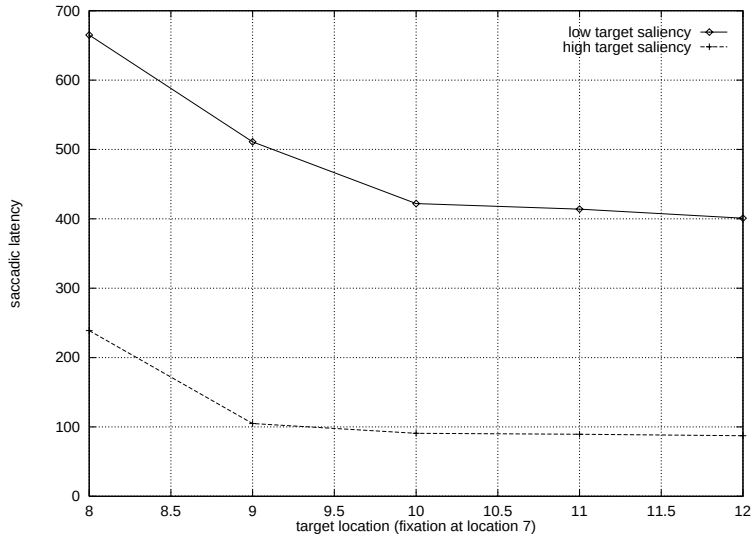

FIGURE 8. Saccadic latency as a function of the distance of the target stimulus from the fixation location. Curves for high and low target saliency are shown.

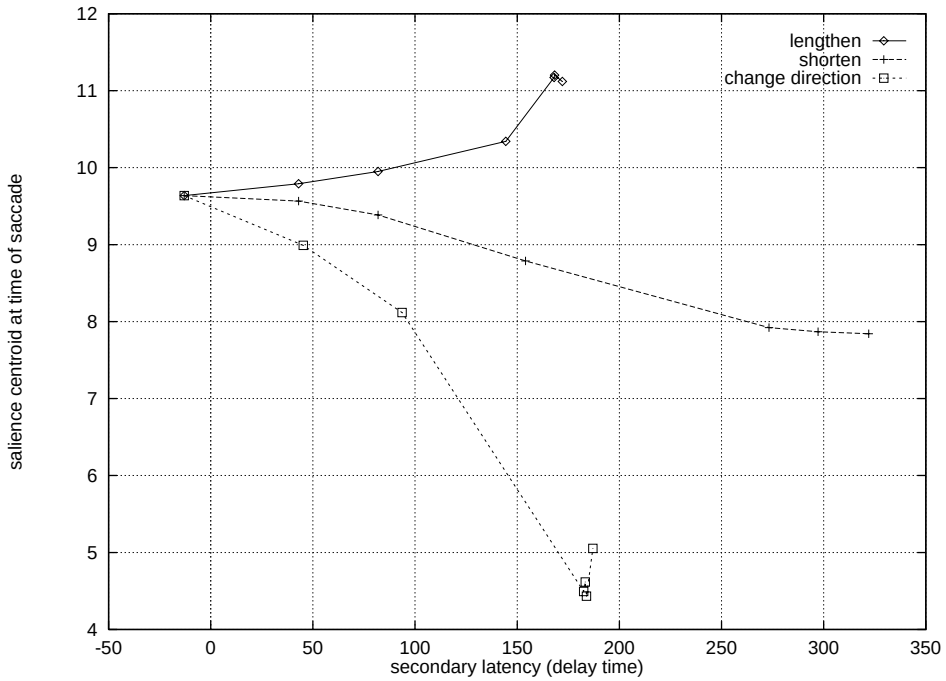

FIGURE 9. Results of the simulation of saccades to stepped targets. The centroid of the feature activity at the time of the saccade as a function of the secondary latency or delay time. Three types of movements of the target are shown - away from fixation (Lengthen), towards fixation (Shorten), and to the other side of the fixation (Change Direction).

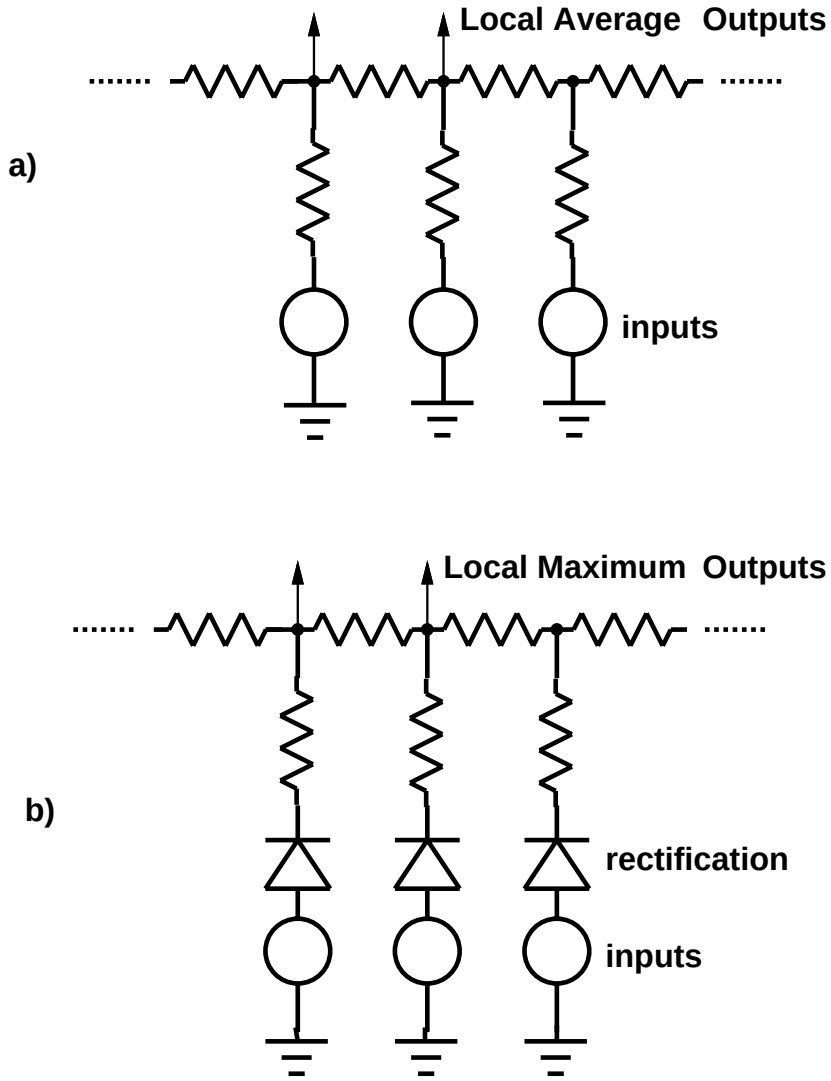

FIGURE A-1. a) A resistive network model of the local spatial average operation. b) A resistive network model of the local spatial maximum operation.

Figure Captions

FIGURE 1. A schematic view of our neural network model of spatial visual attention.

FIGURE 2. A model of the attentional modulation signal generation.

FIGURE 3. Reulen's piecewise linear model of the relationship between saccadic latency and stimulus asynchrony.

FIGURE 4. Simulations of the model exhibiting the gap/overlap effect for various values of the saliency of the target stimulus.

FIGURE 5. Simulations of the model exhibiting the gap/overlap effect for various values of the saliency of the fixation stimulus.

FIGURE 6. The configuration used in the simulation of the global effect.

FIGURE 7. Simulations of the model responses to target and distractors. Each curve represents a different saccadic latency. Latency values are given in 1000's of time steps in the legend. The vertical axis is the centroid of the modulated feature activity, and is taken as the saccade command position. Increasing saccadic latency, l, is seen to reduce the global effect.

FIGURE 8. Saccadic latency as a function of the distance of the target stimulus from the fixation location. Curves for high and low target saliency are shown.

FIGURE 9. Results of the simulation of saccades to stepped targets. The centroid of the feature activity at the time of the saccade as a function of the secondary latency or delay time. Three types of movements of the target are shown - away from fixation (Lengthen), towards fixation (Shorten), and to the other side of the fixation (Change Direction).

FIGURE A-1. a) A resistive network model of the local spatial average operation. b) A resistive network model of the local spatial maximum operation.